

分类号: TP391

密 级: 公开

学校代码: 10363

学 号: 2210330178



安徽工程大学

Anhui Polytechnic University

硕士学位论文

题 目 基于图像增强技术的无人驾驶
车道线检测

论 文 作 者 王静品

校 内 导 师 葛愿 教授

校 外 导 师 易建伟 高级工程师

专业学位类别 电子信息

研究方向 (领域) 人工智能与系统集成

论文提交日期: 2024 年 5 月 31 日

分类号：TP391

单位代码：10363

密 级：公开

学 号：2210330178

题 目 基于图像增强技术的无人驾驶车道线检测

题 目 DRIVERLESS LANE DETECTION BASED ON
IMAGE ENHANCEMENT

学 生 姓 名 : 王静品

校 内 导 师 : 葛愿 教授

校 外 导 师 : 易建伟 高级工程师

专 业 (领 域) : 电子信息

研究方向 (领域) : 人工智能与系统集成

论文答辩日期： 2024年5月19日

安徽工程大学学位论文原创性声明

本人郑重声明：我恪守学术道德，崇尚严谨学风。所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已明确注明和引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品及成果的内容。论文为本人亲自撰写，我对所写的内容负责，并完全意识到本声明的法律后果由本人承担。

学位论文作者签名：王静品

日期：2024年5月31日

安徽工程大学学位论文版权使用授权书

学位论文作者完全了解学校有关保留、使用学位论文的规定，同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅或借阅。本人授权安徽工程大学可以将本学位论文的全部内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。

保密 ☐，在 ____ 年解密后适用本版权书。

本学位论文属于

不保密 ☒。

学位论文作者签名：王静品

指导教师签名：李俊

日期：2024年5月31日

日期：2024年5月31日

基于图像增强技术的无人驾驶车道线检测

摘 要

恶劣环境会严重影响车道线检测算法的性能，夜间场景下车道线图像亮度较低且不均匀；雾天环境下光线散射，车道线特征模糊；雨夜环境下雨条纹会遮掩车道线信息，针对三种场景设计三种图像增强网络对恶劣环境下的车道线图像进行增强，然后使用改进的车道线检测网络对增强后的图像进行车道线检测。本文从提升车道线检测效果的角度出发，针对恶劣环境下的图像增强和车道线检测进行研究，主要研究内容如下：

(一)、夜间图像增强及车道线检测的研究。针对夜间场景下车道线图像整体较暗、光照不均匀，直接导致检测准确度低，漏检率和误检率较高的问题，设计一个基于 Retinex 理论的图像增强网络。首先在 Retinex 理论的基础上设计反射率估计网络，通过在编码层使用注意力机制提升网络对夜间低光照区域的关注；其次引入深度可分离卷积提升网络的运行速度；最后在(Efficient Residual Factorized ConvNet, ERF-Net)网络的编码层引入自注意力蒸馏，提升网络对车道线信息的认知能力。实验结果表明改进的夜间图像增强网络可以在保留图像颜色信息的基础上提升低光照部分的亮度，并且不破坏图像的细节信息，车道线检测网络在保证实时性的同时提升了检测精度。

(二)、雾天图像增强及车道线检测的研究。为了解决雾天环境

下，车道线受到雾气的遮挡，检测网络精度降低，设计一个基于 CycleGAN 的无监督型去雾网络。首先在循环生成对抗网络的生成器中引入多层特征融合模块提升不同编码层特征信息的利用率，通过重新设计多个残差块结构以降低网络的计算量，使网络可以适用于无人驾驶场景；此外使用判别器从全局和局部两个角度分别判断生成的图像，使网络可以达到更好的博弈效果。实验结果表明改进的去雾网络可以有效的去除不同浓度的雾，同时增强的图像可以提升车道线检测网络的检测效果。

(三)、雨夜环境下的图像增强及车道线检测研究。针对夜间雨天环境下图像整体亮度较低，去雨网络无法去除隐藏在黑暗中的雨条纹，车道线信息会受到复杂光照条件的影响，同时雨条纹会遮挡车道线的问题，首先设计一个端到端的夜间雨天图像增强网络，通过使用(Progressive Image Deraining Networks, PReNet)网络和坐标注意力预先检测夜间图像中雨条纹的位置，然后使用图像恢复网络根据学习到的雨条纹先验知识对图像进行增强。实验结果表明端到端的增强网络可以去除不同光照下的雨条纹，同时可以提升车道线检测网络的检测效果。

关键词：图像增强，车道线检测，Retinex 理论，循环生成对抗网络，夜间雨条纹位置预测

DRIVERLESS LANE DETECTION BASED ON IMAGE ENHANCEMENT

ABSTRACT

Harsh environments can seriously affect the performance of lane line detection algorithms, with low and uneven brightness of lane line images in nighttime scenarios, light scattering and blurred lane line features in foggy environments, and rain streaks obscuring lane line information in rainy environments. Three image enhancement networks are designed for the three scenarios to enhance lane line images in harsh environments. Then, the enhanced images are used to perform lane line detection using the improved lane line detection network. Lane line detection using an improved lane line detection network. In this paper, image enhancement and lane line detection in harsh environments are studied from the point of view of improving the effectiveness of lane line detection, and the main research contents are as follows:

(1) A study of nighttime image enhancement and lane line detection. An image enhancement network based on Retinex theory is designed to address the problem that the overall darkness and uneven illumination of the images of off-road lanes in nighttime scenes directly lead to low detection accuracy and high leakage and false detection rates. Firstly, the reflectance estimation network is designed based on Retinex theory, which improves the network's attention to the low-light region at night by using the attention mechanism in the coding layer; secondly, the depth-separable convolution is introduced to improve the network's operation

speed; and lastly, the self-attention distillation is introduced in the coding layer of the ERF-Net network, which improves the network's cognitive ability to the information of the lane lines. The experimental results show that the improved nighttime image enhancement network can enhance the brightness of the low-light part of the image while preserving the color information of the image and without destroying the detailed information of the image. The lane line detection network improves the detection accuracy while ensuring real-time performance.

(2) A study of foggy image enhancement and lane line detection. To solve the foggy environment, lane lines are obscured by fog and the accuracy of the detection network is reduced. An unsupervised dehazing network based on CycleGAN is designed. Firstly, a multi-layer feature fusion module is introduced into the generator of the cycle generative adversarial network to improve the utilization of feature information from different coding layers, and the network can be applied to scenarios by re-designing the structure of multiple residual blocks in order to reduce the computational amount of the network; in addition, a discriminator is used to judge the generated images from global and local perspectives respectively, so that the network can achieve a better gaming effect. The experimental results show that the improved dehazing network can effectively remove different concentrations of fog. At the same time, the enhanced image can improve the detection effect of the lane line detection network.

(3) A study of image enhancement and lane line detection in rainy nighttime environments. For the nighttime rainy environment where the overall brightness of the image is low, the rain removal network is unable to remove the rain streaks that are hidden in the darkness, and the lane line information will be affected by the complex lighting conditions, while the rain streaks will partially obscure the lane lines. An end-to-end

nighttime rainy day image enhancement network is first designed to pre-detect the location of rain streaks in nighttime images using a PReNet network and coordinate attention. Then, an image recovery network is used to enhance the images based on the learned a priori knowledge of rain streaks. The experimental results show that the end-to-end augmented network can remove the rain streaks under different lighting and simultaneously improve the detection of the lane line detection network.

Jingpin Wang(Electronic Information)

Supervised by Prof.Yuan Ge

KEY WORDS: image enhancements, lane line detection, retinex theory, cycle generative adversarial network, nighttime rain streak location forecast

目录

摘 要.....	I
ABSTRACT.....	III
第 1 章 绪论.....	1
1.1 研究背景和意义.....	1
1.2 国内外研究现状.....	2
1.2.1 夜间图像增强研究现状.....	2
1.2.2 图像去雾技术研究现状.....	4
1.2.3 图像去雨技术研究现状.....	5
1.2.4 车道线检测技术研究现状.....	7
1.3 论文主要研究内容及创新点.....	8
1.4 本文章节安排.....	9
第 2 章 基于 Retinex 理论和自注意力蒸馏的夜间车道线检测	11
2.1 引言.....	11
2.2 基于改进 Retinex 理论的夜间图像增强	12
2.2.1 Retinex 理论	12
2.2.2 RGB 和 HSV 颜色空间.....	13
2.2.3 夜间图像增强网络.....	16
2.2.4 损失函数.....	20
2.3 基于自注意力蒸馏的车道线检测网络.....	21
2.3.1 ERF-Net 网络	21
2.3.2 车道线检测网络.....	24
2.3.3 损失函数.....	25
2.4 实验结果分析.....	27
2.4.1 实验设置.....	27
2.4.2 评价指标.....	28

2.4.3 夜间图像增强结果分析.....	30
2.4.4 车道线检测结果分析.....	33
2.5 本章小结.....	36
第 3 章 基于多层特征融合的雾天车道线检测.....	37
3.1 引言.....	37
3.2 大气散射模型.....	38
3.3 循环生成对抗网络.....	39
3.4 FCA 注意网络.....	40
3.5 改进的 CycleGAN 去雾网络.....	41
3.5.1 多层特征融合模块.....	42
3.5.2 轻量化的残差模块.....	43
3.5.3 全局判别器和局部判别器.....	44
3.5.4 损失函数.....	45
3.6 实验结果及分析.....	47
3.6.1 实验设置.....	48
3.6.2 合成雾天数据集上的实验结果分析.....	49
3.6.3 真实雾天数据集上的实验结果分析.....	52
3.6.4 消融实验.....	54
3.6.5 车道线检测结果分析.....	55
3.7 本章小结.....	56
第 4 章 基于 Uformer 的雨夜车道线检测.....	57
4.1 引言.....	57
4.2 Uformer 网络.....	58
4.3 PReNet.....	61
4.4 夜间雨天图像增强网络.....	63
4.4.1 夜间雨条纹位置检测模块.....	64
4.4.2 损失函数.....	66
4.5 实验结果分析.....	67
4.5.1 实验设置.....	67

4.5.2 合成的夜间雨天数据集上的实验结果分析.....	68
4.5.3 真实的夜间雨天数据集上的实验结果分析.....	70
4.5.4 夜间雨天数据集上的车道线检测.....	71
4.6 本章小结.....	72
第 5 章 总结与展望.....	73
5.1 总结.....	73
5.2 展望.....	74
参考文献.....	75
攻读学位期间发表的学术论文.....	83
致谢.....	84

第 1 章 绪 论

1.1 研究背景和意义

国家双碳目标推动了更多厂商加大对新能源汽车的投入，奇瑞、比亚迪等知名厂商纷纷推出更加智能的新能源汽车^[1]。随着汽车产量的增加，它在给人们通行方式提供便捷的同时，也带来了一系列的交通安全问题，如交通堵塞、交通事故等^[2]。根据国家统计局公布的数据，2022 年共发生交通事故 256409 起，因为汽车交通事故死亡人数达到 42012 人，受伤人数达到 149650 人。因此可以得出结论，道路交通事故已经演变成了对人们生命和财产安全构成重大威胁的严重问题。在道路交通事故的报告中显示有 90% 的交通事故是由于驾驶员自身的主观过失所引发，这种主观过失通常表现为驾驶者的注意力分散，驾驶员无法及时获取路况信息并采取正确反应，最终导致交通事故的发生^[3]。

以上种种情况推动着自动驾驶技术的发展。传统的安全装置，如安全气囊、防抱死系统、安全带等均是采用在事故发生后才会触发，无法起到预防交通事故的作用。与传统的安全装置不同，自动驾驶技术具备主动性，能够通过采取一系列的措施来预防交通事故的发生。根据驾驶者对车辆的掌控程度将自动驾驶技术分为 6 级， $L_0 \sim L_3$ 等级是驾驶系统辅助人类进行驾驶，一般称其为高级驾驶员辅助系统； L_4 等级是真正意义上的自动驾驶，是人类辅助驾驶系统进行驾驶； L_5 等级是汽车完全由驾驶系统控制，即无人驾驶^[4]。高级辅助驾驶系统可分为车辆控制和车辆预警两部分，车辆预警系统包括车道偏离和驾驶员疲劳预警，而车辆控制系统则包括车道保持和自动泊车等功能。车道偏离预警和车道保持系统的有效性取决于对车道线的准确检测^[5]。

车道线检测作为自动驾驶的重要组成部分，具有多重重要意义。首先，车道线检测能够帮助车辆准确识别和跟踪车道，从而确保车辆在道路上保持稳定的行驶轨迹，这可以避免车辆偏离车道、发生违规行驶或者与其他车辆发生碰撞的情况，提高行车安全。其次，车道线检测还能够提高驾驶效率，通过实时监测车道线，车辆可以自动进行车道保持，以此提升驾驶者的驾驶效率。近年

来,很多学者对车道线的检测方法进行研究,已经提出了很多有效的方法去检测车道线,且在正常白天环境下可以达到较高的准确度,但是这些方法对于恶劣环境下的车道线检测效果较差。通常道路行驶会遇到的恶劣环境有三种,分别是夜间、雾天和雨天。相比于视野正常的白天环境,夜间和雨雾等恶劣环境更容易发生交通事故^[6]。夜间驾驶环境下,驾驶员的视野会受到限制,环境光较弱并且受到对向车辆的灯光影响,这将导致车道线检测出现漏检、错检等情况,使车辆更易偏离车道。雾天环境下,能见度较低,车载视觉传感器所捕获的路况图像通常存在图像质量降低和细节丢失的问题,导致一些关键车道线特征受到雾霾等影响而难以准确检测。雨天环境下,受到雨水附着的影响,镜头会模糊不清,图像会变形,导致拍摄到的图像不能被正常检测。通过图像增强可以提升恶劣环境下采集到的图像质量,以此对后续的高级视觉任务起到积极作用。因此,本文首先针对恶劣环境图像的增强方法和车道线检测方法进行研究,设计图像增强网络对低质量图像进行增强,然后使用车道线检测网络进行车道线检测。

1.2 国内外研究现状

1.2.1 夜间图像增强研究现状

白天的光照充足,车道线特征和背景有较高的对比度,驾驶员可以准确地根据车道线的形状进行判断和决策,即使在意外事件发生时也可以快速的做出反应。然而夜间环境下光照不足,整体环境相对较暗,驾驶员的视野受限,难以清晰地辨认车道线特征,导致驾驶的可视范围变窄。此外夜间道路上常常存在干扰光,包括对向车辆灯光的干扰、路边灯光的影响等,这些干扰光源会使驾驶员的注意力分散,降低其对道路情况的感知能力。为了增加夜间行车的安全性,需要提升夜间车道线的检测精度。夜间图像增强技术的应用可以突出车道线特征并降低图像中的噪声,从而提高对车道线的检测精度。通过优化夜间道路图像的质量,能够使车道线特征更加显著,为后续图像处理算法提供更准确的特征信息,进一步改善车道线检测的效果。

对于低照度的图像增强研究可以分为两个阶段,首先是以传统图像处理算

法为主的增强方式，通常是以直方图均衡化(Histogram Equalization, HE)和 Retinex 理论为基础进行改进。最初的 HE 通过重新分配图像的灰度级别来实现更均匀的灰度分布，从而提升整体亮度。尽管这种方法可以增强图像的亮度，但由于其强制合并灰度级别，可能导致图像色彩失真，并在局部区域产生过度增强的现象。为了应对上述现象，Reza.A 等人^[7]通过将图像分割成许多小块，在每个小块上应用 HE，并使用插值方法将这些小块拼接在一起，通过引入对比度限制，避免过度增强噪声的问题。然而，当处理具有较大范围亮度变化的图像时，该方法可能会产生一些伪影现象。涂毅晗等人^[8]首先将低照度图像转换到 HIS 颜色空间，对光照分量使用 3 个尺度进行分块，对每个块使用 HE，对反射分量使用多方向梯度增强，通过组合后重新转换回 RGB 颜色空间以实现夜间图像增强。赵丽等研究者^[9]运用梯度信息进行引导滤波处理图像，通过将 HE 与梯度域引导滤波相结合，实现了在增强图像亮度的同时保持图像细节清晰的效果。Retinex 理论的变体算法也经常被应用于夜间图像增强。Retinex 理论由 E.Land 提出^[10]，该理论通过对亮度分量的增强来改善图像的视觉效果。以此为核心思想扩展出许多算法，如带有颜色恢复的多尺度 Retinex(Multi-Scale Retinex with Color Restoration, MSRRCR)算法^[11]。传统图像处理算法普遍存在一个缺点，即耗时时间较长且对于夜间光照不均匀的区域恢复效果较差。

目前已经有很多学者将深度学习应用到低光照图像增强领域，2017 年 Lore^[12]使用两阶段架构进行增强，一个阶段进行对比度增强，另一个阶段进行去噪。同年 Shen 等人^[13]第一次将 MSR 理论和卷积神经网络(Convolutional Neural Networks, CNN)相结合，提出了 MSR-Net 算法，利用 CNN 网络端到端的学习低光照图像到正常亮度图像的映射。Wei 等人^[14]提出了 RetinexNet 网络，RetinexNet 网络由图像分解、调整 and 重建网络三个部分组成。首先，图像分解网络将低光照图像分解为反射分量和光照分量，然后图像调整网络对光照分量进行调整，并对反射分量进行噪声抑制，最后图像重建网络将经过处理的反射分量和改进的光照分量重新组合，恢复成具有正常光照的图像。由于成对出现的场景图像数据集较难获取，且大部分是合成图像，不足以适用真实夜间场景，因此很多研究者将研究方向转向无监督型网络。生成对抗网络(Generative Adversarial Networks, GAN)^[15]在低级视觉处理方面表现出卓越的性能，许多研

研究者开始采用 GAN 作为基础网络。Jiang 等人^[16]通过引入自正则化感知损失有效地控制低光照输入图像与增强后图像之间的特征距离，同时使用全局判别器和局部判别器从全局和局部两个层面进行增强。吴佳奇等人^[17]设计了一种融合空间注意力和通道注意力的双分支估计网络，通过在两个分支网络之间引入多尺度特征分解机制，可以有效提高图像的亮度且避免颜色失真的发生。Anoosheh 等人^[18]设计了一个图像翻译模型将图像从夜间风格转换到白天风格。除此之外，Guo 等人^[19]提出了(Zero-Reference Deep Curve Estimation, Zero-DCE)网络，该网络也是无监督学习方法，并且不同于常用的 CNN 和 GAN 网络，其使用神经网络拟合一条曲线将不同亮度的像素点映射到一个新的亮度上。

1.2.2 图像去雾技术研究现状

在雾天环境下，图像中的车道线模糊不清，进而对车道线检测系统的性能产生负面影响。目前对于去雾方法的研究主要分为两个阶段，第一个阶段是使用物理模型进行去雾。基于物理模型的去雾方法分析了造成图像质量变低的原因，通过建立模型求解得到去雾图像。对于去雾模型来说，常用的是 Nayer 等人^[20]提出的大气散射模型，通过计算模型中的透射率和大气光值进行求解。经典的算法是 He 等人^[21]提出的暗通道先验理论，该理论认为在无雾图像中，每个像素的 RGB 三个颜色通道中必定有一个通道的灰度值较低，基于此可以求出透射率。对于大气光值，取图像最亮像素的前 10% 得到近似值，然而这样会导致在处理和大气光值相似的场景时，去雾算法失效。Berman 等人^[22]提出了雾线理论进行去雾，使用该理论可以恢复无雾图像，但是当雾的浓度很高时，该方法的去雾效果将降低。高涛等人^[23]提出了将亮通道和暗通道结合进行去雾。即使部分基于先验知识的去雾算法实现起来较为简单，但是这些算法普遍耗时严重，对有雾图像的要求较高，如果所需去雾图像与先验知识冲突，则会导致去雾效果严重下降。为了克服上述问题，深度学习技术被应用到图像去雾领域。

随着深度学习的发展，很多研究者都采用深度学习的方式从数据层面进行去雾^{[24][25]}。初期的神经网络去雾基本都是和大气散射模型联合使用的，许多算法通过设计网络求出有雾图像的透射率，然后再利用大气散射模型恢复无雾图像。例如，Cai 等人^[26]使用 DehazeNet 估算出雾天图像的透射率，然后通过大气

散射模型计算出无雾图像。Ren 等人^[27]将多尺度的 CNN 网络用到透射率的计算上，从粗尺度网络获得每个图像传输图的粗略结构，再通过细尺度网络细化。但是当传输图计算不准确时，去雾结果会受到严重的影响。Li 等人^[28]将大气散射模型重新进行设计，通过监督学习的方式端到端的生成去雾结果。以上这些算法是通过计算中间参数进行去雾的，在没有约束条件的情况下会产生很大的误差，这将导致网络输出的无雾图像产生颜色失真等情况。Chen 等人^[29]通过门控网络融合不同层次的特征，然后解码特征映射得到残差图，最后将残差图和雾图相加得到去雾图像。Qin 等人^[30]提出了 (Feature Fusion Attention Network, FFA-Net) 网络，该网络将特征注意模块和跳跃连接结合成块结构，通过多次调用块结构和组结构进行去雾，此方法在端到端的去雾中取得了较为可观的效果。

鉴于数据获取难度较大，且对真实世界的雾天没有泛化能力，基于 GAN 的无监督去雾方法便油然而生。Engin^[31]基于 (Cycle Generative Adversarial Network, CycleGAN) 设计了 Cycle-Dehaze 网络，主要是针对损失函数进行优化，将循环感知损失引入损失函数中，同时改进拉普拉斯金字塔，得到高分辨率的去雾图像。Sun^[32]将迭代去雾模型嵌入 CycleGAN 的生成器中，从物理模型获得细节信息，并设计一致性损失去除雾霾。张成等人^[33]利用暗通道先验将图像分解输出透射率和去雾图像，再使用判别器对输出进行处理和判断，从而取得去雾结果。但志平^[34]提出了一种基于双判别器异构 CycleGAN 的多通道注意力去雾算法，该网络构建了多阶通道注意力模块，其中整合了一阶和二阶特征统计量以优化图像处理效果。Chaitanya 等人^[35]将 (All-in-One Dehazing Network, AOD-Net) 嵌入到编码器和解码器中间，并将 (Structural Similarity Index, SSIM) 损失加入到损失函数中以提高去雾质量。

1.2.3 图像去雨技术研究现状

在雨天环境下，车道线检测受到多重干扰，包括雨水遮挡和反射光线干扰，这种情况加剧了车道线检测算法的复杂性，并可能导致车道线轮廓模糊不清、检测精度降低等问题。许多研究者为了有效应对雨天环境对车道线检测的影响，采用多种去雨算法对图像进行处理。

按照时间顺序，去雨算法分为两大类：一类是早期依赖于稀疏编码和字典

学习的方案；另一类则是运用深度学习技术的现代去雨算法。采用稀疏编码技术进行去雨时，能够将含有雨水的输入图像分解为一组基向量的稀疏线性组合形式，进而有效地还原出无雨状态下的清晰图像。Lin 等人^[36]运用形态成分分析将有雨图像分解为雨线层和背景层，再通过字典学习和稀疏编码区分并去除雨线部分。汤红忠等人^[37]通过双重字典对高频图像进行稀疏表达，有效区分并剔除了高频图像中的雨条纹部分，继而将提取出的高频无雨信息与低频图像结合，实现图像去雨。

基于模型的方法通过在雨层和背景层上增加先验来限制去雨过程。然而，在真实自然条件下，雨纹层和背景层结构复杂，难以完全分离。为了适应更多的雨天场景，研究者们使用基于深度学习的方式对雨条纹进行去除。Fu 等人^[38]首先将雨天图像分解为低频层和高频雨纹层，将高频层输入残差网络中处理，通过结合增强的高频输出和低频背景层来恢复图像。Yang 等人^[39]提出多任务学习网络，利用雨纹掩模定位有雨区域并借助空洞卷积网络进行雨条纹检测。针对雨滴密集分布及其引起的多方向雨纹问题，该模型采用了循环的方式渐进的去雨条纹，以优化复杂雨景下的处理性能。Zhang 等人^[40]将网络分为多流密集连接网络和残差感知密度分类模块，残差感知密度分类模块可以自动确定雨的密度信息，多流密集网络根据估计出来的雨密度标签去除相应的雨条纹。尽管该网络可以去除各种大小的雨条纹，但在实际应用中，由于难以准确衡量雨条纹的大小，导致去除效果较差。Ren 等人^[41]提出了多阶段渐进式去雨网络，通过使用递归机制来融合不同阶段的特征信息，该方法虽然可以去除部分图像中的雨条纹，但是恢复的图像亮度较低。Deng 等人^[42]提出了一个端到端的细节恢复去雨网络，该网络分为两个子网络，雨条纹残差网络学习从雨天图像到雨条纹的映射关系，细节修复网络修复去雨时损失的细节信息，最后将输入雨图像减去雨条纹图像，再加上细节修复网络学习到的信息得到最终的去雨图像。Zamir 等人^[43]将图像的恢复任务分解为几个子任务，利用多阶段架构进行图像修复，该网络采用编码器-解码器结构，在每个阶段通过逐像素自适应调整权重，以生成无雨图像。王美华等人^[44]提出一种自适应选择卷积去雨网络，该网络通过构建选择卷积网络选择不同卷积核对应的维度信息进行学习，最后通过残差修正网络直接学习雨条纹图像和残差修正系数，减少了对背景的误判。Wang 等

人^[45]提出了卷积字典去雨网络(Rain Convolutional Dictionary Network, RCD-Net),使用图像内在特征对雨的形状进行编码,通过多阶段迭代进行去雨。Wang 等人^[46]设计了 Uformer 网络,该网络运用非重叠窗口自注意力机制减少计算成本,并通过深度卷积增强上下文捕捉能力,同时该网络使用了跳跃连接,使编码器的信息可以传递到解码器中。

1.2.4 车道线检测技术研究现状

车道线检测的目的在于自动检测并跟踪道路图像中的车道线。借助对摄像头或传感器数据的分析,车道线检测算法能够有效识别车道线的准确位置和几何特征,并为其生成对应的标记。在无人驾驶系统中,车道线检测是关键技术之一。车道线检测为车辆提供了重要的环境感知信息,使其能够预测车道变化、其他车辆的行驶轨迹和动作,从而增强驾驶决策能力,提高驾驶的安全性。对于车道线检测的研究最初是使用显性特征进行车道线检测。根据车道线的颜色、边缘等特征信息使用图像处理和主动特征提取等方法进行车道线检测。一般车道线在被车载视觉传感器采集到的图像中多处于图像的下半部分,因此早期的研究者一般将图像的下 2/3 的区域划分为感兴趣区域。然后进行特征信息的提取,以便更好地识别车道线,最后采用车道线模型拟合的方法,实现对车道线的检测。Cai 等人^[47]同时考虑车道线的颜色特征和车辆的位置信息进行车道线检测。宋琴琴等人^[48]采用颜色阈值分割技术进行车道线检测,首先用高斯滤波预处理降噪,继而将 RGB 图像转换到 HSV 色彩空间,设定颜色阈值以识别车道线,最后在潜在车道线区域内运用 Canny 算子检测边缘。化春键等人^[49]通过对 Lab 空间中 b 分量和灰度图像加权增强对比度,利用掩膜寻找连续边缘并通过改进霍夫变换及最小二乘法检测车道线。对于以上使用车道线特征进行检测的算法来说,大部分算法只适用于车道线比较清晰的图像,但是实际驾驶环境比较复杂,车道线会出现磨损等情况,该类算法没有较好的鲁棒性,因此很多研究者开始使用深度学习的方式进行车道线检测。

初期的网络算法设计是将 CNN 和传统算法结合检测车道线。Kim 等人^[50]首先提取出图像中的边缘特征,然后使用 CNN 和(Random Sample Consensus, RANSAC)算法结合的方法来检测车道线。He 等人^[51]首先使用逆透视变换将图

像从正视图转换到俯视图，然后使用 RANSAC 算法提取出待选的车道线，重新将角度转换回正视图，获取正视图中的待选区域，将两个待选区域输入到 CNN 网络中进行训练，最终获得检测出的车道线。Neven 等人^[52]首先使用编码器将图像中的每个像素标记为车道线或背景，然后编码器中的另一分支将特征嵌入到空间结构中，对编码器输出的特征图进行融合和聚类，以获取更加精确的车道线。Pan 等人^[53]设计了空间卷积神经网络(Spatial CNN, SCNN)进行车道线检测，研究者使用特征图切片卷积替换常规逐层卷积，在隐藏层后添加一个空间卷积神经网络从多个维度提升空间探索能力。实验证明此网络可以提升车道线的检测效果，但是由于多维度操作导致检测速率降低。Qin 等人^[54]提出了(Ultra Fast, UFast)网络，该网络将车道线检测视为像素行分类问题，通过将图像切割为多个 anchor，每个 anchor 对应 y 轴坐标预测，模型进一步确定各 y 坐标行是否属于车道线，由此推导出 x 坐标值，最后整合这些点以描绘出整体车道线。杜黎等人^[55]提出了一种快速的车道线检测方法，首先设计一个特征提取模块，准确地提取车道特征，然后和龙格库塔方法相结合，对残差结构的输出进行加权融合，在解码层使用基于行锚检测方法的车道位置预测模块提高网络的检测速度。

1.3 论文主要研究内容及创新点

本文针对目前无人驾驶在复杂的路况环境下的车道线检测精度较差，误检率较高的情况，将无人驾驶复杂环境分为三种情况进行研究，分别是夜间环境下的车道线检测，雾天环境下的车道线检测和雨夜环境下的车道线检测。为了使车道线检测网络可以更加准确的检测出恶劣环境中的车道线，首先设计图像增强网络对恶劣环境下的图像进行增强，然后使用改进的车道线检测网络对车道线进行检测。论文的主要研究内容分为以下三个部分：

(1) 夜间环境下的车道线检测

针对夜间场景下车道线图像存在整体较暗、光照不均匀导致检测准确率低，漏检率和误检率较高的问题，首先设计一个无监督型的夜间增强网络。通过将图像转换到 HSV 颜色空间处理避免颜色分量的损失，引入亮度调整函数生成同一场景的不同亮度图像，将两个亮度图像输入到反射率估计网络中，在编码器

网络和跳跃连接中添加(Convolutional Block Attention Module, CBAM)注意力机制提升网络的增强效果,通过 Retinex 理论获得增强后的图像。其次在 ERF-Net 网络中引入自注意力蒸馏,提升车道线检测网络的检测效果。

(2) 雾天环境下的车道线检测

为了解决雾天环境下,车道线受到雾气的遮挡,检测网络精度降低。设计一个基于 CycleGAN 的无监督型去雾网络,为了充分使用不同编码层的特征,设计多层特征融合模块将不同编码层的特征进行融合,然后使用(Frequency Channel Attention Networks, FCA-Net)网络加强网络对不同通道的关注度。同时为了提升判别器的判别效果,将判别器扩展为全局和局部判别器,从不同的角度对生成图像进行判别。最后使用车道线检测网络对增强后的图像进行检测。

(3) 雨夜环境下的车道线检测

针对夜间雨天环境下图像整体亮度较低,去雨网络无法去除隐藏在黑暗中的雨条纹,车道线信息会受到复杂光照条件的影响,同时雨条纹会遮挡车道线部分。设计一个端到端的夜间雨天图像增强网络,首先使用雨条纹检测模块检测夜间条件下的雨条纹位置,然后将学习到的雨条纹先验知识输入到 Uformer 网络中进行图像复原,最后使用车道线检测网络对图像增强效果进行检验。

1.4 本文章节安排

本文主要围绕无人驾驶复杂环境下的车道线检测展开研究。论文中各章节的主要研究内容如下:

第 1 章:绪论。主要论述复杂环境下图像增强和车道线检测的研究背景和意义,并着重介绍了国内外对夜间图像增强、雾天图像增强、雨天图像增强和车道线检测的研究现状与研究进展,此外介绍本文各章节的研究内容。

第 2 章:基于 Retinex 理论和自注意力蒸馏的夜间车道线检测。本章将网络分为两个部分,分别是夜间图像增强网络和车道线检测网络。对于夜间图像增强网络,首先介绍了 Retinex 理论,然后分析了 HSV 颜色空间的优势,最后介绍了改进的夜间图像增强网络。对于车道线检测部分,在 ERF-Net 网络中引入自注意力蒸馏提升车道线检测能力。最后在合成数据集和真实数据集上对夜间增强网络进行定量和定性分析,同时在车道线数据集上检测改进的车道线检测

能力。

第 3 章：基于多层特征融合的雾天车道线检测。本章首先介绍了大气散射模型，由此引出图像通道对去雾能力的重要性，然后介绍了本章的骨干网络，即循环生成对抗网络。其次解释了多层特征融合模块和轻量化残差模块等多个模块的详细结构和作用。最后对网络进行定量和定性分析，并使用第二章的车道线检测网络展示去雾效果。

第 4 章：基于 Uformer 的雨夜车道线检测。本章首先介绍了图像复原网络 Uformer 和去雨网络 PReNet，然后分析了改进的夜间去雨网络，解释了夜间预先检测出夜间雨条纹位置的必要性。最后使用创建的夜间城市雨天景观数据集对网络进行训练，同时在合成数据集和真实夜间雨天环境对网络进行定量和定性分析，并展示夜间雨天图像增强前后车道线检测效果。

第 5 章：总结与展望。针对本文工作进行总结，阐明整体结构与各模块间的功能，并分析存在的不足与改进措施，同时对未来研究进展做出展望。

第 2 章 基于 Retinex 理论和自注意力蒸馏的夜间车道线检测

2.1 引言

夜间图像受到环境亮度不足的影响，车载视觉传感器采集到的图像整体对比度较差、亮度也比较低，车道线的边缘细节模糊不清，当光照严重不足时，无法从图像中识别到车道线，其次，夜间环境中存在大量的干扰光源，如路灯、车灯等，这些光源可能会产生反射和光晕现象，进一步降低车道线的可见性，对后续车道线检测产生严重的影响。为了更加直观的看出夜间图像和正常光照图像之间的差异，将 SICE 测试数据集中的 229 张图像使用直方图的形式表示出来，如图 2-1 所示，横坐标为 0 到 255 的灰度值分布，纵坐标表示各个灰度值出现的频次。在正常光照条件下，图像的灰度值分布相对均匀，边缘细节信息较为突出，然而夜间图像整体均处在像素值较低的区域，这样的分布会使图像呈现黑色，边缘细节等信息会被隐藏。

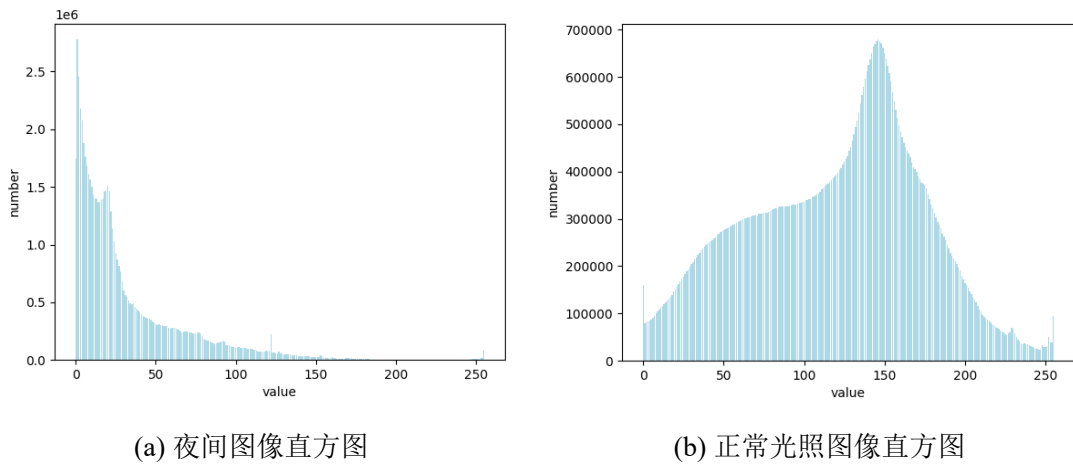


图 2-1 灰度值分布直方图对比

为了解决夜间环境下车道线检测会受到环境复杂光源和低光照的影响，本章首先将图像从 RGB 色彩空间转换至 HSV 色彩模型，保留了色调(Hue, H)、饱和度(Saturation, S)，将 Retinex 理论应用到亮度值(Value, V)上。设计一个亮度调整函数来获得一个异常的亮度，然后将 Retinex 理论应用到其 V 通道上。最后

将增强的亮度通道和保留的色调、饱和度通道重新组合成正常光照图像。在增强完图像后使用改进的车道线检测网络检测车道线，将车道线检测看作语义分割任务，利用改进的 ERF-Net 网络^[56]将图像划分为车道线和背景图像。引入了自注意力蒸馏，利用网络本身层的注意力图作为下层的蒸馏目标。本章的研究路线如图 2-2 所示。

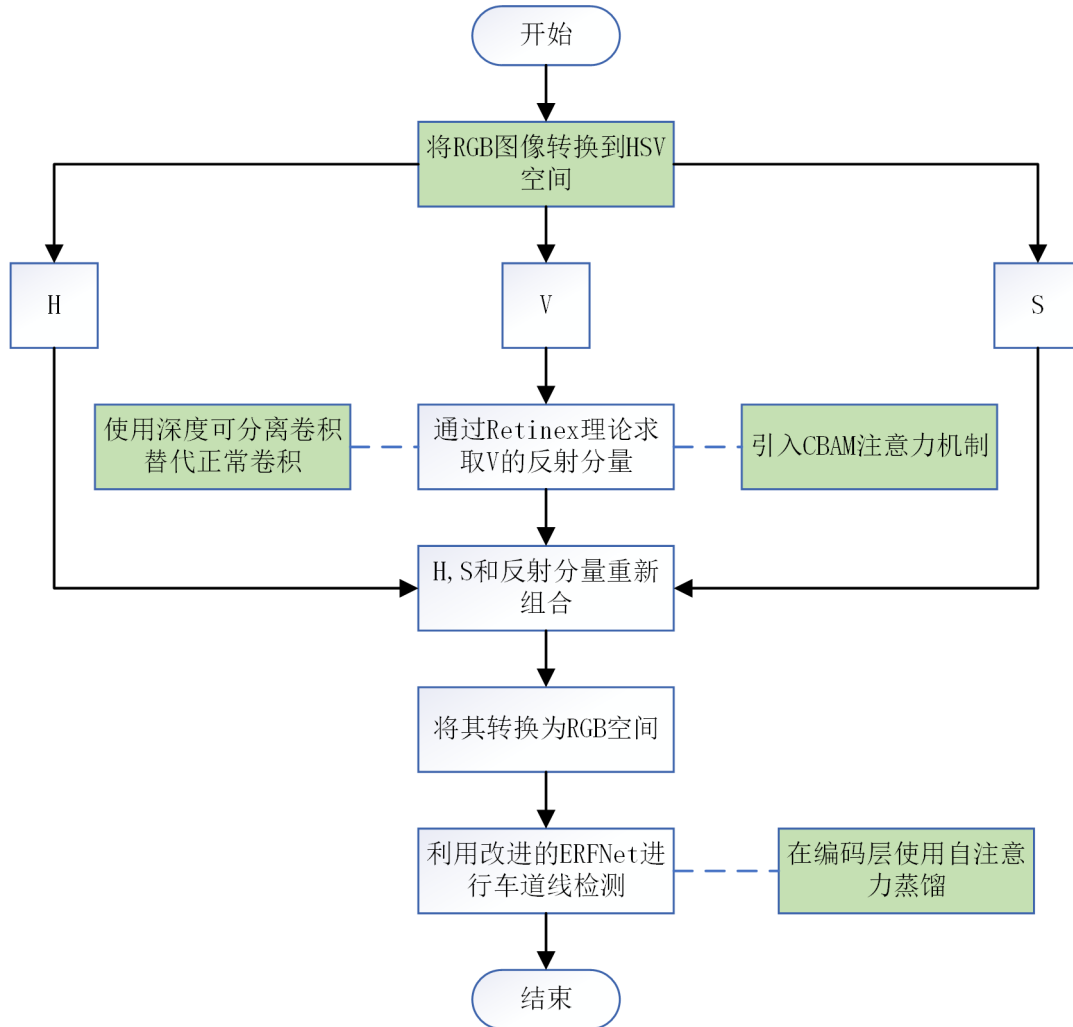


图 2-2 夜间车道线检测研究路线图

2.2 基于改进 Retinex 理论的夜间图像增强

2.2.1 Retinex 理论

Retinex 理论是图像增强和颜色恢复领域的重要理论之一。Retinex 假设人类视觉系统会对物体反射的光谱进行自适应的调整，以消除来自光照的影响，从而更好地辨别物体的颜色。假设当人们突然遇到停电的情况，在起初阶段，

人眼在暗环境中可能观察不到明显的视觉信息，然而随着眼睛逐渐适应该环境，能够隐约观察到一些景物，这表明颜色与环境中的光线无关，仅与物体表面的反射光相关。**Retinex** 理论模型如图 2-3 所示，**Retinex** 理论认为图像由反射与照度两部分构成，核心在于忽视光线作用，凸显物体本质属性。在 **Retinex** 理论框架下，处理图像时应着重保留反映轮廓与纹理等高频信息的反射分量，同时需分离并尽可能剔除易导致色彩失真的照度分量。

将入射光用 L 表示，反射物体用 R 表示，人眼观察到的图像用 S 表示。入射光 L 照射到物体的表面，通过物体 R 的反射进入到人的眼中，从而观察到图像。该理论模型可以用公式 2-1 进行描述：

$$I(x, y) = R(x, y) * L(x, y) \quad (2-1)$$

其中 $I(x, y)$ 代表的是人眼所观察到的图像， (x, y) 是像素坐标， $R(x, y)$ 是反射图像，是物体本身存在的固有属性， $L(x, y)$ 是照度图像，是由场景的光线特点所决定的。**Retinex** 理论的本质是保留图像中的反射分量，对光照分量进行去除或增强，以此来提升图像质量。

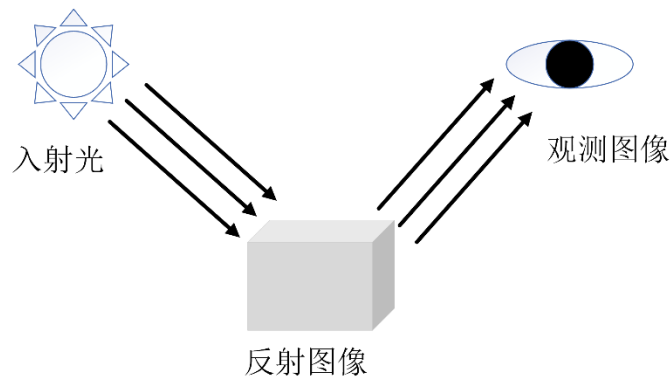


图 2-3 Retinex 理论模型图

2.2.2 RGB 和 HSV 颜色空间

图像的表现形式逐渐从早期的灰度图像转变为现在的彩色图像，为了实现规范性的描述，需要确立一种通用的色彩空间模型标准。在数字图像处理中，通常采用三维独立变量构建的空间坐标系来精确描绘色彩信息，而不同变量测量方式的选择会衍生出多种不同的色彩空间模型。**RGB**、**HSV** 和 **YUV** 都是常用的颜色空间模型^[57]。**RGB** 基于红绿蓝三原色合成其他颜色。在 **RGB** 模型中，

每种颜色通过不同强度的三种原色的组合来表示。通过调整这三种颜色通道的数值，可以呈现出各种颜色。将三原色数值进行归一化处理，并采用三维坐标系中的 x 、 y 、 z 轴来表示颜色。如图 2-4 所示，当三原色的值都为 0 时图像呈现黑色，当三者坐标值均为 1 时图像呈现白色。三个坐标轴上的点代表着三种原色。RGB 色彩模型具有色域范围广的特点，所以 RGB 色彩模型会应用到各种硬件设备上。虽然日常使用的图片均是 RGB 颜色模型，但是 RGB 空间中三个分量皆与亮度紧密相关，一旦亮度变动，各分量亦受影响，故 RGB 颜色空间不利于在图像处理中使用。因此本章将数据集中的图像从 RGB 颜色空间转换到 HSV 颜色空间进行训练。

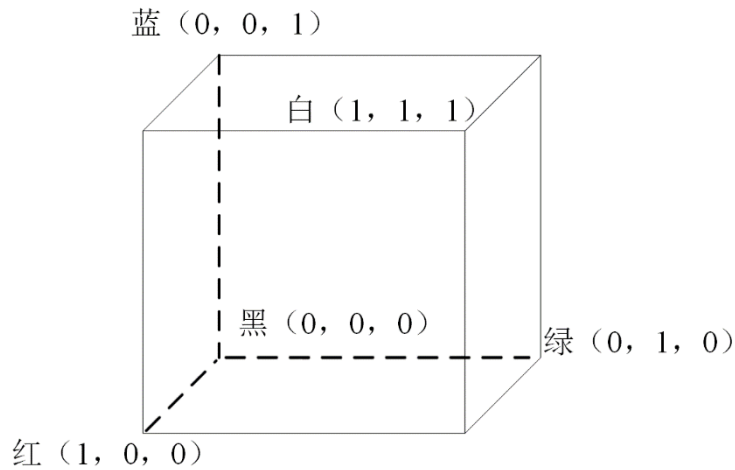


图 2-4 RGB 色彩模型

不同于显示设备，人眼视觉多采用 HSV 颜色模型来解析颜色，该模型通过色调、饱和度和亮度这三个维度来描述颜色特征。色调界定颜色种类，饱和度反映颜色的纯净度与鲜艳度，高饱和度色彩更纯正，低饱和度趋于灰色调，亮度决定颜色的明暗，亮度高则颜色明亮，亮度低则颜色偏暗。对比 RGB 颜色空间，HSV 颜色空间的优势可以总结为以下两点，首先 H ， S ， V 三个通道相互正交，可以实现亮度和色彩分量的解耦，由于本章所使用的图像增强网络是无监督型网络，输入夜间图像是没有对应的同场景图像作为训练目标，网络无法学习到从低光照图像到正常光照图像的颜色映射，因此当采用 HSV 颜色模型时，只对其中的亮度通道操作可以有效的避免网络学习对颜色产生的干扰。其次低光照图像的 V 通道值对应 RGB 三个分量中像素的最大值，文献^[78]指出，该最大值通道拥有最高的熵，这对于夜间图像增强网络意味着可提供更为丰富的信息资

源。

从 RGB 颜色空间到 HSV 颜色空间的转换公式为：

$$V = \max\{R, G, B\} \quad (2-2)$$

$$S = \frac{V - \min\{R, G, B\}}{V} \quad (2-3)$$

$$H = \frac{1}{6} \begin{cases} \frac{G-B}{V - \min\{R, G, B\}} & V = R, \min\{R, G, B\} = B \\ \frac{R-B}{V - \min\{R, G, B\}} + 1 & V = G, \min\{R, G, B\} = B \\ \frac{B-R}{V - \min\{R, G, B\}} + 2 & V = G, \min\{R, G, B\} = B \\ \frac{G-R}{V - \min\{R, G, B\}} + 3 & V = B, \min\{R, G, B\} = R \\ \frac{R-G}{V - \min\{R, G, B\}} + 4 & V = B, \min\{R, G, B\} = R \\ \frac{B-G}{V - \min\{R, G, B\}} + 5 & V = R, \min\{R, G, B\} = G \end{cases} \quad (2-4)$$

在对 V 通道处理完后需要将其重新转换回 RGB 颜色空间，其转换公式为：

$$J = \lfloor 6H \rfloor \quad (2-5)$$

$$F = 6H - J \quad (2-6)$$

$$P = V \cdot (1 - S) \quad (2-7)$$

$$Q = V \cdot (1 - S \cdot F) \quad (2-8)$$

$$T = V \cdot (1 - S \cdot (1 - F)) \quad (2-9)$$

$$(R, G, B) = \begin{cases} (V, T, P) & J = 0 \\ (Q, V, P) & J = 1 \\ (P, V, T) & J = 2 \\ (P, Q, V) & J = 3 \\ (T, P, V) & J = 4 \\ (V, P, Q) & J = 5 \end{cases} \quad (2-10)$$

为了证明 HSV 颜色空间的优势，设计一个简单的实验^[79]，从 SICE 数据集

中随机选择一组相同场景的低光照图像和正常光照图像，将两者均转换到 HSV 颜色空间，然后取低光照图像的色调 S ，饱和度 H 和正常光照图像的亮度通道 V ，在通道维度上进行拼接，将拼接后的图像转换回 RGB 颜色空间。实验结果如图 2-5(a)和(d)所示，重建后的图像与正常光照图像的颜色基本一致，因此对于夜间图像的增强，使用 HSV 颜色模型更适用于后续图像处理操作，即只对 HSV 颜色模型中的亮度通道进行增强，颜色部分全部来自于夜间图像。

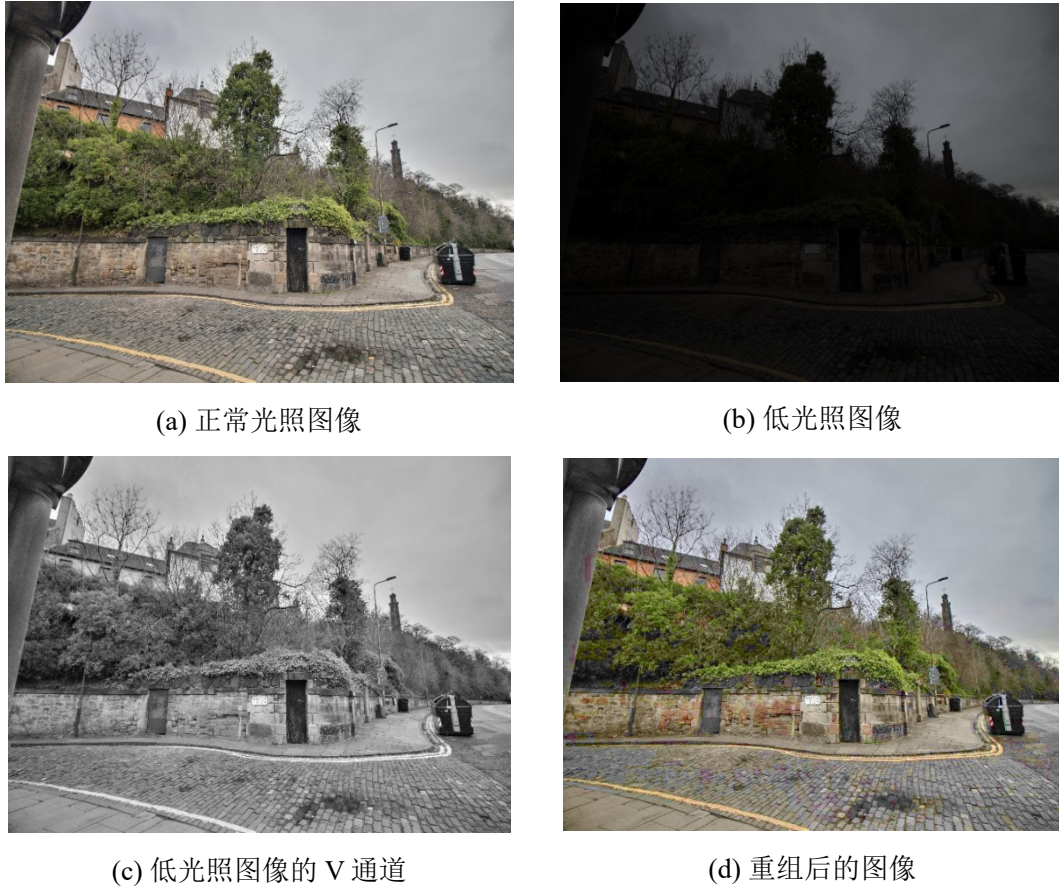


图 2-5 HSV 颜色模型实验结果

2.2.3 夜间图像增强网络

众多基于 Retinex 理论的图像增强网络分解夜间图像为光照和反射分量，通过增强光照分量亮度并对反射分量进行去噪，以此实现夜间图像增强。如 RetinexNet 网络，但是这样的算法可能会导致图像在网络重建时出现颜色失真，增强效果不稳定，对夜间不均匀光照环境的适应性较差。基于以上问题，为了适应各种光照条件，将反射分量作为增强后的亮度图像。由于 Retinex 理论假设在不同的光照条件下，同一场景的图像具有相同的反射分量，同时本章所设计

的夜间图像增强网络为无监督型网络，缺少同一环境下的真实正常光照图像。因此设计一个随机亮度调整函数生成一个同一场景的不同亮度图像。该随机亮度调整函数使用具有随机指数的幂函数来计算，具体的公式如 2-11 所示：

$$V'(x) = V(x)^\gamma \quad (2-11)$$

其中 $V(x)$ 是 V 通道的像素值， $V'(x)$ 是通过调整后的 V 通道值， γ 是随机值，根据 V 通道的值改变取值范围。由于幂函数是单调递增的函数，这样可以保持输入图像的亮度和调整亮度之间的梯度方向一致，同时幂函数的随机指数可以保持亮度的多样性。

本章抛弃传统 Retinex 理论的图像重建阶段，直接通过图像增强网络获得增强的夜间图像反射分量。将初始夜间图像和调整后的光照图像同时输入到夜间图像增强网络中进行训练。根据 Retinex 理论可以将 $V(x)$ 和 $V'(x)$ 分解为两部分，如公式 2-12 所示。

$$\begin{cases} V = R \times L \\ V' = R' \times L' \end{cases} \quad (2-12)$$

其中 R 为反射分量， L 为照度分量。将 L 看作是计算 R 的中间变量，则上式可改为公式 2-13 形式。

$$\begin{cases} R = V \times \frac{1}{L} \\ R' = V' \times \frac{1}{L'} \end{cases} \quad (2-13)$$

这样可以不用计算 V 和 $R \times L$ 之间的差值，可以避免在重建图像时丢失信息。该增强网络采用常用的 U-Net 网络，如图 2-6 所示。其有 4 个编码层和 4 个解码层。首先使用两个卷积层和两个 ReLU 激活函数将输入的 V 通道图像转换成适应神经网络的特征图。4 个编码层和 4 个解码层的卷积操作均使用深度可分离卷积，以此减少网络的参数量，提升网络的运行速度。对于下采样操作，使用步数为 2 的 2×2 最大池化将特征图尺寸缩小。在解码层恢复图像尺寸的阶段，采用双线性插值替代反卷积，因为反卷积的参数量较大，双线性插值可以使网络更易训练。为了提升网络对夜间图像黑暗处的关注度，在编码层的卷积操作后添加 CBAM 注意力机制。同时为了提升图像重建后的效果，在解码器重建图

像的过程中，首先在编码层相应尺度特征图处添加 CBAM 注意力机制，随后采用跳跃连接机制，将经过处理的下采样特征图与上采样阶段相应尺寸的特征图相互衔接，此操作有助于增强上采样过程中的信息丰富度，最后使用 1×1 的卷积将特征图转换回光照分量的逆。

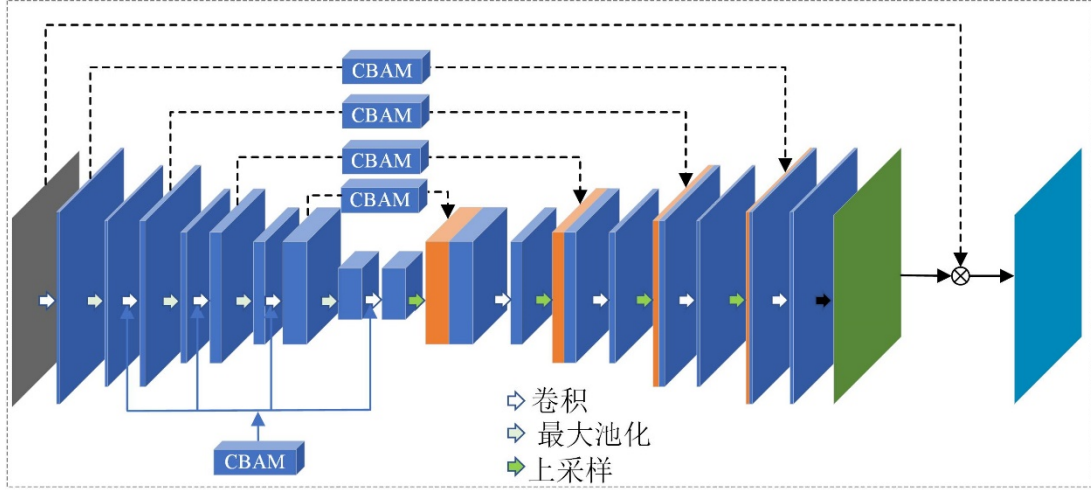


图 2-6 夜间图像增强网络结构图

为了提升网络的运行速度，网络中的部分卷积使用深度可分离卷积进行替换，普通卷积和深度可分离卷积的具体实现方式如图 2-7 所示。图 2-7(a)是普通卷积的示意图，由图可知，在卷积操作过程中，输入图像的每个通道都要与卷积层中的所有卷积核进行运算，紧接着，在同一通道内部，将各个卷积核计算出的结果汇总求和，以得到单个通道的输出，最后将所有卷积核的通道输出逐一整合形成最终的卷积层输出结果。输出的结果如公式 2-14 所示。

$$\begin{cases} O_1 = K_1 \otimes I_1 + K_1 \otimes I_2 + K_1 \otimes I_3 \\ O_2 = K_2 \otimes I_1 + K_2 \otimes I_2 + K_2 \otimes I_3 \\ O_3 = K_3 \otimes I_1 + K_3 \otimes I_2 + K_3 \otimes I_3 \end{cases} \quad (2-14)$$

其中 $\otimes$ 表示卷积操作， I 为输入的三个通道， K 为卷积核， O 为单通道输出。

在普通卷积中，每一个卷积核都会对所有输入通道进行综合处理，这一特性会显著提升网络参数的数量，进而可能对系统的整体运行速度产生消极影响。为了有效应对这一问题，许多研究者使用深度可分离卷积^[58]替代普通卷积。深度可分离卷积将卷积过程分为两个阶段，如图 2-7(b)所示，首先是专注于单一通道的深度卷积，每个卷积核仅针对单个输入通道进行卷积操作，然后逐点卷积使用 1×1 的卷积核对每个位置的所有通道进行加权组合，同时逐点卷积引入了(Rectified Linear Unit, ReLU)激活函数增强模型捕捉复杂模式的能力。

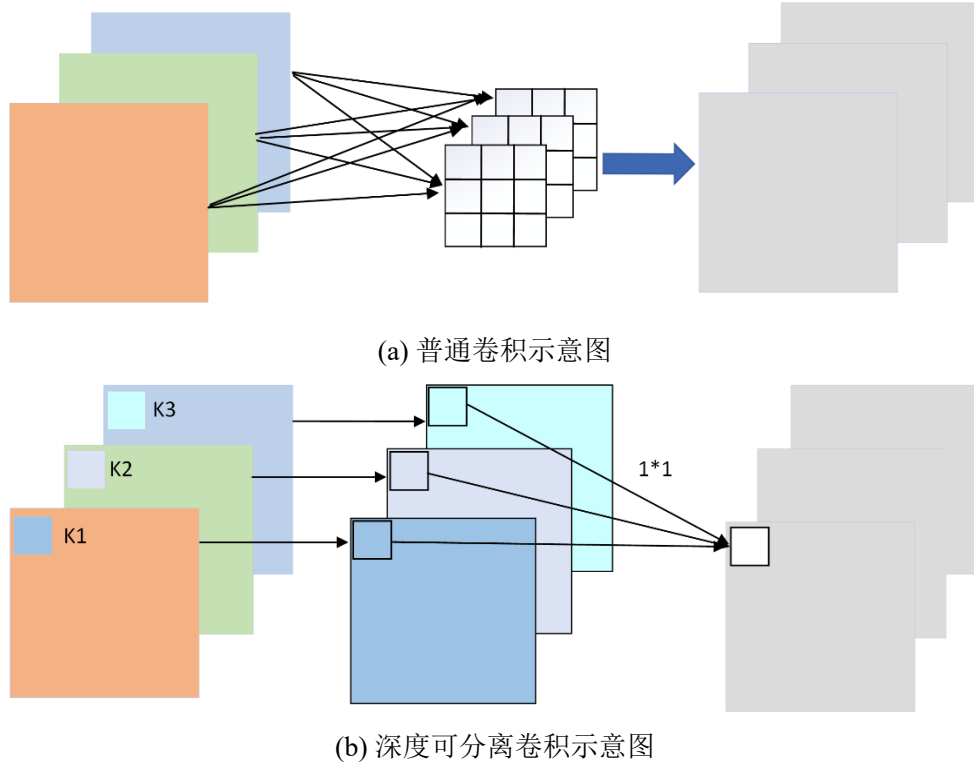


图 2-7 普通卷积和深度可分离卷积

注意力机制能够有针对性地将有限的计算能力集中在任务的关键部分。通过运用注意力机制，网络能在海量输入信息中找准重心，更加专注地处理与当前任务密切相关的数据，从而提升整体处理效果。注意力机制可分为三大类：空间注意力、通道注意力以及混合维度注意力。空间注意力侧重于在特征图的空间维度上进行权重分配，针对特征图的不同区域给予不同的重视程度。而在通道注意力方面，通过在空间维度压缩特征图，为各个通道分配不同的权重，从而增强模型对不同通道特征的辨识能力。混合维度注意力通常运用诸如 CBAM 的网络^[59]，其首先沿通道方向和空间方向分别推理出注意力分布图，然后将这两个维度上的注意力信息与原始特征图逐点相乘，以此实现对特征图的自适应优化。不同于普通的通道注意力，如(Squeeze-and-Excitation Networks, SENet)只通过全局平均池化将空间信息嵌入通道中，CBAM 网络同时使用全局平均池化和全局最大池化，通过两者的结合抑制信息的丢失。CBAM 网络结构如图 2-8 所示，该网络整合了通道注意力机制和空间注意力机制这两个关键模块，通过串行的方式将两者连接在一起。借助这种双重注意力引导策略，CBAM 网络能够在夜间图像处理场景中实现对关键视觉特征的有效提取与优化，提升网络生成图像的质量。

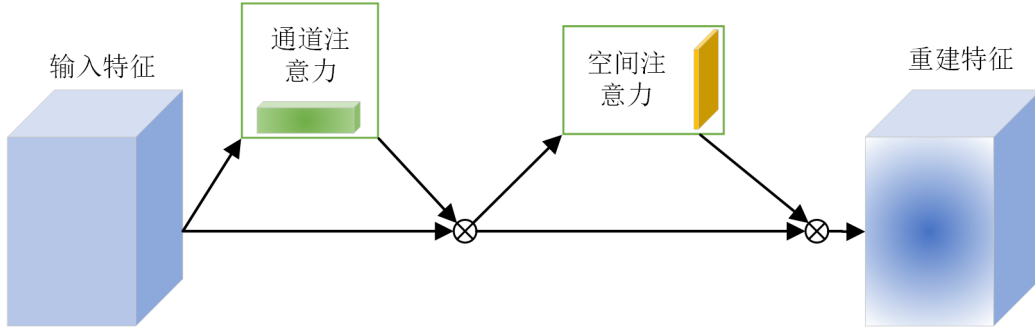


图 2-8 CBAM 注意力结构图

2.2.4 损失函数

为了使图像细节恢复更为平滑，在增强图像亮度的同时保持图像空间结构不变，使增强后的图像拥有均匀的光照和良好的视觉效果。本章将夜间增强网络的损失函数定义为公式 2-15。

$$L = L_1 + L_2 + L_3 + \alpha L_4 \quad (2-15)$$

其中 L 为总的损失函数， L_1 为曝光损失， L_2 为反射率一致损失， L_3 为空间结构损失， L_4 为照明平滑损失， α 为照明平滑损失的权重。损失函数的具体定义如下：

曝光损失：为了控制生成的反射率图的亮度，计算反射率的平均亮度和给定正常图像曝光值的差异。曝光损失 L_1 可以定义为：

$$L_1 = \|R - E\|_2^2 \quad (2-16)$$

其中 R 为生成的反射率， E 为维数和 R 相同的矩阵，且其所有值均为 0.7。

反射率一致损失：根据 Retinex 理论， V 和 V' 的反射率应当保持一致，所以，反射率一致损失可以定义为：

$$L_2 = \|R - R'\|_2^2 \quad (2-17)$$

其中 R 和 R' 分别是 V 和 V' 生成的反射率， $\|\cdot\|_2$ 为 L_2 范数。

空间结构损失：为了保证输入图像的空间结构可以保留到反射率图上，引入空间结构损失，该损失计算了输入图像与反射率图每个像素的水平 and 垂直梯度之间的差异。空间结构损失 L_3 可以定义为：

$$L_3 = \|\nabla R_m - \nabla V_m\|_2^2 \quad (2-18)$$

其中 R_m 和 V_m 分别表示生成的反射率和输入的平均池化结果。 ∇ 表示水平和垂直方向上的一阶差分运算。

照明平滑损失：Retinex 理论中照明应当是平滑的，这样图像的细节可以保留在反射率图上。利用 S 来代表 $1/L$ ，则照明平滑损失可以定义为：

$$L_4 = \|\nabla S\|_2^2 + \|\nabla S'\|_2^2 \quad (2-19)$$

2.3 基于自注意力蒸馏的车道线检测网络

2.3.1 ERF-Net 网络

车道线检测可以简单的理解为使用语义分割方法将一张输入图像划分为车道线区域和背景区域。CNN 虽然可以实现语义分割效果，但是在准确性和分割效率上较差。(Efficient Neural network, ENet)网络^[60]虽然在运行速度上有很大优势，但是在网络结构上会丢失大部分的信息，最终导致精度变低。为了实现检测速度和检测效果之间的平衡，本章车道线检测网络采用 ERF-Net 网络为骨干网络。

ERF-Net 网络是一种端到端的编码器-解码器结构，如图 2-9 所示。编码器部分的功能是对输入的图像进行特征提取并对图像进行下采样，在下采样模块后使用残差模块对不同感受野特征进行提取，解码器结构整合了上采样单元与残差连接机制，通过运用反卷积操作对编码阶段得到的特征图进行上采样，通过残差模块恢复细节信息，最后网络输出和原图相同尺寸的分割结果图。

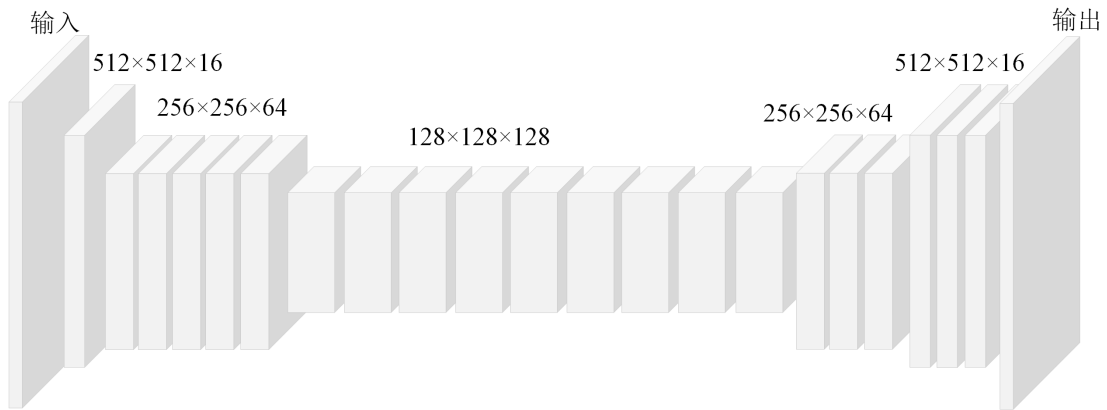


图 2-9 ERF-Net 网络结构图

ERF-Net 详细网络结构如表 2-1 所示：

表 2-1 ERF-Net 网络结构

	层	名称	尺寸
Encoder-1	1	Downsampler	512×512×16
Encoder-2	2	Downsampler	256×256×64
	3-7	5×Non-bt-1D	256×256×64
Encoder-3	8	Downsampler	128×128×128
	9	Non-bt-1D (dilated 2)	128×128×128
	10	Non-bt-1D (dilated 4)	128×128×128
	11	Non-bt-1D (dilated 6)	128×128×128
	12	Non-bt-1D (dilated 8)	128×128×128
	13	Non-bt-1D (dilated 2)	128×128×128
Encoder-4	14	Non-bt-1D (dilated 4)	128×128×128
	15	Non-bt-1D (dilated 6)	128×128×128
	16	Non-bt-1D (dilated 8)	128×128×128
Decoder-1	17	Deconvolution	256×256×64
	18-19	2×Non-bt-1D	256×256×64
Decoder-2	20	Deconvolution	512×512×16
	21-22	2×Non-bt-1D	512×512×16
Decoder-3	23	Deconvolution	1024×1024×C

ERF-Net 网络模型共有 23 层，其中根据下采样的位置，将整个编码阶段分为 4 个编码层，Downsampler 模块通过 3×3 的卷积和最大池化对输入图像进行下采样，然后使用 Concat 的操作将两者结合得到 $512 \times 512 \times 16$ 的特征图，最后使用 ReLU 激活函数将特征图传入下一层。这样可以在初始阶段减小输入图像的尺寸，同时降低视觉信息的冗余。

在编码层和解码层中均含有一种特殊的残差模块，与常规的残差模块不同，

ERF-Net 网络提出了一种新颖的残差结构 (Non-bottleneck-1D), 常规残差模块和 Non-bottleneck-1D 的网络结构如图 2-10 所示。传统的残差模块如图 2-10(a) 所示, 直接将输入跨接到 3×3 卷积输出部分, 这样会导致多次使用残差模块时, 参数量大幅提升。为了减少所需的参数量, 瓶颈模块通过 1×1 卷积层对输入特征进行降维处理, 然后再通过 3×3 卷积层对其进一步运算。虽然瓶颈模块可以减少参数量, 但是其将通道进行压缩, 导致信息丢失, 降低了模型精度。Non-bottleneck-1D 结构采用一维卷积技术, 分别在垂直和水平方向上执行特征提取操作: 首先运用 3×1 的一维卷积沿列方向对输入进行卷积处理, 以捕获行间的特征; 随后通过 1×3 的一维卷积沿行方向进行卷积运算, 以提取列间的特征信息。Non-bottleneck-1D 模块速度比瓶颈模块更快, 同时可以保持与常规残差模块相近的准确性。

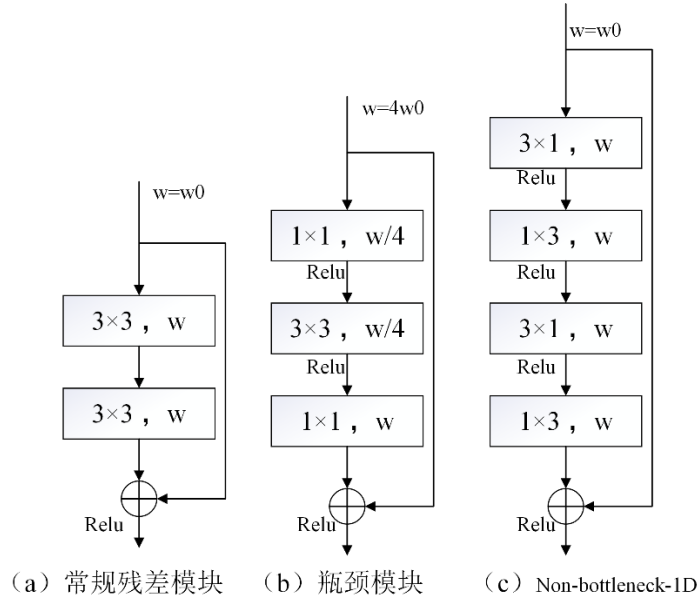


图 2-10 不同残差模块网络结构图

假设输入和输出的通道数为 C , 通过计算量进行直观分析, 常规残差模块的计算量为:

$$(3 \times 3 \times C \times C + 3 \times 3 \times C \times C) = 18C^2 \quad (2-20)$$

瓶颈模块的计算量为:

$$1 \times 1 \times C \times \frac{C}{4} + 3 \times 3 \times \frac{C}{4} \times \frac{C}{4} + 1 \times 1 \times \frac{C}{4} \times C = \frac{17}{16}C^2 \quad (2-21)$$

Non-bottleneck-1D 模块的计算量为:

$$(3 \times 1 \times C \times C + 1 \times 3 \times C \times C + 3 \times 1 \times C \times C + 1 \times 3 \times C \times C) = 12C^2 \quad (2-22)$$

通过对比可以发现，Non-bottleneck-1D 模块比常规残差模块计算量小，随着网络深度的增加，网络节省的计算资源将逐渐变大。

2.3.2 车道线检测网络

使用语义分割的方法进行车道线检测会受到车道线特征变化的影响，当车道线受到遮挡和磨蚀，车道线特征会变得不完整，有些区域甚至会出现特征消失的情况，使得车道线检测愈加困难。为了克服车道线特征稀疏性问题，在ERF-Net 的编码层中融入了自注意力蒸馏^[61]，使用较小的网络模型去学习大型教师网络。自注意力蒸馏不需要教师网络，其使用该网络的深层指导前一层网络。自注意力蒸馏将当前网络层生成的注意力图反馈给上一层，并且不需要使用任何的标签信息。自注意力蒸馏通过层与层之间的注意力图进行学习，可以使靠近输入端的编码层学习到深层的语义信息，浅层的注意力图可以获得更加丰富的上下文信息，更加聚焦于目标任务学习。同时浅层学习到的优势也会传导到深层，从而解决特征稀疏问题。

改进后的 ERF-Net 网络结构简图如图 2-11 所示，在编码器的第一、二、三、四层均加入自注意力蒸馏，让编码器的第一、二、三、四层学习第二、三、四及四层以后的注意力图信息。网络能够更有效地掌握精确的车道线细节信息，从而提升整体的识别精度。

自注意力蒸馏的使用依赖于每个网络层输出的注意力图，首先需要对注意力图进行计算，生成注意力图的过程可理解为构建一个函数模型，该模型旨在探寻一种映射机制，其功能是从多维度的通道特征向量中提炼出一个具有二维表达的注意力分布。在注意力图中，每一个单独的元素都代表着该元素在其所在层输出中的相对重要程度。映射函数如公式 2-23 所示。

$$f_{\text{sum}}^2(A_m) = \sum_{i=1}^{C_m} |A_{mi}|^p \quad (2-23)$$

其中 A_m 为编码器第 m 层的输出， C_m 表示其通道， A_{mi} 表示 A_m 的第 i 个通道。

$|\cdot|^2$ 为概率绝对值函数，该函数首先取所有通道上的元素的绝对值和，然后为每个通道添加一个指数调节参数。

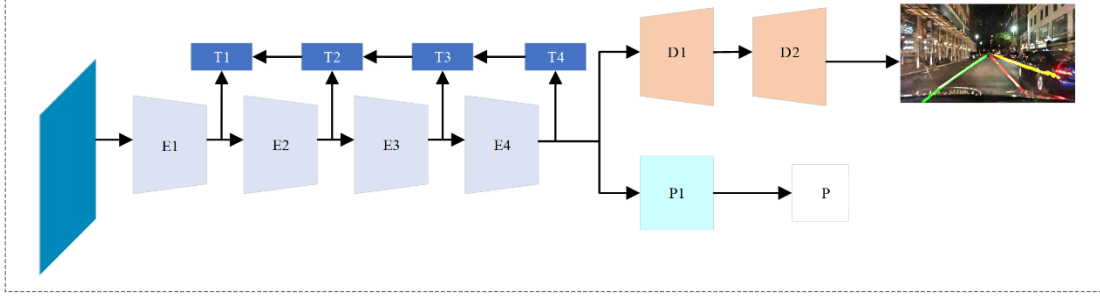


图 2-11 车道线检测网络结构图

由于需要使用当前注意力图去学习下一层的注意力图，两个注意力图的尺寸大小需要保持一致，对下一层的注意力图进行双线性插值上采样处理，然后对上采样后的注意力图实施 **Spatial Softmax** 变换，将其转化为一维向量形式，最后通过 **Softmax** 函数运算得出每层所需的注意力图分布。注意力图生成器的具体公式如下：

$$T_G = \Phi \left(\mathcal{B} \left(\mathcal{F}_{\text{sum}}^2 (.) \right) \right) \quad (2-24)$$

其中 $\mathcal{B}$ 为双线性上采样操作， Φ 为 **Spatial Softmax** 操作。

通过运用自注意力蒸馏，底层网络可以借鉴高层网络所包含的深层语义特征，将深层的注意力图迁移到浅层网络结构中，促使浅层网络可以明确模型应当着重关注的区域，进而对该区域进行更为精细和准确的分割处理。浅层通过注意力图学习到的内容也可以通过网络传导到深层，从而提升网络的检测性能。

在编码器尾端和解码器的前端处添加车道线分类网络，弥补其网络结构导致的全局信息损失。首先，对编码器输出的特征进行卷积处理以深化特征提取，紧接着通过 **ReLU** 激活函数对其进行非线性转化。随后再次进行卷积操作，并在通道维度上应用 **Softmax** 函数以调整特征权重。进一步采取最大池化缩小特征图尺寸，随后经由两个全连接层处理，输出一个维度为 4 的一维向量。最后，运用 **Sigmoid** 函数对这四个元素进行激活，从而得到四条车道线的分类预测结果。

2.3.3 损失函数

语义分割损失：在语义分割场景中通常使用交叉熵损失函数，对于车道线检测任务，需要将检测网络的输出通过 **Softmax** 函数转化为概率的形式，然后再使用交叉熵损失函数处理概率值。对于每个像素点，交叉熵损失函数通过对

比预测标签与实际标签之间的概率分布差异，从而量化两者之间的差异程度。

Softmax 函数的公式为：

$$S'_i = \frac{e^{x_i}}{\sum_j^N e^{x_j}} \quad (2-25)$$

其中 x_i 表示检测网络第 i 个类别的输出， x_j 表示检测网络第 j 个类别的输出， N 表示类别的总数。 S'_i 表示类别 x_i 的输出概率值。交叉熵损失函数的公式如下：

$$L_{seg}(S_i, S'_i) = -\sum_{i=1}^N S(x_i) \cdot \log(S'(x_i)) \quad (2-26)$$

其中 $S(x_i)$ 是第 i 个类别的真实值， $S'(x_i)$ 是第 i 个类别的预测概率值。

IOU 损失：在车道线检测任务中，IOU 损失可以用于衡量模型预测的车道线与真实车道线标签之间的重叠程度，可以帮助模型学习准确地预测车道线。其计算公式如下：

$$L_{IOU} = 1 - \frac{N_p}{N_p + N_g - N_o} \quad (2-27)$$

其中 N_p 为预测的车道像素数量， N_g 为真实车道像素数量， N_o 为预测车道区域和真实车道区域重叠部分的车道像素数量。

车道存在性损失：对于网络中的车道线存在分支，通过判断检测到的车道线是否存在来确保检测的准确性。对于该部分的损失仍然可以使用二值交叉熵损失函数。其公式如下：

$$L_{exist} = -\frac{1}{n} \sum_{i=1}^N S(x_i) \cdot \log(S'(x_i)) \quad (2-28)$$

其中 n 是车道线的总条数，在本章标签数据中将车道线存在标记为 1，否则标记为 0。

注意力蒸馏损失：为了确保浅层注意力图可以正确的学习深层语义信息，使用自注意力蒸馏损失对其进行约束。损失函数定义如下：

$$L_{distill}(A_m, A_{m+1}) = \sum_{m=1}^3 L_{MSE}(T_G(A_m), T_G(A_{m+1})) \quad (2-29)$$

其中 $T_G(A_m)$ 表示对注意力图进行尺度调整操作。

网络的整体损失函数可以定义如下：

$$L = L_{\text{seg}}(s, \hat{s}) + \alpha L_{\text{IOU}}(s, \hat{s}) + \beta L_{\text{exist}}(b, b) + \gamma L_{\text{distill}}(A_m, A_{m+1}) \quad (2-30)$$

其中 α 、 β 、 γ 是相应的权重系数。

2.4 实验结果分析

本节将介绍夜间图像数据集、车道线检测数据集，此外将介绍实验平台及训练过程参数设置，在夜间图像数据集以及真实夜间环境下对增强网络效果进行定性分析，将车道线检测网络同其他网络进行对比分析，最后将夜间图像和通过夜间图像增强网络增强后的图像输入到车道线检测网络检测分析。为了直观显示各部分网络的效果，使用图像质量评价指标和车道线检测指标对检测结果进行定量分析。

2.4.1 实验设置

本章使用 Pytorch 作为深度学习框架，变量类型为张量型，可以部署到 GPU 中进行计算。本章实验是在拥有 Intel(R) Xeon(R) Platinum 8255C CPU 和 NVIDIA RTX3090 GPU 的设备上进行的，显存为 24G，操作系统为 Ubuntu。

夜间车道线检测分为两个网络，对于第一部分的夜间图像增强网络，采用 SICE 数据集^[62]去训练网络。具体的说，将 SICE 训练集中同一场景的图像保留一张作为验证集，其余的作为训练集。为了方便网络训练，将图像大小调整为 512×512。设置学习率为 10^{-4} ，设置照明平滑损失系数为 10，对其训练 500 轮。使用 Adam 优化器更新网络参数。其中每 10 轮使用验证集评估一次，取最好的模型为最终模型。

对于第二部分的车道线检测网络，目前常用的车道线数据集为 CULane 数据集、Tusimple 数据集和 BDD100K 数据集^[63]。其中 CULane 和 Tusimple 数据集有车道线标签信息，而 BDD100K 无车道线标签，需要人工通过 labelme 去生成标签。有很多车道线检测网络是在 Tusimple 数据集上实验的，但是其场景单一，数据量少，均是在高速路段拍摄的图片，无法处理特定情况的车道线，因

此本实验使用 CULane 数据集，它包含了更多复杂场景的驾驶图像。CULane 数据集包含高速道路和城市街区的车道线图像，其中每张图像均有车道线标签数据，该数据集由 89k 张训练集，9k 张验证集和 35k 测试集组成，且包含夜间、弯道等多类别车道线图像。

将 CULane 数据集的图像调整为 288×800 。使用随机梯度下降法优化网络参数。将网络的初始学习率设置为 0.01，动量大小设置为 0.9，权重衰减设置为 10^{-4} ，学习率的衰减函数如公式 2-31 所示。为了防止过拟合或者显卡内存溢出，将 batch size 设置为 8，训练 12 个 epoch。损失函数的三个权重系数 α 、 β 和 γ 设置为 0.1。

$$lr = lr_{init} \times \left(1 - \frac{epoch}{epochs}\right)^{0.9} \quad (2-31)$$

其中 lr_{init} 为初始的学习率， $epoch$ 为当前的训练轮数， $epochs$ 为设定好的总轮数。

2.4.2 评价指标

在当前的图像处理研究中，图像增强的效果通常使用多种图像质量评价指标来评估。一般而言，图像质量评价体系可被划分为两大类别：一是主观评价，二是客观评价。主观评价是基于人眼观察和情感感受的评估方法，能够反映人们对图像质量的真实感受。但是主观评价的结果可能受到个体之间的差异和主观喜好的影响，不同的观察者可能对图像增强效果有不同的看法。因此，在主观评价之外引入了几种客观评价指标，以便对图像增强的效果进行全面而准确的分析。

1) 峰值信噪比(Peak Signal to Noise Ratio, PSNR): PSNR 通过计算增强图像与参考真实图像中对应像素之间差值的均方根误差，来量化评估两幅图像之间的差异程度。PSNR 是基于图像的均方误差进行计算的，PSNR 将均方误差转换为对数形式，以便更直观地表达图像质量的差异。PSNR 的值越大意味着图像与参考图像的相似度更高，图像失真度更低。均方误差的计算公式如下：

$$MSE = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n (x_{i,j} - y_{i,j})^2 \quad (2-32)$$

其中 MSE 为增强图像和真实图像的均方差。 m 和 n 分别代表输入图像的高和宽， i, j 为图像像素点的坐标。 $x_{i,j}$ 为增强后图像， $y_{i,j}$ 为真实图像。

PSNR 的计算公式如下：

$$PSNR = 10 \times \lg \frac{L^2}{MSE} \quad (2-33)$$

2) 结构相似性：结构相似性^[64]用于评估两幅图像之间的相似程度。相较于 PSNR，SSIM 进一步考虑了图像的结构一致性、纹理细节以及人类视觉感知等多个重要因素。SSIM 能够量化它们之间的差异，从而给出一个 0 到 1 之间的相似性分数，分数越高表示图像质量越好。结构相似性计算公式如下：

$$SSIM = l(x, y) * c(x, y) * s(x, y) \quad (2-34)$$

$$l(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \quad (2-35)$$

$$c(x, y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \quad (2-36)$$

$$s(x, y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \quad (2-37)$$

$$SSIM = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (2-38)$$

其中 μ_x 和 μ_y 分别为增强图像 x 和真实图像 y 的均值； σ_x 和 σ_y 分别为增强图像和真实图像方差； σ_{xy} 为增强图像和真实图像的协方差； $l(x, y)$ 是亮度相似性； $c(x, y)$ 是对比度相似性； $s(x, y)$ 是结构相似性； C_1 、 C_2 、 C_3 为常数。

3) 无参考图像质量评价(Naturalness Image Quality Evaluator, NIQE)：上述的指标均是针对有真实环境下的图像作为参考的指标。然而真实环境下的低质量图像缺少与其对应的真实图像，这会导致上述指标无法使用，因此引入无参考图像质量评价。NIQE^[65]旨在度量一幅图像与自然场景图像之间的差异程度，从而判断图像的质量。NIQE 的计算公式如下：

$$NIQE = \sqrt{(v_1 - v_2)^T \left(\frac{\Sigma_1 + \Sigma_2}{2} \right)^{-1} (v_1 - v_2)} \quad (2-39)$$

其中 ν_1 、 ν_2 和 Σ_1 、 Σ_2 分别表示真实图像和低质量图像 MVG 模型中的均值和协方差矩阵。

对于车道线检测网络的性能评价指标一般使用准确率(Accuracy)、漏检率(False Negative, FN)和误检率(False Positive, FP)衡量。其计算公式如下所示：

$$Accuracy = \frac{\sum_{clip} C_{clip}}{\sum_{clip} S_{clip}} \quad (2-40)$$

其中 C_{clip} 为网络预测正确的道路数， S_{clip} 为所有的道路数。

$$F_P = \frac{F_{pred}}{N_{pred}} \quad (2-41)$$

其中 F_{pred} 为预测错误的车道线数量， N_{pred} 为网络预测出的车道线总数。

$$F_N = \frac{M_{pred}}{N_{gt}} \quad (2-42)$$

其中 M_{pred} 为真实存在但是网络未检测到的车道线数量， N_{gt} 为真实存在的车道线总数。

2.4.3 夜间图像增强结果分析

在本节中对本章夜间图像增强网络部分使用定量和定性的方法进行评估，分别在合成夜间数据集和真实夜间环境进行实验。将该网络与几种经典算法进行对比，MSRCR^[11]、RetinexNet^[14]、EnlightenGAN^[16]，Zero-DCE^[19]。首先使用各算法对 SICE 测试集进行实验，实验结果对比如图 2-12 和图 2-13 所示。

如图 2-12(b)所示，MSRCR 虽然提高了图像的亮度，但是图像中的色彩被完全抹除，增强后的图像完全偏离人的视觉感知，如图像中的天空区域和车道线部分，经过 MSRCR 增强后的图像天空区域呈现灰白色，车道线部分则为淡黄色，与图像本身颜色差异较大。如图 2-12(c)所示，RetinexNet 网络增强后的图像可以直观的看出图像失真，出现了过度增强的现象，图像上有明显的噪声。如图中的建筑物下方阴影部分被过度增强为蓝紫色，左侧树木部分被噪声所覆盖。相同的，在图 2-13(c)中，天空区域的色彩过度饱和，车道线部分已经完全

消失。相比之下，两种无监督增强网络取得了较好的增强效果，如图 2-12(d)和图 2-13(d)所示，虽然 EnlightenGAN 增强的图像整体亮度得到了提高，并且部分细节较完整，但是两张图像的天空区域均出现色彩饱和度过高的情况。而 Zero-DCE 算法增强的图像在边缘细节上的恢复效果较差。总体而言，本章算法对颜色的保持优于其他对比算法，在增强图像整体亮度的同时也保持了边缘细节的完整性，且抑制了图像噪声的产生。

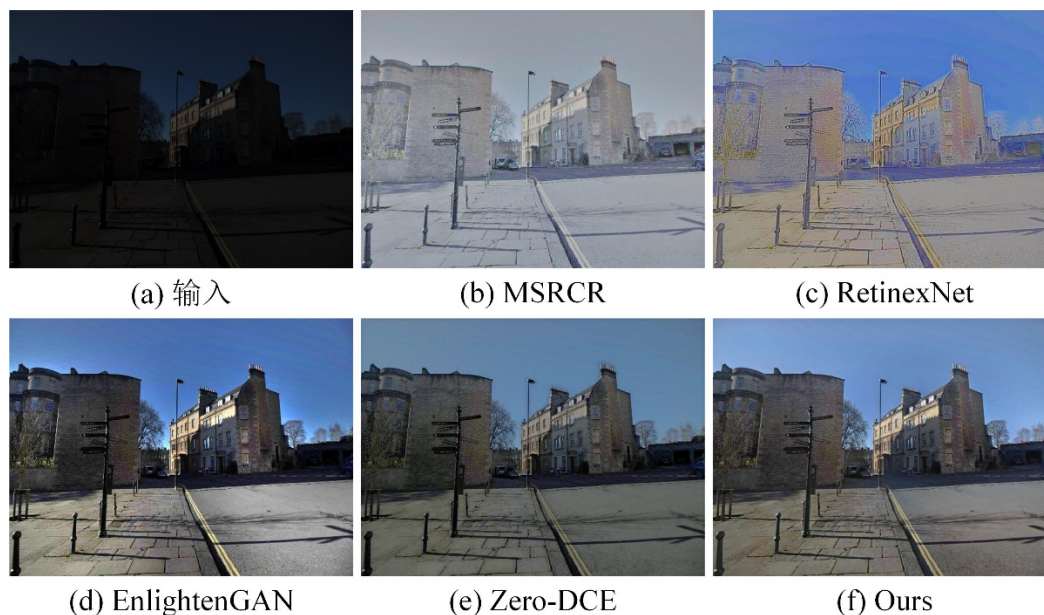


图 2-12 合成数据集场景一实验结果图

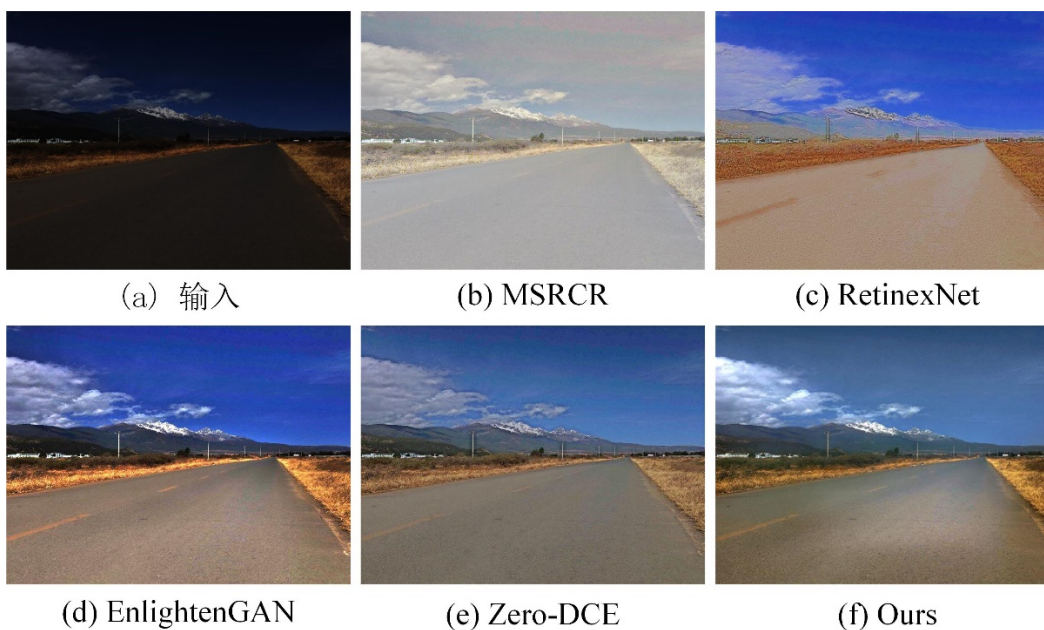


图 2-13 合成数据集场景二实验结果图

为了反映增强网络对真实夜间环境的处理能力，将各算法在 BDD100K 测

试集中重新进行测试，实验结果如图 2-14 和 2-15 所示。MSRCR 和 RetinexNet 算法增强后的图像出现了严重的失真现象，且色彩不均衡，图像细节不清晰，导致车道线部分对比度有所欠缺。如图 2-14(c)所示，RetinexNet 增强后的图像已经无法从中读取到有效的信息，车道线部分无法识别。EnlightenGAN 算法在图像较亮区域会出现过度增强情况，导致车道线被亮光所覆盖，影响车道线检测。如图 2-14(d)中对向行驶的车辆区域，由于车灯处亮度大，导致此区域被过度增强。



图 2-14 真实夜间集场景一实验结果图

在图 2-15(d)的交通灯处可以发现，本章增强图像对亮度较大区域，光晕控制要优于 EnlightenGAN 算法。本章增强算法处理后的图像亮度适中，内容保持较为完整，可以清晰的显示出车道线部分，并且车道线的边缘细节也优于其他算法。

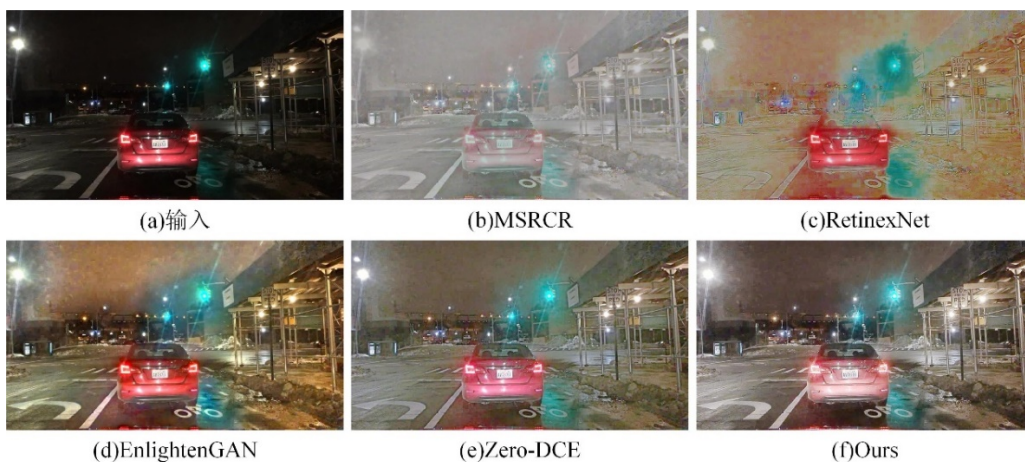


图 2-15 真实夜间集场景二实验结果图

本章使用三个常用的图像评价指标对所选算法和本章增强网络进行定量分析，三个指标分别是 PSNR、SSIM、NIQE，在合成数据集 SICE 的测试集和 LOL 测试集上进行测试，各数据集上的测量结果如表 2-2 所示。其中 PSNR 和 SSIM 为合成数据集上的指标，NIQE 为真实夜间环境下的指标。

表 2-2 各算法在测试集上的定量比较

Method	PSNR(db)	SSIM	NIQE	Time(s)
HE	15.662	0.634	5.721	1.073
MSRCR	14.457	0.617	5.893	1.852
RetinexNet	16.838	0.673	5.603	0.137
EnlightenGAN	19.254	0.817	5.074	0.045
Zero-DCE	18.763	0.815	5.693	0.005
Ours	21.129	0.872	4.917	0.016

通过分析可以得到以下结论：由于本章算法中使用了亮度扰动产生一个异常亮度，以此避免了重建图像带来的损失，增强图像产生的失真较小，因此，PSNR 值高于其他算法。本章方法增强后的图像与参考图像的结构相似度高于其他算法。从时间角度看，本章算法耗费时间少，仅次 Zero-DCE 网络，为后续的车道线识别网络奠定了基础。另外不同于 RetinexNet 和 EnlightenGAN 网络，本章使用的是无监督学习，RetinexNet 网络需要同一场景的不同光照图像，EnlightenGAN 网络需要正常光照图像，因此本章算法比其他算法具有广泛适用性，其可以适应于大部分夜间场景的增强。

2.4.4 车道线检测结果分析

为验证本章车道线检测方法的性能，首先将其与现有其他车道线检测算法一同在 Tusimple 测试集上进行了定量评估。实验中，采用准确率、误检率和漏检率作为性能指标，表 2-3 展示了各车道线检测算法的具体实验结果。

表 2-3 各算法在 Tusimple 测试集的定量比较

Method	Accuracy	FP	FN	Time(s)
UFast	90.73%	0.071	0.034	0.009
SCNN	95.91%	0.063	0.016	0.134
ENet	92.78%	0.093	0.072	0.019
Ours	96.54%	0.059	0.021	0.013

从表 2-3 可以看出，本章提出的车道线检测方法在准确率上表现出明显优势，相较于 SCNN 算法^[53]提高了 0.63%，并与 UFast 算法^[54]相比提升了 5.81%，虽然本算法在时间因素上不及 UFast，但是本算法的漏检率和误检率低于其他

几种算法。

为了充分展示车道线检测网络的检测效果，将其与目前检测精度较高的网络进行比较。包括 SCNN 和 UFast。同时将本章车道线检测网络同使用的骨干网络 ERF-Net 进行对比，以此证明添加自注意力蒸馏的有效性。分别选取 BDD100K 测试集中的不同场景图像作为测试对象，从图 2-16 可以看出，UFast 网络可以检测出部分的车道线，但是没有检测出来最左侧的车道线，并且对车道线末端的检测出现了偏离情况。而骨干网络 ERF-Net 只有最右侧车道线是完整检测出来的，中间车道线出现断断续续的情况，最左侧车道线也出现了偏离。SCNN 网络虽然在路况畅通的情况下检测到了全部的车道线，但是对于最左侧的车道线并没有连续检测出来。相较于其他检测网络，本章车道线检测网络完整的检测出了车道线，并且没有出现错检、误检的状态。

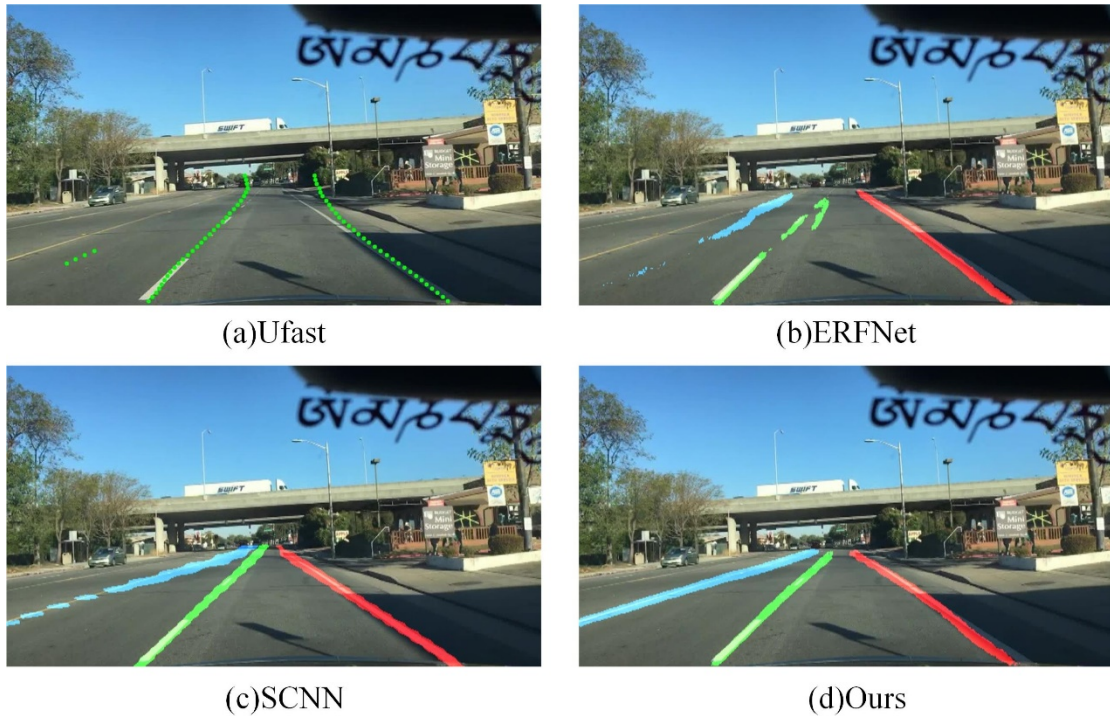


图 2-16 车道线检测实验结果图

本章同时在车辆拥挤的路况下进行实验，实验结果如图 2-17 所示。车辆拥挤会导致车道线被遮挡，车道线检测网络受此影响会降低检测精度。如图 2-17(b)所示，ERF-Net 受到车辆遮挡的影响，完全没有检测到车道线。SCNN 网络只检测出一条车道线。本章网络完整的检测出输入图像中的车道线。

为了凸显出夜间图像增强网络的重要性，对夜间驾驶环境图像和其使用上

述增强网络增强后的图像进行车道线检测，检测结果如图 2-18 所示。从图 2-18(a)和图 2-18(b)中可以看出，增强后的图像车道线信息完整，图像整体质量得到了提升，从而使车道线检测网络可以检测出全部的车道线，而未增强的夜间图像受到夜间环境的影响，车道线检测网络只检测到其中一条车道线。图 2-18(d)和上述类似，车道线部分对比度提升，车道线更加明显，车道线可以得到充分的检测。



图 2-17 拥挤路况车道线检测实验结果图

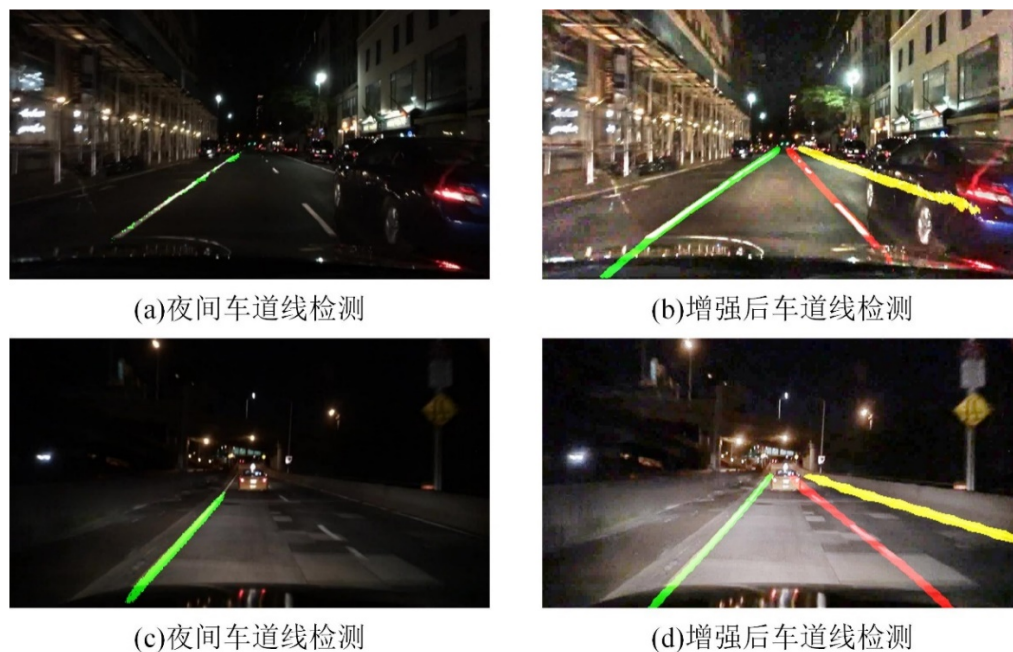


图 2-18 夜间图像增强前后车道线检测实验结果图

2.5 本章小结

在本章中，为解决夜间因光照不足而导致车道线难以识别的问题，将夜间图像增强技术和车道线检测技术结合，构建一种高效的夜间车道线检测算法。该算法首先借助夜间图像增强网络对暗光环境下的道路图像进行亮度与清晰度的提升，随后利用改进的 ERF-Net 模型进行车道线的精确检测。实验结果显示，本章所提夜间增强网络可以在增强图像的同时保留图像细节信息，同时不破坏图像的颜色信息。车道线检测算法整体检测性能优异，特别是在准确率上，相较于其他同类算法取得了显著提升。本章将夜间图像增强网络和车道线检测网络结合进行车道线检测，提升了车道线检测网络在夜间的检测精度。

第3章 基于多层特征融合的雾天车道线检测

3.1 引言

雾天环境对车道线检测和无人驾驶技术都会产生一定的影响。雾中的光线会扭曲车道线的形状，使得车道线检测算法难以准确工作。此外，浓雾中的光线干扰还会增加图像中的噪声，使车道线检测更加困难。因此，采用图像去雾技术提升图像清晰度和视觉质量，旨在增强车道线检测的准确度以及无人驾驶的安全性。

本章鉴于以上问题，设计一种多层特征融合的去雾网络，首先使用 CycleGAN 网络^[66]为主体框架，其中的生成器采用 U-Net 为基础网络，本章的研究路线如图 3-1 所示。

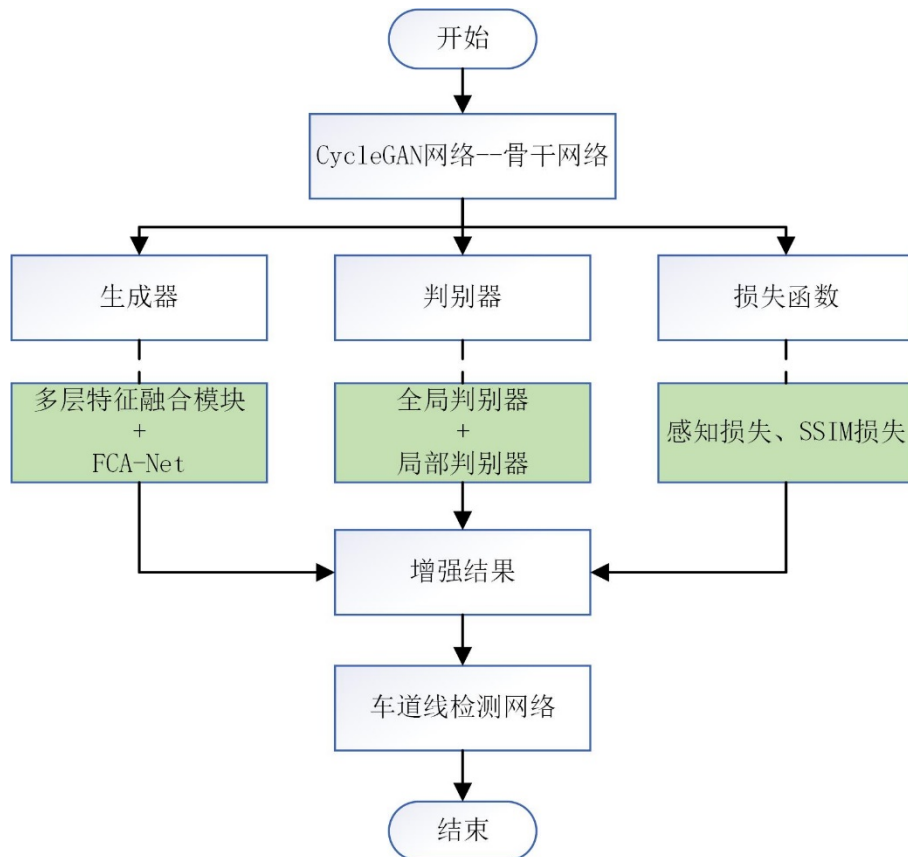


图 3-1 雾天车道线检测研究路线图

为了使编码器的不同层特征能充分与解码层相结合，减少特征的稀释，设计一个多层特征融合模块将编码器的不同层特征与解码层相融合。为提升网络

的去雾能力，受暗通道先验的影响，在生成器中的重要节点添加 FCA-Net 注意力机制^[67]为不同的通道分配不同的权重。为了更好的判断生成器生成的图像效果，防止生成的图像过度偏离原本输入的图像，将初始 CycleGAN 的判别器扩展为两个，一个为全局判别器，在全局层面判断生成的图像是否是生成器生成的，另一个为局部判别器^[68]，对于局部判别器，将输入的有雾图像随机裁剪成四个部分，然后对每个部分进行判别，在计算局部判别器的损失函数时，计算所有块的平均值，这样可以提升判别器的判别能力。将柔性似然估计引入到判别器中，以此提升网络训练的稳定性。在对雾天图像进行去雾后使用第二章设计的车道线检测网络对去雾效果进行检验。

3.2 大气散射模型

在雾天环境下，图像质量会出现显著的退化，主要是由于大气散射和光线衰减的影响。雾天的空气中存在很小的颗粒，如水滴、气溶胶等，它们与光波相互作用，使得光线发生散射。这种散射会导致图像中出现雾化效果，使物体的边界模糊不清、失去细节。雾天环境下的成像原理如图 3-2 所示，在雾天环境下，由于大气中的悬浮微粒会引发光线的散射和吸收，对物体进行光照并反射到成像系统，这导致在成像过程中，大量白色大气光被照射到远处物体表面上，同时物体表面反射光线也会因为大气中悬浮粒子的吸收和散射作用而减弱。因此，雾天环境下成像系统捕捉到的图像往往呈现低对比度和细节模糊等图像退化特征。

Nayer 等人提出了大气散射模型来解决上述问题，大气散射模型的公式如下：

$$I(x) = J(x)t(x) + A(1 - t(x)) \quad (3-1)$$

其中 $I(x)$ 是原始输入的有雾图像， $J(x)$ 是无雾的图像， A 是大气光值， $t(x)$ 是透射率。其中 $t(x)$ 的公式如下：

$$t(x) = e^{-\beta \cdot d(x)} \quad (3-2)$$

其中 β 为大气散射系数， $d(x)$ 为场景深度。

由于原始的有雾图像是已知的，所以利用大气散射模型去除雾霾就是通过

各种先验知识或方法求解式中的 A 和 $t(x)$ 。

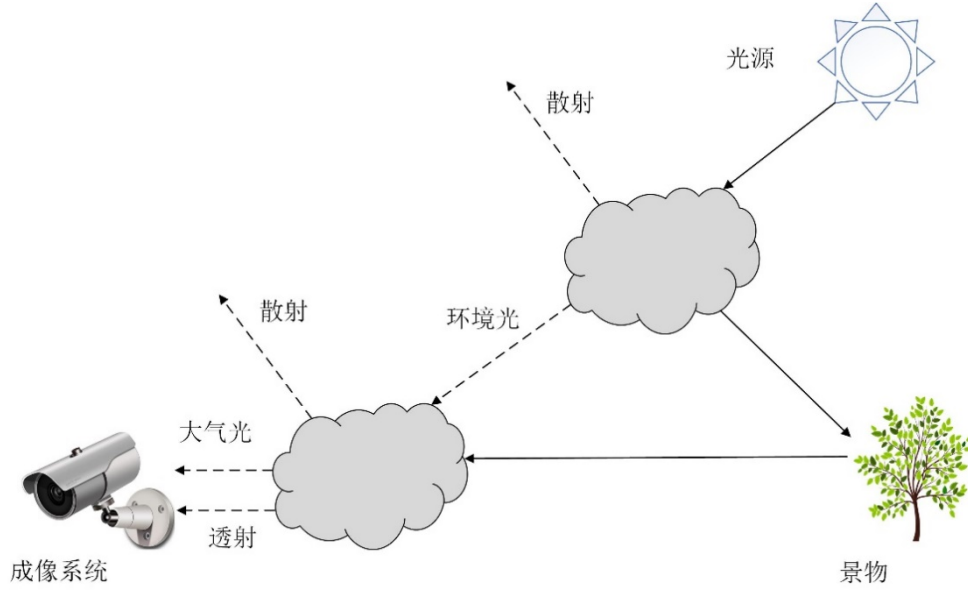


图 3-2 雾天成像原理图

暗通道先验去雾便是其中最为经典的算法。He 等人通过大量的观察与实验得出一个结论：在一张室外无雾图像中，除去天空的其他区域，RGB 三通道中至少存在一个通道的值接近于 0。该理论的公式如下：

$$J^{dark}(x) = \min_{y \in \Omega(x)} \left(\min_{c \in \{r, g, b\}} J^c(y) \right) \quad (3-3)$$

其中 $J^c(y)$ 是无雾图像中 RGB 某一个通道在像素点 y 的值， $\Omega(x)$ 表示以 x 为中心的区域。

具体的计算过程本章不做详述，暗通道先验对于设计去雾网络而言，提供了一个强有力的思路，去雾效果和雾天图像的通道有较大的关联。

3.3 循环生成对抗网络

GAN 网络本质上由两大部分构成，即生成器模块与判别器模块，通过同时训练两者提升所生成图像的质量。GAN 网络的结构如图 3-3 所示。生成器的目标是学习如何生成与训练数据相似的样本。判别器的主要工作是对输入的样本进行辨别。通过不断的训练和优化，判别器得以提升其鉴别能力，精确地区分开生成器生成的图像与真实的图像。通过让两者对抗地进行训练以提高二者的能力，最终达到一个动态平衡，生成器可以生成高质量的图像，而判别器可以

准确地判断真伪。

GAN 网络对于输入为随机图像时，网络也会生成随机的图像，为了使网络生成的图像可以与输入图像相对应，其训练过程需要成对的数据集进行训练。对于雾天环境，要获取其同场景的对应无雾图像难度较大，因此许多研究者使用 CycleGAN 作为去雾网络的基础框架。CycleGAN 可以将不同域之间进行迁移，对于去雾来说，CycleGAN 将图像从有雾风格域转换成无雾风格。CycleGAN 由两个 GAN 网络组成，通过两个双向循环对生成器进行约束，配合两个判别器使生成器生成逼真的预期图像。

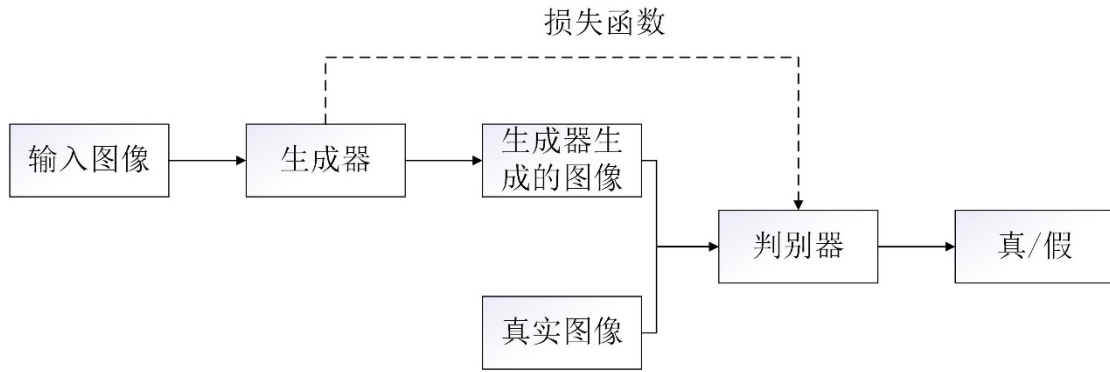


图 3-3 生成对抗网络结构图

3.4 FCA 注意网络

He 等人^[21]提出的暗通道先验理论说明了去雾过程中对通道处理的重要性，除此之外，真实雾天环境不会像合成雾天图像一样，真实雾天的雾分布不均匀，需要对有雾的地方进行去雾处理，而对无雾区域或者雾度较薄的区域分配较低的权重，以此增强网络的去雾效果。目前大部分研究已经使用注意力机制增强去雾效果，常规的注意力机制采用全局平均池化的方法将特征图在通道维度压缩为一个标量，这样就会导致网络无法充分利用输入的特征，对于不同的输入特征在通过全局平均池化后可能会得到相同的标量，在实际图像中，不同的通道包含了不同的语义信息。为了充分利用不同通道之间的信息，FCA 注意力网络使用了离散余弦变换。离散余弦变换的公式如下：

$$f_{h,w}^{2d} = \sum_{i=0}^{H-1} \sum_{j=0}^{W-1} x_{i,j}^{2d} \cos\left(\frac{\pi h}{H}\left(i + \frac{1}{2}\right)\right) \cos\left(\frac{\pi w}{W}\left(j + \frac{1}{2}\right)\right) \quad (3-4)$$

其中 f^{2d} 是二维离散余弦变换的频谱， x 为输入的图像， W 和 H 分别是输入图

像的宽度和高度。

FCA 注意力网络通过多谱注意力模块引入了不同的频率分量，多谱注意力模块将输入在通道维度划分成 n 个部分，然后为每一部分分配频率分量。每个部分的输出如公式 3-5 所示。

$$Freq^i = 2DDCT^{u,v} \left(X^i \right) = \sum_{h=0}^{H-1} \sum_{w=0}^{W-1} X_{h,w}^i B_{h,w}^{u,v} \quad (3-5)$$

其中 B 为二维离散余弦变换函数，对每个部分使用不同的频率分量，然后将每个部分的频率分量拼接得到多谱向量，对多谱向量使用全连接层建立权重，最后使用 Sigmoid 函数进行归一化得到结果。

FCA-Net 从频域的角度切入，这样会填充现有通道注意力方法中特征信息的不足，用离散余弦变换引入更多的频域分量。FCA-Net 使网络可以自适应的从特征图中学习权重，提升网络对不同雾浓度的处理。在网络的下采样层和上采样层均加入 FCA-Net，并使用 FCA-Net 在多层特征融合模块中融合多层特征。

3.5 改进的 CycleGAN 去雾网络

本章去雾网络基于循环生成对抗网络进行设计，网络的总体框架如图 3-4 所示。该网络由 2 个生成器和 4 个判别器组成， G_A 的作用是将有雾图像转换为无雾图像， G_B 的作用是将无雾图像映射为有雾图像。判别器 D_A^{global} 从全局的角度判断图像是生成器 G_A 生成的还是真实输入的无雾图像。判别器 D_A^{Local} 将图像分为 4 个块，并对每个块进行判断。同理，判别器 D_B^{global} 和 D_B^{Local} 判断雾天图像。网络的整体流程是首先将非成对的雾天图像和无雾图像输入到网络中。生成器 $fake_B$ 学习有雾图像到无雾图像的映射关系，生成无雾图像，然后生成器 G_B 学习无雾图像到有雾图像之间的关系，将无雾图像重新转换成有雾图像 rec_A 。判别器 D_A^{Local} 和判别器 D_A^{global} 判断 $fake_B$ 是生成器生成的假图像还是真实图像。输入图像和重建图像 rec_A 通过计算循环一致性损失约束网络。另一个循环分支与上述同理，不再详述。通过生成器和判别器之间的博弈训练使图像去雾效果更

好。

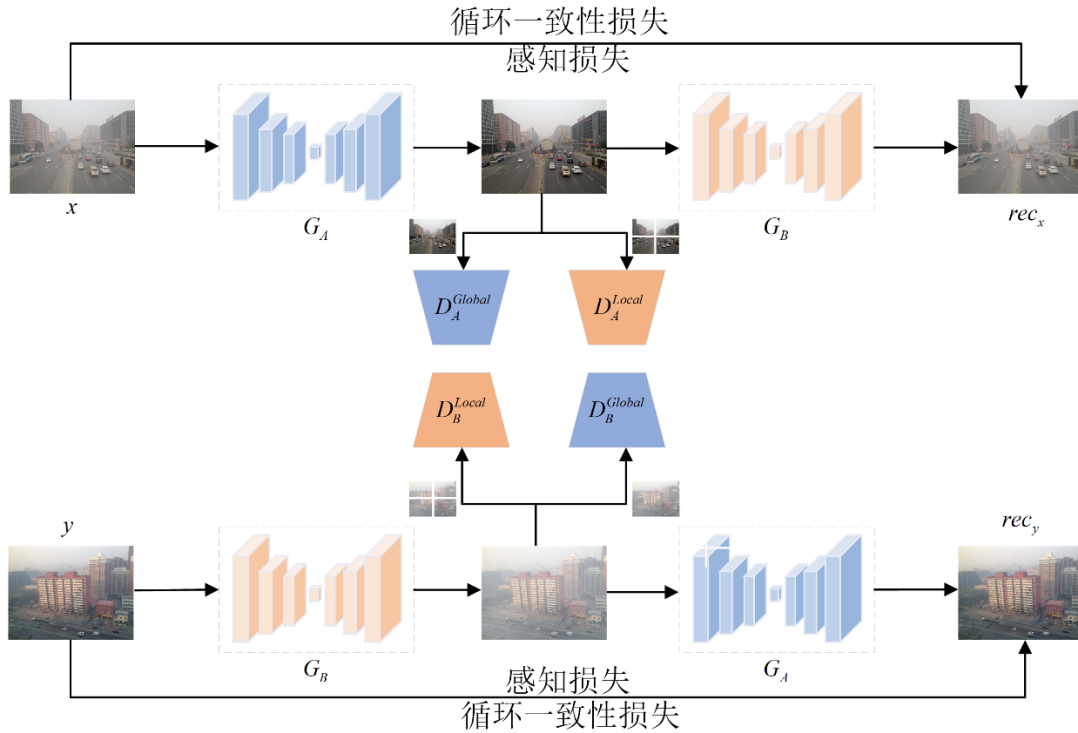


图 3-4 网络整体结构图

3.5.1 多层特征融合模块

网络的生成器采用 U-Net 为骨干网络，U-Net 网络是将编码层和解码层相应尺度的特征连接在一起，这样虽然可以引入编码层的特征信息，但是忽略了不同尺度的编码层所包含的信息。在 U-Net 网络中，浅层的编码层提取的特征由于和输入比较近，能够保留丰富的颜色、纹理及边缘等详细特征，因其感受野较小且重叠区域有限，所以可以保证网络获得更多的细节。随着深度加深，编码层的感受野逐渐扩大，相邻感受野的重叠也随之增多，此时网络能够捕获更多关于图像整体结构和高级属性的信息，如语义内容和全局亮度等。解码层在恢复图像时是将深层的编码层和浅层的解码层连接在一起的，浅层的解码层和深层的编码层相同，都缺少纹理和边缘等细节信息，随着上采样过程，语义和亮度信息也会逐渐变少，导致最终恢复出的结果较差。

为了解决以上问题，在编码层的三个下采样处引入多层特征融合模块，对于三个不同尺寸特征图使用不同的融合模块，使解码层可以融合不同层级的编码层特征信息。生成器网络结构如图 3-5 所示。

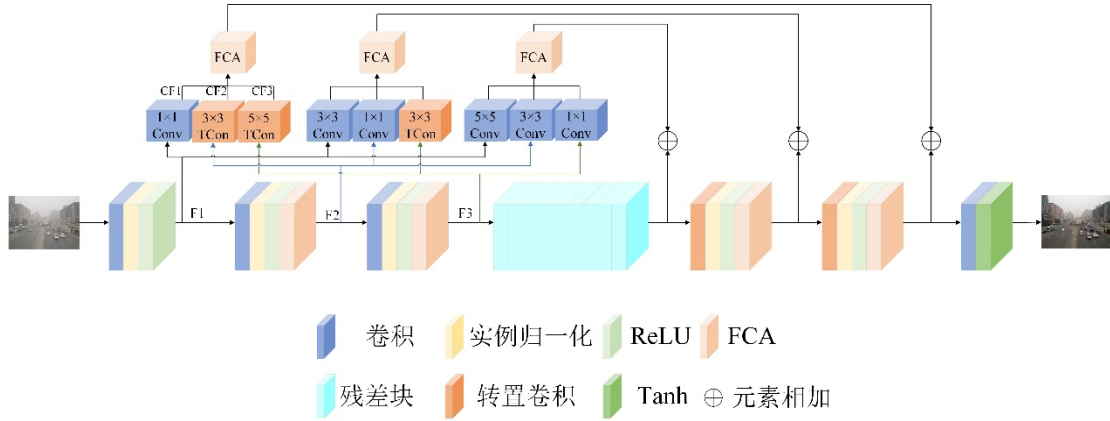


图 3-5 生成器网络结构图

编码层有三个下采样层，通过三个下采样层提取不同层的特征。解码层包含两个上采样层和一个输出层。假设三个编码层输出的特征图分别为 $F1$, $F2$, $F3$ 。此时三个特征图的尺寸不一致，包含的内容信息也不相同，这时将 $F1$ 通过 1×1 的卷积核保持 $F1$ 的尺寸不变， $F2$ 通过 3×3 的转置卷积将分辨率调整到和 $F1$ 一致， $F3$ 通过 5×5 的转置卷积将分辨率调整到和 $F1$ 一致，每个卷积和转置卷积后都添加 LeakyReLU 增强非线性。将卷积和转置卷积后的结果相加并输入到 FCA-Net 中，为每个通道分配权重，融合上述特征得到 $CF1$ ，此时再将 $CF1$ 与解码层对应尺度的特征 $DF1$ 相融合，最后通过上采样恢复去雾图像。剩下的 $CF2$ 和 $CF3$ 只有不同特征图的卷积和转置卷积不同，其他与上述相同。这样解码层便可以融合不同编码层的特征，使图像恢复可以包含更多的细节和边缘信息。

3.5.2 轻量化的残差模块

编码层和解码层中间的残差网络是用来深度提取特征的，残差网络由 9 个残差块组成，可以提取到更加复杂的特征。通常去雾网络中使用的残差块较为简单，且叠加多次后参数量会变大，因此重新设计残差块，残差块的结构如图 3-6 所示，将残差网络中的卷积拆分成 3 部分，首先使用 1×1 的卷积将输入图像的通道数量变少，使用 1 个卷积核为 3 的卷积提取特征，然后使用 1 个卷积核为 3 的卷积将特征图尺寸进行还原，这样将大大减少网络模型的参数量。最后为了确保深层特征也可以注意到图像中的有雾区域，在残差块的尾部添加 FCA-Net，使网络在深层特征处为有雾通道附加权重。

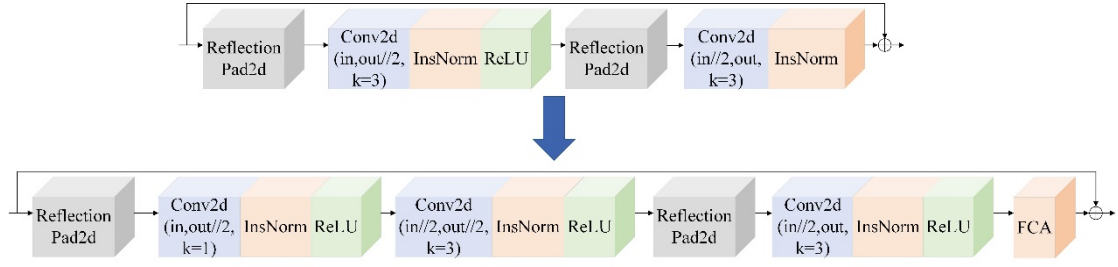


图 3-6 残差网络结构图

3.5.3 全局判别器和局部判别器

传统的 CycleGAN 使用单个判别器去判断生成器生成的图像，是从全局的角度判断图像，对于局部信息变化的处理能力较弱，如雾浓度较大的图像或者雾分布不均匀的图像，因此使用两个判别器判断生成器生成的图像。全局判别器和局部判别器均使用 PatchGAN^[69]作为基础网络架构，PatchGAN 判别器相对于正常 GAN 网络的判别器输出的是一个矩阵的平均值。包含 4 个 4×4 的卷积，输出通道从 64 到 512，最后一个卷积的输出通道为 1。为了提升 PatchGAN 的鉴别性能，使用谱归一化卷积^[70]代替原先的普通卷积，在每个谱归一化卷积后添加 ReLU 非线性函数，最终使用 Sigmoid 激活函数输出一个 64×64 的块，使用这些块的值去判断输入的图像是生成器生成的还是真实图像。对于全局判别器 D_A^{global} ，直接将生成器生成的图像输入到全局判别器中，全局判别器是用来保证生成的去雾图像在整体语义上的准确性和连贯性，使重建后的图像信息与上下文保持一致。对于局部判别器 D_A^{local} ，首先将输入图像随机裁剪成四部分，即 4 个 128×128 的块。局部判别器作用在雾浓度分布不同的区域，可以判断出局部纹理、颜色等特征恢复的是否合理。

在局部判别器的输出端直接计算所有块的平均值会影响最终判别的效果，使用柔性似然估计^[71]计算局部判别器的输出。使用柔性似然估计可以加强判别器对图像细节信息的关注，从而提高图像去雾的质量。首先计算每个块和概率最小块似然的接近程度来获得权重矩阵。使用 SoftMax 和 SoftMin 函数求解权重矩阵的最大化目标函数和最小目标函数。然后将每个块的权重矩阵乘以对应的似然度，最后相加得到最终的输出。这样处理后判别器会更加关注低概率区域的变化，生成器就会生成细节更为丰富、较为真实的图像。计算公式如下：

$$P = \sum_i (W_i \times \ln l_i) \quad (3-6)$$

$$W_i = \frac{e^{\alpha \times \ln l_i}}{\sum_i e^{\alpha \times \ln l_i}} \quad (3-7)$$

其中 P 是判别器的输出, i 是图像被分成的每个块, W_i 是权重矩阵, α 为 SoftMax 和 SoftMin 函数, 将 α 设置为 1 计算 SoftMax 函数, α 设置为-1 计算 SoftMin 函数。

3.5.4 损失函数

原始的 CycleGAN 只设置了三个损失函数, 但是经过实验发现, 网络训练后图像风格迁移的比较严重, 一些图像信息直接被破坏, 如天空区域出现了楼宇和草等等, 为了避免这种情况, 在本章的损失函数中对其进行重新设计, 使用对抗损失、身份损失、循环一致性损失、感知损失、颜色损失、SSIM 损失和总变分损失训练网络。

1) 对抗损失: 对抗损失让生成器和判别器进行博弈学习。采用最小二乘损失来计算, 以此匹配生成图像的和目标无雾图像之间的分布。对抗损失的公式如下:

$$I_{adv}^{global}(G_A, D_A^{global}, x, y) = (D_A^{global}(G_A(x)) - 1)^2 + (D_A^{global}(y) - 1)^2 + (D_A^{global}(G_A(x)))^2 \quad (3-8)$$

$$L_{adv}^{global}(G_B, D_B^{global}, x, y) = (D_B^{global}(G_B(y)) - 1)^2 + (D_B^{global}(x) - 1)^2 + (D_B^{global}(G_B(y)))^2 \quad (3-9)$$

$$L_{adv}^{local}(G_A, D_A^{local}, x_p, y_p) = (D_A^{local}(G_A(x_p)) - 1)^2 + (D_A^{local}(y_p) - 1)^2 + (D_A^{local}(G_A(x_p)))^2 \quad (3-10)$$

$$L_{adv}^{local}(G_B, D_B^{local}, x_p, y_p) = (D_B^{local}(G_B(y_p)) - 1)^2 + (D_B^{local}(x_p) - 1)^2 + (D_B^{local}(G_B(y_p)))^2 \quad (3-11)$$

$$L_{adv} = L_{adv}^{global}(G_A, D_A^{global}, x, y) + L_{adv}^{global}(G_B, D_B^{global}, x, y) + L_{adv}^{local}(G_A, D_A^{local}, x_p, y_p) + L_{adv}^{local}(G_B, D_B^{local}, x_p, y_p) \quad (3-12)$$

其中 x 是有雾图像, y 是无雾图像, x_p 和 y_p 是随机裁剪后有雾和无雾图像。

G_A 是将有雾图像转换为无雾图像的生成器, G_B 是将无雾图像转换为有雾图像

的生成器。 D_A^{global} 和 D_A^{local} 分别是针对生成器 G_A 的全局和局部判别器， D_B^{global} 和 D_B^{local} 分别是针对生成器 G_B 的全局和局部判别器。

2) 身份损失：身份损失可以增强去雾图像的一些细节信息，使网络输出的图像和输入图像在结构上保持一致。该损失函数定义如下：

$$L_{id} = E_{y \sim P_{data}(y)}(\|G_A(y) - y\|_1) + E_{x \sim P_{data}(x)}(\|G_B(x) - x\|_1) \quad (3-13)$$

其中 $\|\cdot\|_1$ 表示 L1 范数。

3) 循环一致性损失：CycleGAN 网络之所以不需要成对的雾天图像和无雾图像就是因为有循环一致性损失，该损失可以使生成器生成的图像保留原始图像的信息，并且防止模式崩溃的发生，避免了生成器总是生成同一个可以欺骗判别器的图像。循环一致性损失的定义如下公式：

$$L_{cyc} = E_{x \sim P_{data}(x)}(\|G_B(G_A(x)) - x\|_1) + E_{y \sim P_{data}(y)}(\|G_A(G_B(y)) - y\|_1) \quad (3-14)$$

其中 $G_B(G_A(x))$ 和 $G_A(G_B(y))$ 分别是重建后的有雾图像和无雾图像。循环一致性损失的目标就是输入的有雾图像或无雾图像经过网络的一个循环后仍能恢复至初始的图像，即 $G_B(G_A(x)) \sim x$ 和 $G_A(G_B(y)) \sim y$ 。

4) 感知损失：感知损失可以从有雾图像中恢复纹理信息，并且可以保持输入图像整体结构和一些细微细节信息。和传统的均方误差函数相比，感知损失更加注重图像的感知质量。感知损失通过预训练的神经网络来计算两张图像之间的差异。感知损失的计算公式如下：

$$L_{per} = \|\phi(G_B(G_A(x))) - \phi(x)\|_2 + \|\phi(G_A(G_B(y))) - \phi(y)\|_2 \quad (3-15)$$

其中 $\|\cdot\|_2$ 表示 L2 范数， ϕ 表示使用在 ImageNet 上预训练的 VGG16 模型的第 2 个和第 5 个池化层提取特征图操作。由于本章网络是无监督学习网络，没有成对的雾天和无雾数据，因此将重建后的无雾图像 $G_A(G_B(y))$ 和真实的无雾图像 y 进行比较。

5) 颜色损失：使用颜色损失来比较去雾后的图像和真实输入的无雾图像之间的颜色差异，从而使去雾后的图像更加接近真实图像。由于有雾图像的对比度通常比无雾图像低，因此先将去雾后的图像和真实无雾图像都用高斯滤波器

变得模糊，这样只考虑颜色信息的变化，再用均方误差计算差值。颜色损失的公式如下：

$$L_{col} = \|(G_A(x))_{blur} - y_{blur}\|_2 + \|(G_B(y))_{blur} - x_{blur}\|_2 \quad (3-16)$$

其中 $(G_A(x))_{blur}$ 是去雾后图像通过高斯滤波器后的模糊图像， y_{blur} 是真实无雾图像通过高斯滤波器后的模糊图像。 $(G_B(y))_{blur}$ 是有雾图像通过高斯滤波器后的模糊图像， x_{blur} 是真实雾天图像通过高斯滤波器后的模糊图像。

6) SSIM 损失：为了使生成图像在结构相似性和对比度相似性上保持一致，添加 SSIM 损失函数，SSIM 损失的公式如下：

$$L_{ssim} = \left\| \begin{aligned} &(1 - ssim(\phi(G_B(G_A(x))), \phi(x))) + \\ &(1 - ssim(\phi(G_A(G_B(y))), \phi(y))) \end{aligned} \right\| \quad (3-17)$$

$$ssim(i_1, i_2) = \frac{(2u_{i_1}u_{i_2} + C_1)(2\sigma_{i_1}\sigma_{i_2} + C_2)}{(u_{i_1}^2 + u_{i_2}^2 + C_1)(\sigma_{i_1}^2 + \sigma_{i_2}^2 + C_2)} \quad (3-18)$$

其中 i_1, i_2 表示预训练 VGG16 模型的第 2 个和第 5 个池化层对重建后的图像和真实图像提取的特征图。 u 表示取平均值操作， σ 表示取方差操作。

7) 总变分损失：为了促进网络输出结果的空间平滑性，减少生成图像的噪声，在损失函数中加入总变分损失。总变分损失的定义如下：

$$L_{tv} = \|\nabla_x G_A(x) + \nabla_y G_A(x)\|_1 + \|\nabla_x G_B(y) + \nabla_y G_B(y)\|_1 \quad (3-19)$$

其中 ∇_x 和 ∇_y 表示水平和垂直方向上的微分运算。

总体损失函数如下：

$$L = \lambda_1 L_{adv} + \lambda_2 L_{id} + \lambda_3 L_{cyc} + \lambda_4 L_{per} + \lambda_5 L_{col} + \lambda_6 L_{ssim} + \lambda_7 L_{tv} \quad (3-20)$$

其中 λ 表示各损失的权重。

3.6 实验结果及分析

本节将介绍网络训练及测试时所需的雾天图像数据集，此外将介绍实验平台及训练过程参数设置。在本节中，将本章方法和其他几种去雾算法进行比较，分别在合成的雾天数据集和真实雾天数据集进行测试，由于全参考评价指标需

要成对的图像数据进行计算，所以采用全参考图像质量评价指标和无参考图像质量评价指标对去雾结果进行定量评价。其中全参考图像质量评价指标选取 PSNR 和 SSIM，用于评价去雾方法在合成数据集上的表现。无参考评价指标选取 NIQE、雾感知密度评估(Fog Aware Density Evaluator, FADE)^[72]和无参考图像空间质量评估(Blind/Referenceless Image Spatial Quality Evaluator, BRISQUE)^[73]，用于评价真实雾天环境下的去雾效果。PSNR 和 SSIM 的值越高越好，无参考的三个图像质量评价指标值越小越好，生成的雾天图像质量越高。

3.6.1 实验设置

本章网络是基于 CycleGAN 进行设计的，是不需要成对雾天数据进行训练的，同时为了增强本章网络对真实雾天环境的泛化能力，所以选择真实雾天环境下的图像进行训练。目前雾天数据集较大且常用的是(Realistic Single Image Dehazing, RESIDE)数据集^[74]，RESIDE 数据集包含户外训练集(Outdoor Training Set, OTS)和真实世界任务驱动测试集等子集。从真实世界任务驱动测试集(Real-world Task-Driven Testing Set, RTTS)和无标签的真实雾天图像(Unannotated Real-world Hazy Images, URHI)中分别选取 3000 张雾天图像作为训练集的雾天图像。对于训练集中的无雾图像，由于不需要成对的数据，所以从 OTS 中随机选择 6000 张无雾图像进行训练。对于网络评估的测试集，选取 SOTS、HSTS 和 URHI 的部分图像作为测试集。

本章实验是在拥有 Intel(R) Xeon(R) Platinum 8255C CPU 和 NVIDIA RTX3090 GPU 的设备上进行的，显存为 24G，操作系统为 Ubuntu。编程语言使用 Python3.8，实验软件采用 PyCharm，深度学习基础框架为 PyTorch。由于输入的雾天图像数据和无雾图像数据可能出现尺寸不匹配的情况，因此在数据输入端口将图像统一裁剪为 256×256 大小，裁剪方式为随机裁剪。网络使用 ADAM 优化器训练生成器和判别器，其中的动量参数 β_1 和 β_2 分别设置为 0.5 和 0.999。初始学习率设置为 2×10^{-4} ，batch-size 设置为 2，当网络损失值接近最小值时，学习率应该变得更小使模型接近这一点，所以使用余弦退火衰减学习率，余弦函数半周期设置为 5。网络总共训练 200 轮，由于 CycleGAN 网络训练时不容易稳定，因此在训练的前 100 轮固定学习率不变，在后 100 轮对学习率进

行衰减，直至衰减到 0。根据 CycleGAN 的训练经验，对抗损失和循环一致性损失的权重比保持在 1:10 时，网络训练可以达到较好的效果，因此对于损失函数的各项权重，分别设置为 $\lambda_1=1$ ， $\lambda_2=5$ ， $\lambda_3=10$ ， $\lambda_4=5$ ， $\lambda_5=0.5$ ， $\lambda_6=0.5$ ， $\lambda_7=0.5$ 。

3.6.2 合成雾天数据集上的实验结果分析

将本章网络去雾效果和其他几种较为经典、去雾效果较好的方法从定量和定性的角度进行比较。选用 DCP^[21]、AOD-Net^[28]、GCA-Net^[29]、CycleGAN^[66]、FFA-Net^[30]、Cycle-Dehaze^[31]和 EPDN^[75]等去雾方法。

首先将去雾方法在合成的户外测试集 SOTS 上进行比较，DCP 去雾经常会导致去雾不彻底，其去雾效果较差，对于天空区域的雾处理尤其较弱，天空区域还会出现颜色失真。如图 3-7(b)和图 3-8(b)所示。天空区域的颜色已经偏离真实图像颜色。同样的，AOD-Net 和 FFA-Net 等方法虽然提升了图像整体的去雾效果并且解决了颜色失真的问题，但是图像部分区域的雾并没有完全去除。如图 3-7(c)的道路中央区域。EPDN 采用了多尺度的生成器和判别器进行去雾，达到很好的去雾效果，但是从图上可以看出，对于天空区域均出现了颜色失真的情况，而且去雾后的图像整体偏暗，不利于后续的图像处理任务。GCA-Net 去雾效果较好，但是其同样使去雾后的图像变暗，如图 3-8(g)所示，建筑物的边缘处颜色变化较大，图像处理的不够平滑，且天空区域颜色饱和度较高，图像质量明显变差。CycleGAN 网络虽然不需要成对的数据，但是会破坏图像的细节信息，且去雾效果也不理想，如图 3-7(f)所示，墙壁上的文字信息和树木分支等信息在网络输出后被破坏，图像较深处的雾没有被清除。相比之下，本章网络可以有效的去除图像中的雾，且可以很好的还原物体原先的颜色等信息，对于细节上的恢复也优于其他几种方法。

对于雾天图像增强效果的定量评价除了第二章的 PSNR、SSIM 和 NIQE 以外，本章添加了两个无参考图像评价指标，分别是雾感知密度评估(FADE)和无参考图像空间质量评估(BRISQUE)，两者对雾天图像的增强效果有着更加准确的评价效果。FADE 通过将特征信息与密度估计模型相结合，可以更好地区分雾天图像中的目标信息和雾霾引起的噪声，从而提高密度估计的准确性和鲁棒

性。BRISQUE 通过对图像的统计分析来评估图像的质量。

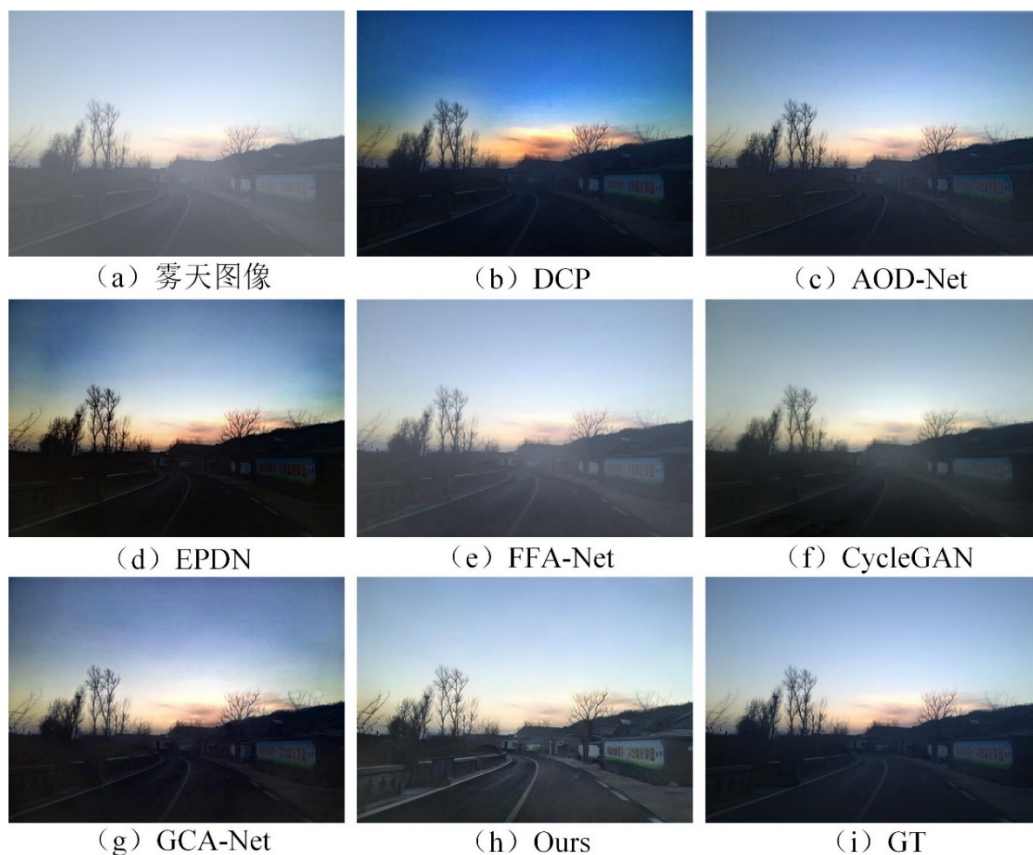


图 3-7 不同算法在 SOTS 上场景一的去雾结果

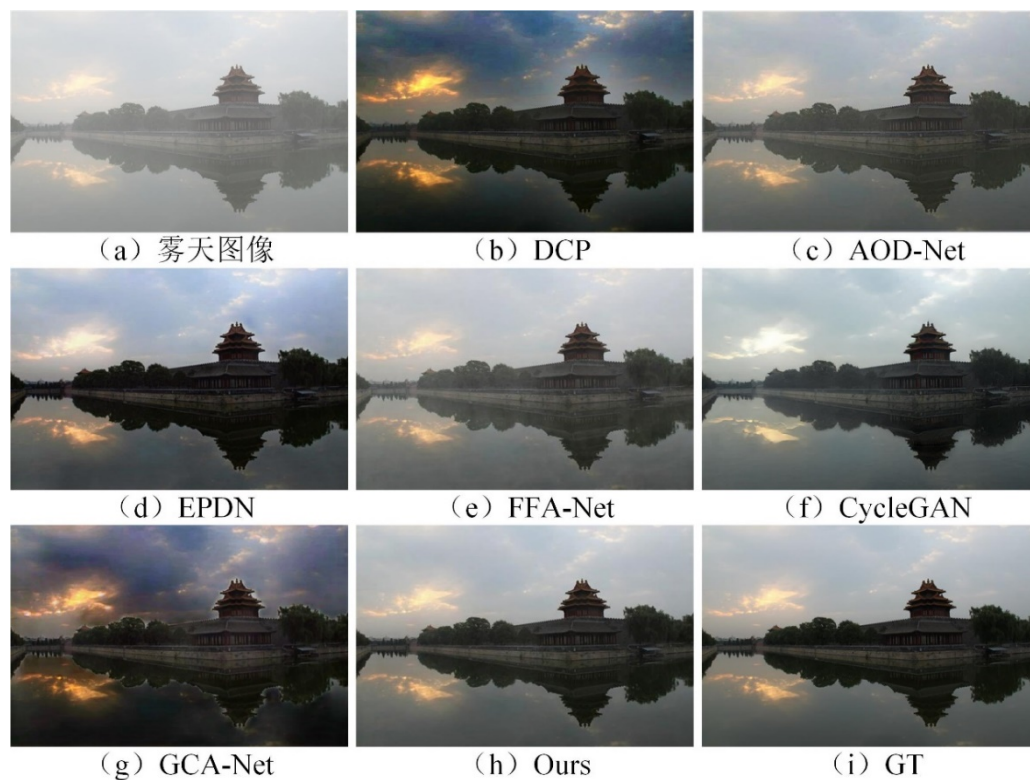


图 3-8 不同算法在 SOTS 上场景二的去雾结果

从定量的角度来看，对所有方法在户外测试集上计算 PSNR 和 SSIM 值，表 3-1 是各种算法在合成数据集 SOTS 和 HSTS 上的定量对比。通过比较可以发现本章方法在 HSTS 数据集和 SOTS 数据集上均优于其他方法。与 AOD-Net、FFA-Net 和 EPDN 等方法相比，本章方法的 PSNR 值分别提高了 2.74dB、2.11dB 和 0.53dB。其他所选算法很多是基于 U-Net 进行设计的，但是这些方法并没有完全使用下采样时不同层的特征，使网络输出的图像在结构和细节上恢复的较差，本章网络充分考虑了不同尺度的特征，且在损失函数中还使用 SSIM 损失函数约束网络，所以本章网络的 SSIM 值要高于其他算法。与 AOD-Net、FFA-Net 和 EPDN 等方法相比，本章方法的 SSIM 值分别提高了 0.077、0.046 和 0.03。

表 3-1 各算法在合成数据集 SOTS 及 HSTS 上的定量比较 (PSNR/SSIM)

Method	Type	SOTS		HSTS	
		PSNR(db)	SSIM	PSNR(db)	SSIM
DCP	Prior	16.77	0.763	16.83	0.772
MSCNN	Supervised	19.06	0.826	18.25	0.803
AOD-Net	Supervised	19.75	0.839	19.59	0.828
FFA-Net	Supervised	20.38	0.861	19.67	0.859
EPDN	Supervised	21.96	0.866	20.17	0.875
GCA-Net	Supervised	21.48	0.843	21.06	0.872
CycleGAN	Unsupervised	18.05	0.722	17.52	0.703
Cycle-Dehaze	Unsupervised	19.11	0.807	18.46	0.797
Ours	Unsupervised	22.49	0.895	23.32	0.905

除了使用全参考图像质量评价指标，本章还使用了无参考图像质量评价指标，表 3-2 是各算法在合成数据集上的无参考比较结果。通过比较可以发现，本章网络的 NIQE 值和 BRISQUE 值均低于其他方法，FADE 的值仅比 Cycle-Dehaze 高 0.007。对比表中的数据可以看出本章网络增强后的图像更加贴近现实中的图像。

表 3-2 各算法在合成数据集上的定量比较 (NIQE/FADE/ BRISQUE)

Method	Type	NIQE	FADE	BRISQUE
DCP	Prior	3.992	0.504	15.358
AOD-Net	Supervised	3.768	0.746	18.065
FFA-Net	Supervised	4.025	0.707	17.632
EPDN	Supervised	3.206	0.866	14.729
GCA-Net	Supervised	3.196	0.611	15.012
CycleGAN	Unsupervised	4.137	0.513	16.183
Cycle-Dehaze	Unsupervised	3.542	0.417	19.762
Ours	Unsupervised	3.074	0.424	14.113

3.6.3 真实雾天数据集上的实验结果分析

本章网络是无监督学习模型，训练是在非成对雾天数据下进行的，并且使用的是真实雾天环境下的数据，因此本章方法对真实世界的泛化能力较强。本章同时在真实环境的数据集上进行了测试，使用的测试数据集是无标签真实雾天图像，使用的测试图像均是除去训练图像以外的图像。DCP 算法虽然可以去除图像中部分区域的雾，但是天空区域颜色严重失真且大部分图像都出现了棋盘格伪影，如图 3-9(b)所示，天空区域出现红色和紫色斑块，建筑物附近边缘特征破损严重。AOD-Net 可以去除图像近景部分的雾，对于图像较深的地方，去雾能力较弱，如图 3-10(c)所示，近景的车道部分和楼宇建筑都比较清晰，但是远景部分的天空和建筑物都被雾笼罩，去雾效果很差，这是因为 AOD-Net 是轻量化网络，且使用的是合成数据对进行训练的，但是真实环境下雾是不均匀的，这就导致 AOD-Net 对图像不同区域的处理程度不同。从图 3-9 上可以发现，FFA-Net 对真实雾天环境的处理很不理想，FFA-Net 同 AOD-Net 相似，都是在合成数据集上进行训练的，这就导致网络对真实世界雾天环境的泛化能力较差。如图 3-9(e)所示，桥体架构和天空区域的雾均没有被消除。

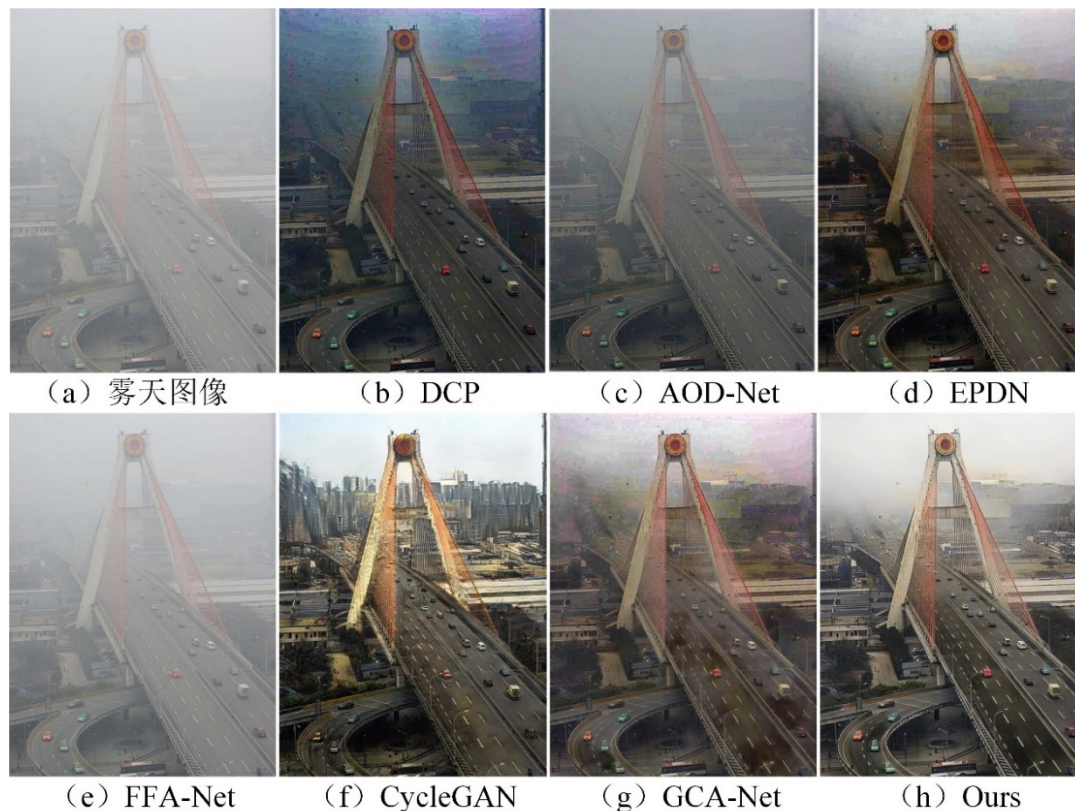


图 3-9 不同算法在真实雾天场景一的去雾结果

EPDN 和 GCA-Net 虽然部分图像的去雾效果较好，但是和在合成数据集上的处理结果相似，都对图像颜色的恢复较差，如图 3-10(d)和图 3-10(g)所示，原先被雾掩盖较深的楼在恢复后，图像的细节信息损失严重，天空区域颜色失真且出现色块。CycleGAN 网络虽然是无监督学习方法，但是其无法控制风格迁移的程度，如图 3-9(f)所示，在桥体附近雾霾浓度较高的区域，网络风格迁移错误，将浓雾区域都转换为了楼宇。相较而言，本章网络对于真实雾天环境的泛化能力较强，对于细节和颜色的恢复也优于其他方法，对于浓雾区域的处理也较好。

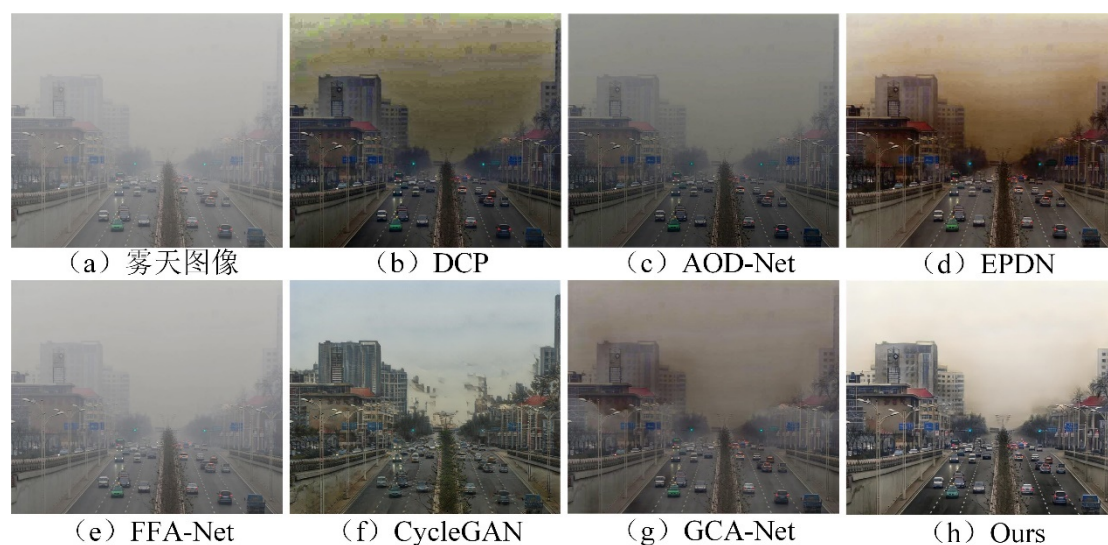


图 3-10 不同算法在真实雾天场景二的去雾结果

本章同时也在非均匀真实环境数据集 NH-Dehaze 上进行了实验，图 3-11 是各算法在 NH-Dehaze 上的实验结果。

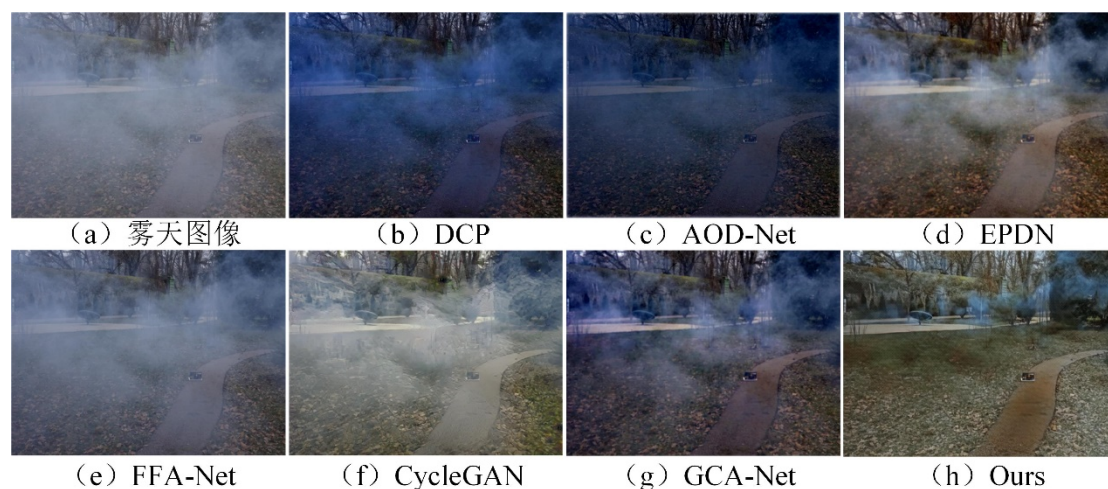


图 3-11 不同算法在 NH-Dehaze 上的去雾结果

与上述在真实雾天环境下的处理类似, DCP、AOD-Net 和 FFA-Net 虽然去除了部分雾霾, 但是从整体上看还有很多的雾没有去除干净。EPDN 和 GCA-Net 虽然部分图像去雾效果较好, 但是图像颜色偏移且恢复后的图像整体较暗。CycleGAN 仍然出现了风格迁移错误的情况。本章方法处理后的图像整体清晰, 物体还原度较高, 虽然部分图像的雾没有完全去除, 但是所剩的雾也较薄, 可以看清细节信息。

使用各算法在真实雾天环境进行定量比较, 表 3-3 是各算法在 URHI 上的 NIQE、FADE 和 BRISQUE 平均值。

表 3-3 各算法在真实雾天数据集 URHI 上的定量比较 (NIQE/FADE/ BRISQUE)

Method	Type	NIQE	FADE	BRISQUE
DCP	Prior	5.070	0.874	27.462
AOD-Net	Supervised	4.997	1.045	29.045
FFA-Net	Supervised	5.234	0.928	28.615
EPDN	Supervised	4.574	0.762	24.983
GCA-Net	Supervised	4.564	0.693	24.766
CycleGAN	Unsupervised	4.145	0.869	29.174
Cycle-Dehaze	Unsupervised	4.032	0.672	24.081
Ours	Unsupervised	3.765	0.701	23.725

通过对比各算法的指标值可以发现本章方法在 NIQE 和 BRISQUE 上均取得了最优的结果。而 FADE 指标也只比 Cycle-Dehaze 低了 0.029, 这些指标反映了本章方法对真实环境去雾具有良好的适应性。

3.6.4 消融实验

本节通过消融实验证实本章网络各部分的作用。将本章网络分为以下几个部分, 第一部分是基础网络 V1, 即原始的 CycleGAN 网络, 第二部分 V2 是在第一部分的基础上加上感知损失、SSIM 损失等损失函数。第三部分 V3 是在第二部分的基础上加上 FCA 注意机制, 最后 V4 再加上多层特征融合模块。因为本章主要针对室外环境进行去雾, 所以在 SOTS 的 outdoor 数据集上进行测试。图 3-12 是消融实验的结果图, 从图中可以看出所提的每个部分均对去雾效果有贡献, 在加入感知损失后, 可以修复被 CycleGAN 风格迁移破坏的部分, 加入特征融合模块可以提升网络对细节信息的恢复。同时通过使用 FCA 注意力机制, 可以使网络聚焦于图像中雾所处的区域, 网络可以达到很好的去雾效果, 如图 3-12 左下角所示, 网络去雾效果趋近于原始去雾图像。



图 3-12 消融实验结果图

3.6.5 车道线检测结果分析

在雾天环境下，一些高级视觉任务会受到雾的影响，如无人驾驶在检测车道线时，雾会遮挡车道线，导致车道线的检测准确率下降。本章使用第二章训练好的车道线检测网络对不同去雾方法增强的车道线图像进行检测，检测结果如图 3-13 所示。从图中可以看出，DCP 和 EPDN 去雾后图像变暗，导致车道线只检测出一条，而 CycleGAN 破坏了图像的车道线信息，导致检测精度下降。本章网络对车道线图像去雾后的检测精度提高，漏检率也降低了，相对于其他方法，本章检测出了第三条车道线，即左边的蓝色车道线。

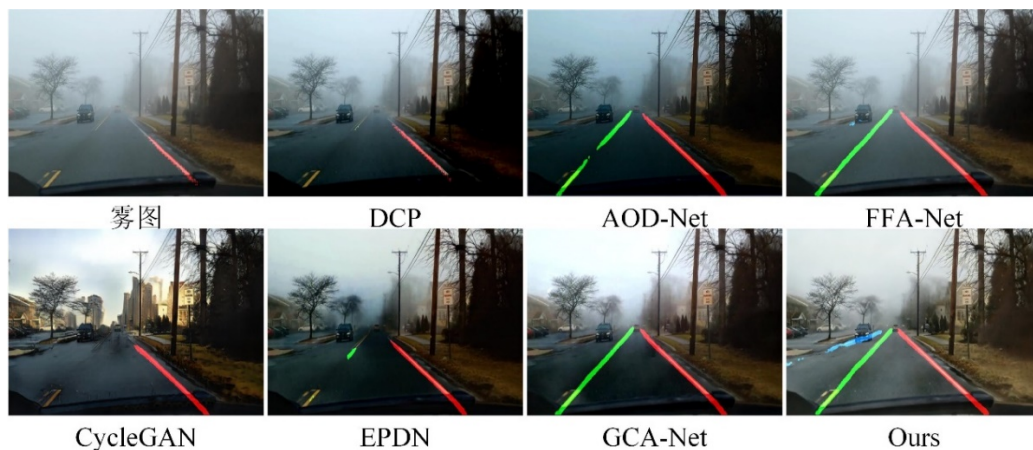


图 3-13 不同方法去雾后的车道线检测结果

3.7 本章小结

针对无人驾驶的车道线检测容易受到雾天环境的影响，导致检测精度变低，误检率升高的情况，本章基于 CycleGAN 网络设计一个多层特征融合去雾框架，该框架是无监督学习模型，不需要成对雾天数据，提升了网络对真实雾天环境处理的泛化能力。为了加强编码器各层特征信息的利用，设计一个多层特征融合模块，增强网络对细节信息和边缘信息的利用，防止特征在上采样时被稀释，同时使用 FCA 注意机制提升网络对不同通道的注意程度，优化网络对非均匀雾的处理。通过在各数据集上的实验证明本章网络对于图像雾霾的处理优于其他算法，本章网络去雾后颜色和细节信息保持较好，对去雾后的高级视觉任务起到了辅助作用，本章网络去雾后的车道线检测精度得到了提升。

第 4 章 基于 Uformer 的雨夜车道线检测

4.1 引言

雨天环境下能见度降低，雨水导致空气中水雾增多，从而降低了可见光的透射率，使得远处的景物变得模糊不清。恶劣的光线条件可能使得车道线和其他道路标记在图像中变得模糊、不清晰。目前已有较多算法可以实现对雨天环境图像的增强，但是在雨天的夜间环境下，图像增强和目标检测将变得更加困难，许多适用于雨天图像增强的网络也处理不了雨天和夜间相结合的复杂环境图像。

在夜间雨天环境下，天空通常较为昏暗，这会使得车道线和其他道路标记在图像中更加不清晰。夜间行驶时，车辆的前灯和路灯照射在路面的雨水上，使光线出现散射现象，这可能使得车道线在图像中受到更多的光线干扰，导致检测系统难以准确捕捉车道线的位置。

鉴于以上问题，本章设计一种基于 Uformer 的夜间雨天图像增强网络，本章的研究路线如图 4-1 所示。

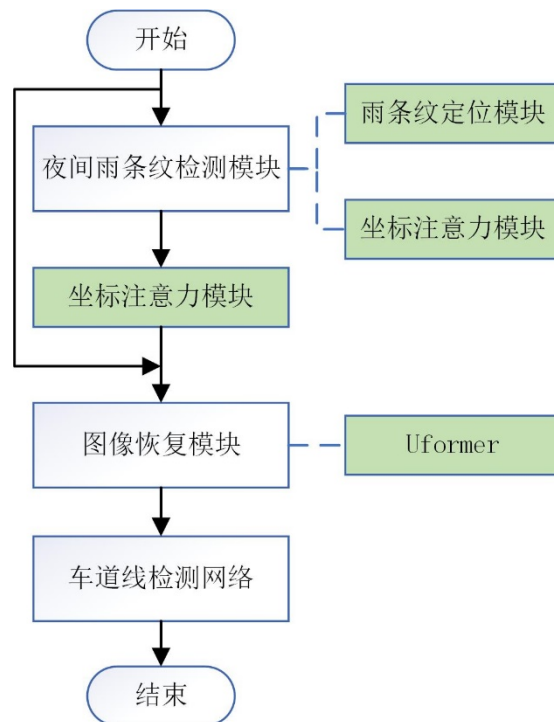


图 4-1 雨夜车道线检测研究路线图

首先使用夜间雨条纹位置检测模块对夜间雨天图像进行检测，得到夜间图像中雨条纹的位置信息，然后使用位置注意力模块强调雨条纹位置信息，提升后续除雨增强模块的增强效果，将上述所得先验信息输入到 Uformer 中进行图像增强。最后使用第二章设计的车道线检测网络对雨夜图像增强效果进行检验。

4.2 Uformer 网络

目前大部分的去雨网络采用基于 CNN 或基于 Transformer 的算法。常规基于 CNN 的网络能够通过卷积层和池化层逐级提取图像特征，具有对局部特征的强大提取能力。CNN 可以通过卷积操作实现参数共享，大幅度削减所需训练的参数总量，从而有效地提升模型训练的效率。此外 CNN 在处理空间结构上具有局部相关性，这很适用于处理具有空间结构的数据，如图像和视频等。但是 CNN 的局部感受野可能会导致信息的丢失，并且网络难以全局理解长距离依赖关系。而基于 Transformer 的网络通过自注意力机制实现全局关注输入序列的能力，能够有效捕捉到远距离的依赖关系。此外，该机制支持并行计算的特性，在特定场景下赋予了 Transformer 更高的计算效率。但是 Transformer 与常规使用的 CNN 不同，其没有循环结构和卷积结构。Transformer 模型由于其较高的计算复杂度，制约了其在图像去雨任务中的广泛应用。面对这一问题，Wang 等人^[46]提出了名为 Uformer 的网络结构，该网络将 CNN 中的常用网络 U-Net 和 Transformer 相结合进行图像复原，Uformer 架构可以理解作为一种将 U-Net 模型内部的卷积层替换成 Transformer 模块的设计，同时保留 U-Net 网络中的编码解码结构和跳跃连接结构。两者的结合可以实现以下效果，Transformer 模块可以建立起每个部分在图像全局方面的依赖关系，而 CNN 结构可以有效地提取单个部分的局部信息。

Uformer 的网络结构如图 4-2 所示。Uformer 网络依旧沿用了编码器-解码器的整体架构，并在内部继续采用跳跃连接机制将相同尺度的特征层相互衔接。Uformer 网络的具体步骤如下。首先将低质量图像 I 输入到网络中，然后使用 3×3 的卷积和 LeakyReLU 函数提取底层的特征 X ，特征图 X 会通过编码器阶段，每一个编码层都包含了局部增强窗口模块(Locally-enhance Window, LeWin)以及下采样层。其中，LeWin 模块运用了非重叠窗口的自注意力机制，旨在捕获较

远距离的依赖关系，从而有效降低计算负荷。对于下采样层，首先会对输入的特征执行重塑步骤，将其转化为二维空间内的特征映射，然后对该映射进行下采样操作。这种下采样策略可以更有效地处理特征，提高计算效率。在编码器部分的尾端，设置有一个瓶颈阶段，此阶段同样整合了 LeWin 模块。这个阶段中的 LeWin 块能够强有力地探测并捕捉到长距离的远程依赖关系。这种方法使得编码器能够更好地处理长距离上的相关信息。

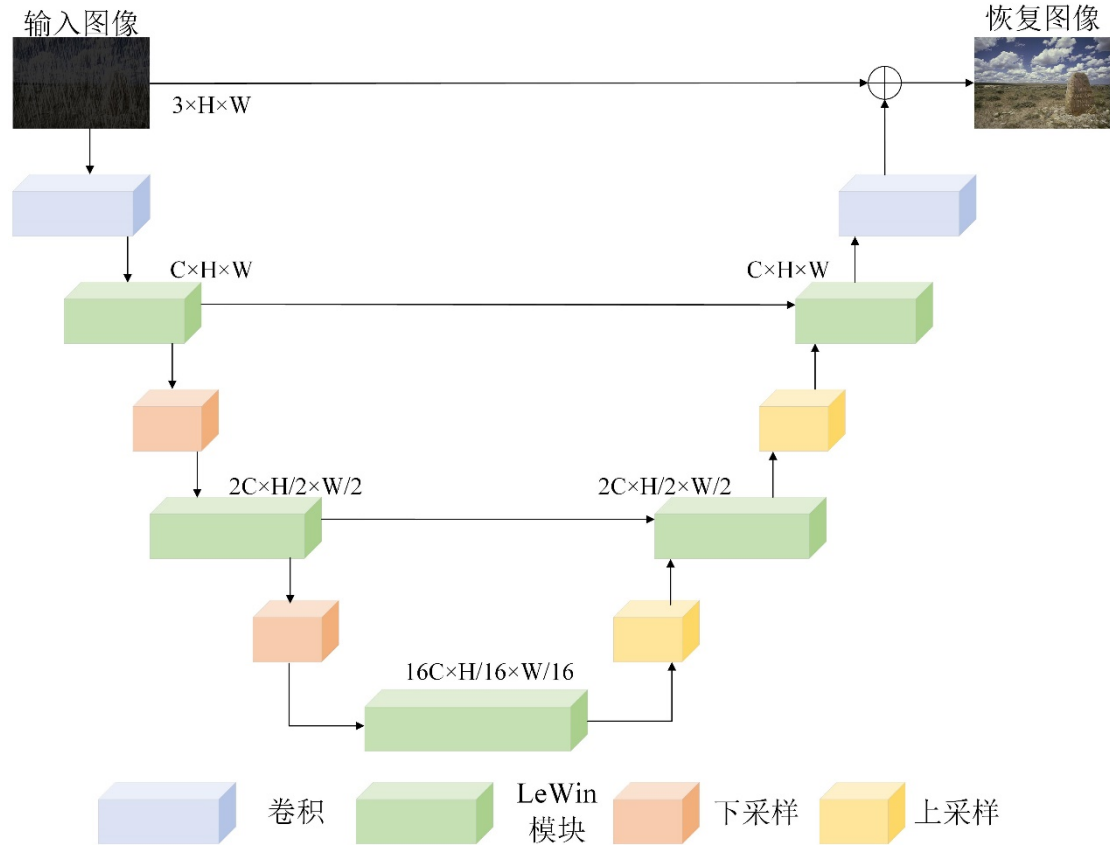


图 4-2 Uformer 网络结构图

在瓶颈阶段过后，有 K 个解码层，用于逐步重构特征信息。每个解码器模块均包含一个上采样层和 LeWin 块。在解码器中，采用 2×2 转置卷积进行上采样。每一层 LeWin 块的输入特征融合了上层上采样结果及编码器对应层的跳跃连接输出。经历 K 个解码器阶段后，通过 3×3 卷积提取残差图像，然后将残差与原始低质量图像相加以生成修复图像。图 4-3(a)展示了 Uformer 中核心的 LeWin Transformer 结构，其中 W-MSA 表示非重叠窗口多头注意力模块，LeFF 为局部增强前馈模块，LN 为层归一化。LeWin 模块的计算公式为：

$$X'_l = W - \text{MSA}(\text{LN}(X_{l-1})) + X_{l-1} \quad (4-1)$$

$$X_l = LeFF(LN(X'_l)) + X'_l \quad (4-2)$$

其中 X'_l 为第 l 层的 W-MSA 模块输出， X_l 为第 l 层的 LeFF 模块输出。

局部增强前馈网络 LeFF 的结构图如图 4-3(b)所示，在传统的 Transformer 模型中，前馈网络受限于捕获图像局部上下文信息的能力。鉴于图像去雨任务对相邻像素上下文的高度依赖性，Uformer 对其前馈网络做了优化：添加了 3×3 深度卷积层以强化模型学习局部特征的能力。具体来说，首先通过线性层对低质量图像特征进行升维处理，接着转化为二维特征图并利用深度卷积层捕获局部特征，随后降维并通过另一线性层满足输入通道数要求。Uformer 中，每个线性层和卷积层后均配备 GeLU 激活函数以提升模型性能。

Uformer 借助非重叠窗口局部自注意力机制，能在较大视野下运作，从而学习图像中的远距离依赖。当给定一个二维图像 X ，将二维图像分割为 $M \times M$ 大小的非重叠窗口 X^i 后，对每个窗口内的特征执行自注意力计算。对于多头注意力中的第 k 个头部，非重叠窗口内自注意力计算公式如下所示：

$$X = \{X^1, X^2, \dots, X^N\}, N = HW / M^2 \quad (4-3)$$

$$Y_k^i = Attention(X^i W_k^Q, X^i W_k^K, X^i W_k^V), i = 1, \dots, N \quad (4-4)$$

$$\hat{X}_k = \{Y_k^1, Y_k^2, \dots, Y_k^M\} \quad (4-5)$$

其中 W_k^Q ， W_k^K ， W_k^V 分别表示第 k 个自注意力头的查询、键和值的映射矩阵。

$\hat{X}_k$ 是第 k 个自注意力头的输出。将所有窗口自注意力计算产生的输出结果进行串联，并通过线性投影生成最终的输出结果。在 W-MSA 模块中，采用相对位置编码来增强自注意力计算的能力，其自注意力计算公式为：

$$Attention(Q, K, V) = Soft \max \left(\frac{QK^T}{\sqrt{d_k}} + B \right) V \quad (4-6)$$

其中 B 为相对位置偏差。与全局自注意力相比，基于窗口的自注意力可以显著的降低计算成本。

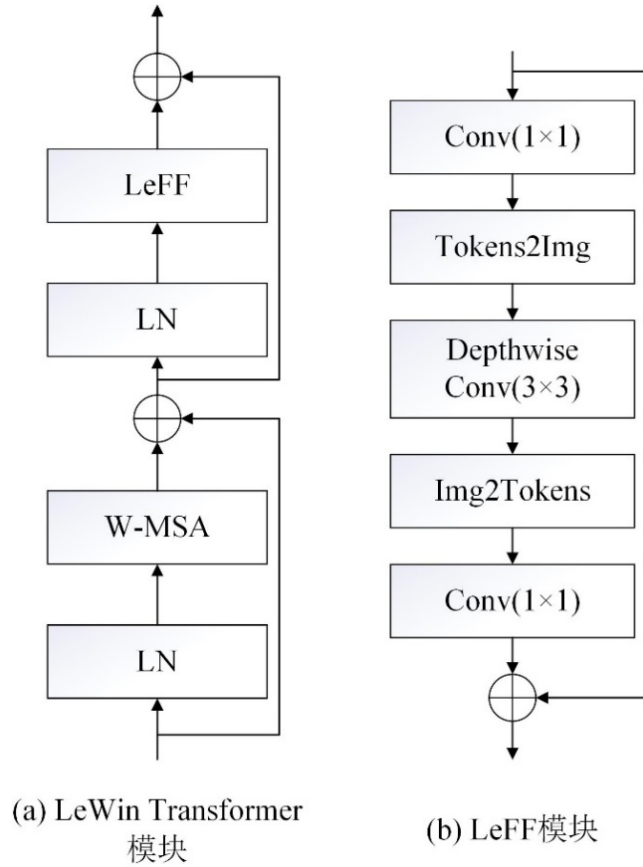


图 4-3 LeWin 和 LeFF 网络结构图

4.3 PReNet

本章的夜间雨条纹位置预测模块是基于 PReNet 网络进行设计的，因此在此小节对 PReNet 网络进行简单介绍。PReNet 网络包含输入层、残差块、循环层和输出层四个部分，PReNet 网络结构图如图 4-4 所示。输入层是由一个卷积核为 3×3 ，padding 为 1 的普通卷积层和一个 ReLU 非线性激活函数组成，输入层的输入是 6 通道，然后输出 32 通道。在 PReNet 网络中将上一个阶段输出的图像和原始的输入低质量图像进行连接，然后作为当前阶段的输入。

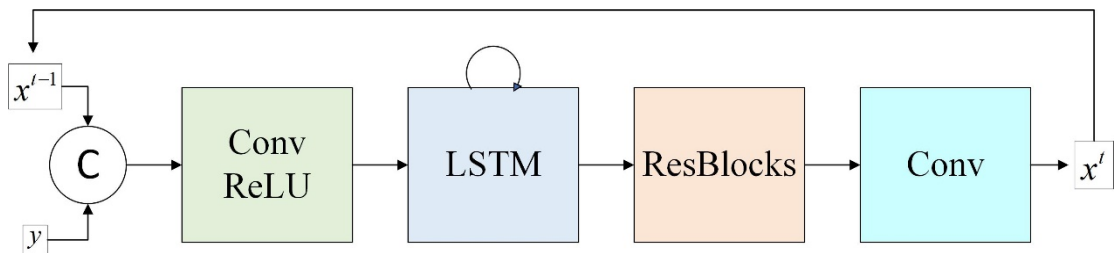


图 4-4 PReNet 基本结构图

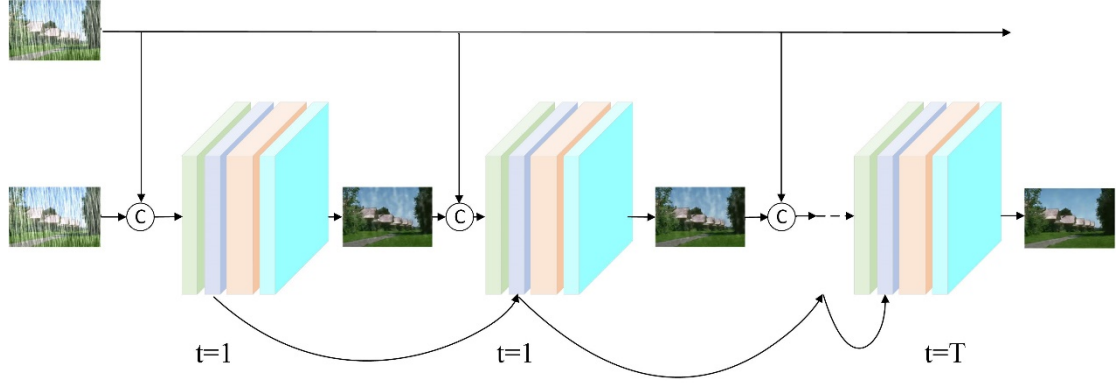


图 4-5 PReNet 的 T 阶段递归图

PReNet 网络的核心组成部分为时间递归神经网络(Long Short-Term Memory, LSTM), 相较于传统的残差网络, LSTM 在提取输入图像特征方面的表现更具优势。LSTM 通过隐藏层联结了短期与长期的时间序列信息, 从而实现了记忆存储功能。如图 4-6 所示, LSTM 的记忆单元架构中, 输入信号充当了网络的先验知识来源, 可以是前一轮训练得到的特征图, 其输出是预测结果。在记忆单元内部, 共设置了三个关键门控机制, 即输入门、输出门以及遗忘门, 这些门控以图中蓝色模块标识, 绿色模块代表了乘法运算操作, 虚线部分则表示对上一状态函数的引用。 S_t 为记忆单元在 t 时刻的状态。记忆单元状态的更新和输出公式如下:

$$i_t = \sigma(W^{(i)}X_t + U^{(i)}S_{t-1}) \quad (4-7)$$

$$f_t = \sigma(W^{(f)}X_t + U^{(f)}S_{t-1}) \quad (4-8)$$

$$O_t = \sigma(W^{(o)}X_t + U^{(o)}S_{t-1}) \quad (4-9)$$

$$\tilde{S}_t = \tanh(W^{(c)}X_t + U^{(c)}S_{t-1}) \quad (4-10)$$

$$S_t = f_t \circ S_{t-1} + i_t \circ \tilde{S}_t \quad (4-11)$$

$$O_t = o_t \circ \tanh(S_t) \quad (4-12)$$

其中 i_t 、 f_t 和 o_t 是三个门的输出。 $\tilde{S}_t$ 是记忆单元 t 时刻更新的状态, S_t 和 O_t 分别是记忆单元的最终状态和最终输出状态。 $W^{(i)}$ 、 $W^{(o)}$ 、 $W^{(f)}$ 、 $W^{(c)}$ 、 $U^{(i)}$ 、 $U^{(o)}$ 、 $U^{(f)}$ 和 $U^{(c)}$ 是不同的系数矩阵。

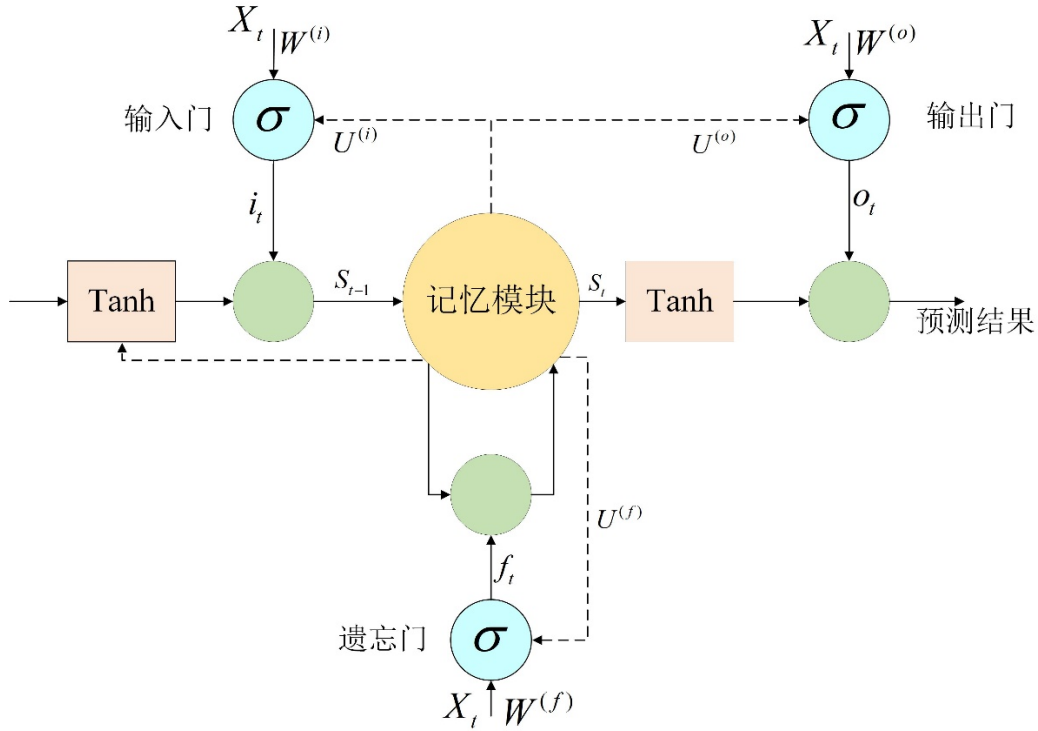


图 4-6 LSTM 记忆单元网络结构图

PReNet 网络的残差模块是由串联的 5 个残差子模块组成的，每个残差子模块具有相同的网络结构，是由两个带有 ReLU 激活函数的常规卷积组成的，其中的卷积核为 3×3 ，padding 为 1。

PReNet 网络的输出层是一个卷积核为 3×3 的普通卷积，将输入特征图的通道数重新转换为 3。

4.4 夜间雨天图像增强网络

雨水会降低图像的质量和可见度，在夜间环境下，由于低光照的影响使得雨条纹更加难以和其他图像特征分开，夜间的雨天图像在不同的空间位置上差异较大，只有靠近光源的雨条纹可以被检测出来，因此在夜间图像上准确地找到雨条纹的位置将对后续的夜间雨天图像增强十分有利。本节鉴于以上问题设计一个夜间雨天图像增强网络。网络总体架构如图 4-7 所示。首先使用改进的 PReNet 网络对夜间图像中的雨条纹进行检测，然后将雨条纹预测图输入到空间位置注意力模块对雨条纹区域进行深度定位，得到最终的雨条纹预测图，将其作为先验知识和输入图像同时输入到图像恢复模块，最后使用 Uformer 模块对图像进行增强和恢复。

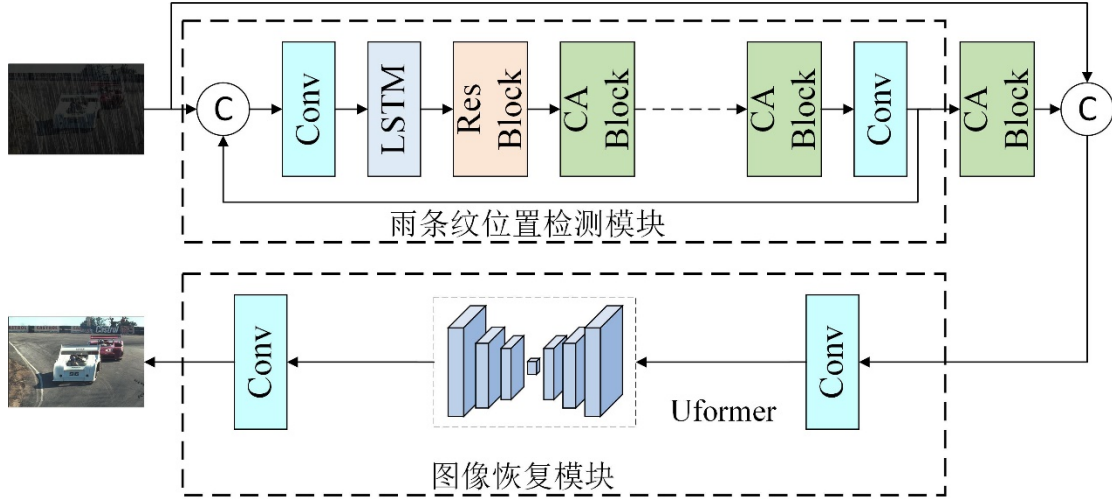


图 4-7 雨夜图像增强网络结构图

4.4.1 夜间雨条纹位置检测模块

能否获取夜间图像中雨条纹的位置直接影响了图像恢复网络对雨条纹的去除效果，因此在去除雨条纹之前使用循环残差结构来挖掘雨条纹位置的先验知识，它可以为后续增强网络提供雨条纹的位置信息。使用循环残差的结构是因为其可以学习雨条纹的高频特性来逐步捕获雨条纹。由于雨条纹空间位置的重要性，因此在 PReNet 网络的残差模块后添加坐标注意力模块(Coordinate Attention, CA)以强调雨条纹位置区域，此外在 PReNet 网络检测出雨条纹位置后，再次使用一个 CA 模块加强雨条纹区域的权重。

CA 模块通过赋予提取出的特征不同的权重，有效地引导网络聚焦于重要特征。目前广泛使用的注意力机制，例如 SE 模块和 ECA 模块，虽然强调了通道层面的信息处理，但却忽略了位置信息的关键作用。为更有效地捕获位置信息并考量通道之间的相互关联性，CA 模块将常规的二维全局池化操作分解为两次独立的一维编码步骤，将位置信息融入到通道注意力计算中。这样不仅能精准定位特征的位置特性，还能同步对通道间的关系以及长期依赖性进行编码处理。注意力的生成包括两个过程，即坐标信息嵌入和坐标注意力生成。CA 模块的网络结构如图 4-8 所示。

图中 C、H、W 分别为特征图的通道数、高度和宽度，其中 X Avg Pool 和 Y Avg Pool 分别代表了沿水平方向执行的全局平均池化操作以及沿垂直方向实施的全局平均池化处理。

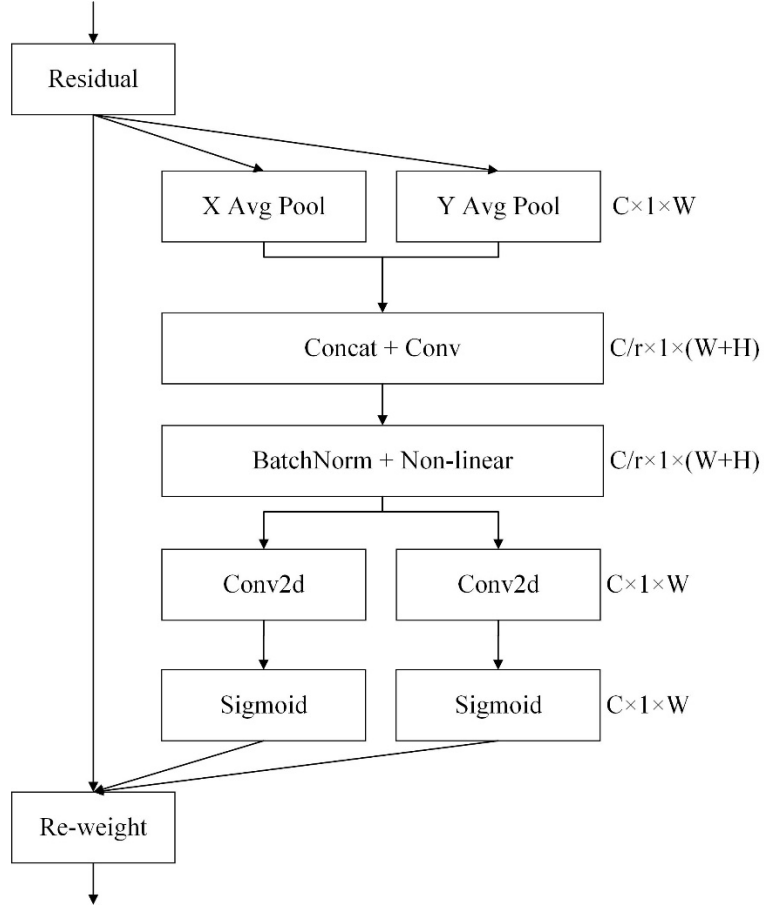


图 4-8 CA 模块结构图

CA 模块对输入特征的每个通道进行特征聚合时，分别使用尺寸为 H 的平均池化沿水平坐标方向和尺寸为 W 的平均池化沿垂直坐标方向。公式 4-13 可以用来表示高度 h 处的第 c 个通道的输出。第 c 个通道的宽度输出公式如公式 4-14 所示。

$$z_c^h(h) = \frac{1}{W} \sum_{0 \leq i < W} x_c(h, i) \quad (4-13)$$

$$z_c^w(w) = \frac{1}{H} \sum_{0 \leq j < H} x_c(j, w) \quad (4-14)$$

利用两个不同方向的变换完成特征聚合，生成了两个不同方向的特征图。这种方式能够更有效地协助注意力模块定位到关键区域，因为其兼具了捕捉空间长距离依赖关系和保持另一空间维度上精确位置信息的能力。将得到的两个特征图使用 Concat 连接并使用 F_1 变换函数，生成中间特征向量 f 。

$$f = \delta \left(F_1 \left(\begin{bmatrix} z^h \\ z^w \end{bmatrix} \right) \right) \quad (4-15)$$

其中 δ 为非线性激活函数。

之后将特征向量 f 沿着空间维度切分成两个向量 f^h 和 f^w ，并使用两个 1×1 卷积分别将这两个向量转换为与输入特征 X 具有相同通道的特征向量。

$$g^h = \sigma(F_h(f^h)) \quad (4-16)$$

$$g^w = \sigma(F_w(f^w)) \quad (4-17)$$

其中 σ 为 Sigmoid 函数。

CA 模块将得到的两个特征向量作为注意力的权重和输入特征 X 进行加权，CA 模块的输出公式如下。

$$y_c(i, j) = x_c(i, j) \times g_c^h(i) \times g_c^w(j) \quad (4-18)$$

将通过 CA 模块预测的雨条纹位置图和输入图像连接起来输入到 Uformer 网络中进行图像恢复。

4.4.2 损失函数

Uformer 网络在训练时通常使用 Charbonnier 损失函数，Charbonnier 损失函数是 L1 损失函数的一种，Charbonnier 损失通过计算 x 和 y 之间的差异对网络进行约束，但是这种损失函数并未考虑到人类的视觉感知效果。Charbonnier 损失会导致最终恢复的图像无法达到最优。

因此本章网络使用 MS-SSIM 损失和 L1 损失结合的方式来替换 Charbonnier 损失。在 MS-SSIM 中，图像会被逐步下采样多次，以构建一系列不同分辨率的图像，形成一个多尺度的图像金字塔结构。接下来，针对金字塔中各个分辨率层级的图像进行结构相似性的评估，最终将这些评估值结合起来得到最终的 MS-SSIM 评估值。这种方法能够更全面地捕捉图像的结构信息，保留图像中边缘细节信息。MS-SSIM 的计算公式如下：

$$MS-SSIM(x, y) = [l_M(x, y)]^{\alpha_M} \cdot \prod_{j=1}^M [c_j(x, y)]^{\beta_j} [s_j(x, y)]^{\gamma_j} \quad (4-19)$$

其中 α 、 β 和 γ 是其对应子函数的重要性调节参数。

MS-SSIM 损失函数的公式如下：

$$\mathcal{L}_{MS-SSIM}(P) = 1 - [MS - SSIM(\tilde{p})] \quad (4-20)$$

其中 P 为图像中片元集合， $\tilde{p}$ 为 P 的中心像素。通过对 P 进行 SSIM 计算并取均值得到。 $\mathcal{L}_1$ 损失函数在处理含有噪声的数据时鲁棒性更强，并且 $\mathcal{L}_1$ 损失对图像边缘的保持效果较好，因此将其与 MS-SSIM 损失结合使用。本章网络的具体损失函数如下：

$$\mathcal{L} = \alpha \cdot \mathcal{L}_{MS-SSIM} + (1 - \alpha) \cdot \mathcal{L}_1 \quad (4-21)$$

其中 α 为 0 到 1 的常数。

4.5 实验结果分析

本节将介绍网络训练及测试时所需的夜间雨天图像数据集，此外将介绍实验平台及训练过程参数设置。在本节中，将本章方法和组合形式的算法进行比较，分别在合成的夜间雨天数据集和真实夜间雨天数据集进行测试，其中图像质量评价指标选取 PSNR 和 SSIM，用于评价夜间雨天图像增强方法的表现。

4.5.1 实验设置

目前常用的雨天数据集只适用于正常光照条件下的去雨算法训练，这些数据集中几乎没有夜间和雨天相结合的图像，如 Rain200H、Rain100H^[76]等数据集。因此本章在 Hu 等人^[77]提供的 Cityscapes 数据集基础上进行改进，通过调整图像的对比度将图像转换成夜间雨天图像，具体合成效果如图 4-9 所示。



(a) 夜间雨天图像

(b) 正常环境图像

图 4-9 夜间雨天图像数据集

Cityscapes 数据集包含 9432 张训练图像和 1188 张测试图像，Cityscapes 数据集所呈现的效果更加接近真实雨天环境。本章网络使用合成的 Night-Rain-

Cityscapes 数据集进行训练，使用该数据集中的合成测试集对网络进行测试，同时为了反应网络在真实雨夜环境下的增强效果，使用真实雨夜环境图像对各网络增强效果进行检测。

本章的实验平台采用 Ubuntu 系统，CPU 的型号为 Intel(R) Xeon(R) Platinum 8255C，GPU 为 NVIDIA RTX3090 GPU。编程语言使用 Python3.8，实验软件采用 PyCharm，深度学习基础框架为 PyTorch。使用 Adam 优化器进行训练，初始的学习率设置为 2×10^{-4} ，最后通过余弦调节函数降低至 1×10^{-6} 。为了便于训练，输入图像的大小裁剪为 256×256 ，训练轮数设置为 250，batchsize 设置为 8。损失函数系数 α 取值为 0.5。

4.5.2 合成的夜间雨天数据集上的实验结果分析

由于目前研究者对于夜间雨天图像的处理多采用两阶段的图像增强方式，因此本章使用四种去雨效果较为显著的网络和夜间图像增强网络进行组合，同本章网络进行对比。其中夜间图像增强网络选用本文在第二章设计的夜间图像增强网络，使用 NIE-Net 指代其名称。去雨网络选用 MPRNet^[43]、DRDNet^[42]、Uformer^[46]和 RCDNet^[45]。与上述网络不同的是本章网络属于端到端的夜间雨天图像增强网络。本章首先在合成的雨夜测试集，即 Night-Rain-Cityscapes 测试集上进行实验，实验结果如图 4-10 至图 4-12 所示。

非端到端的增强网络在去除雨条纹时很难清除彻底，如图 4-10(c)和图 4-11(c)所示，被黑暗部分掩盖的雨条纹没有完全去除，RCDNet 网络对于细小部分的雨条纹去除效果较好，但是在去除较长的雨条纹时效果较差。如图 4-11(c)中的车标附近所示。MPRNet 网络可以去除图像中的大部分雨条纹，但是图像中会残留去雨的痕迹，降低了图像的质量，如图 4-12(b)所示，墙体部分和树木部分有较为明显的去雨痕迹。Uformer 网络可以达到较好的去雨效果，但是在增强过程中会引入噪声，如图 4-10(d)所示，车辆前端天空区域出现错误增强。DRDNet 网络增强后的图像出现失真的情况，雨条纹区域被白色所掩盖，如图 4-12(e)所示，左侧房屋部分出现了较为浓厚的雾，自行车和车道线部分的图像被破坏，直接影响整张图像的观感。相较于对比的算法而言，本章的端到端网络可以更好的修复图像，对于细节的保留较为完整。

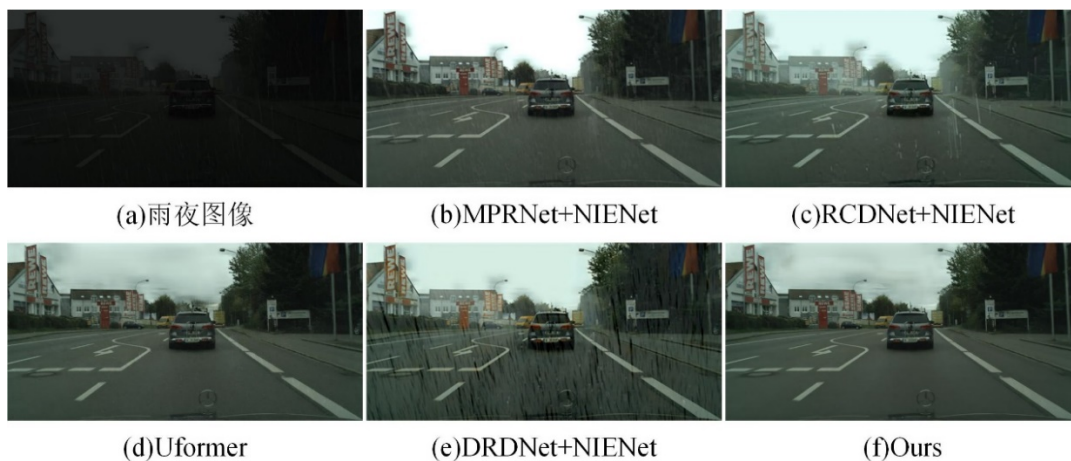


图 4-10 合成雨夜场景一增强结果图

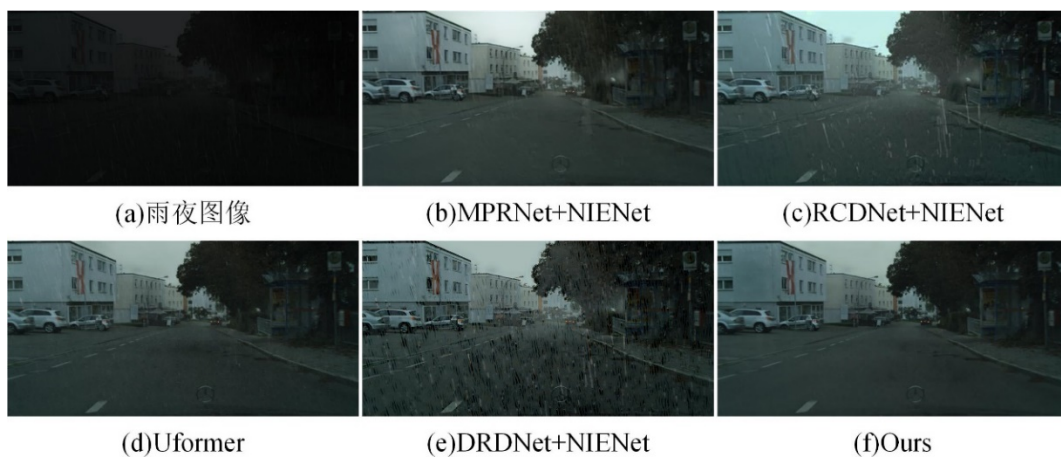


图 4-11 合成雨夜场景二增强结果图



图 4-12 合成雨夜场景三增强结果图

为了评估各种算法在数据层面的增强效果，对组合网络和本章网络在合成的 Night-Rain-Cityscapes 测试集以及合成的 Night-Rain200H 测试集上进行了 PSNR 和 SSIM 值的测量。表 4-1 是各种算法在合成数据集上的定量对比。从表

中可以看出本章网络增强后图像的 PSNR 值和 SSIM 值均高于其他组合算法，同时由于是端到端的网络，耗时也优于其他网络。

表 4-1 各算法在合成测试集上的定量比较 (PSNR/SSIM)

Method	PSNR(db)	SSIM	Time(s)
MPRNet+NIENet	25.43	0.749	0.711
RCDNet+NIENet	25.71	0.717	0.364
Uformer	28.49	0.802	0.214
DRDNet+NIENet	22.65	0.641	0.848
Ours	29.82	0.895	0.105

4.5.3 真实的夜间雨天数据集上的实验结果分析

为了验证本章网络对于真实夜间雨天环境下的增强效果，将本章算法和组合网络在真实夜间雨天环境进行实验，实验结果如图 4-13 和图 4-14 所示。在真实夜间环境下，去雨算法的性能会受到严重的影响，对于图像中亮度正常的区域，所有去雨算法均取得了较好的去除效果，但是对于隐藏在黑暗中的雨条纹，部分网络无法检测到雨条纹的具体位置，导致去雨效果较差。如图 4-14 中的红色物体，由于红色物体附近的亮度较高，雨条纹的去除效果较好，但是其他黑暗区域的雨条纹没有被去除。本章网络可以通过雨条纹检测模块预先检测出图像中的雨条纹位置，包括黑暗场景中的雨条纹，因此本章网络的整体去雨效果优于其他组合对比网络。

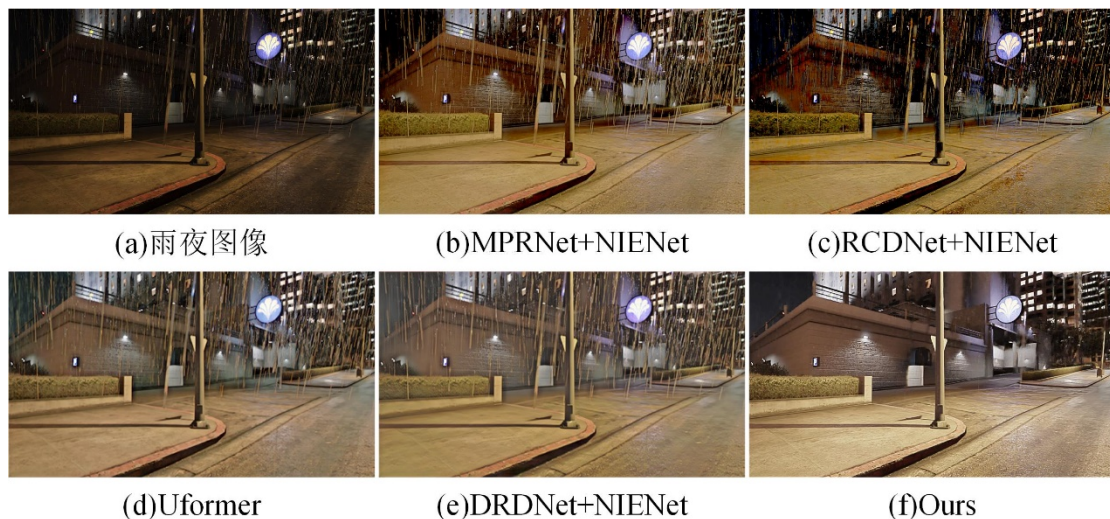


图 4-13 真实雨夜场景一增强结果图

本章同时在真实夜间雨天环境下对增强图像的 NIQE 值进行了测量，表 4-2 是各种算法在真实雨夜的定量对比。从表格数据可以清晰看出，本章网络增强

后的图像具有较低的 NIQE 值，相较于表现最好的 MPRNet 组合网络低了 0.568，同时比端到端的 Uformer 网络降低了 0.995，这表明增强后的图像在视觉效果上优于其他组合算法。

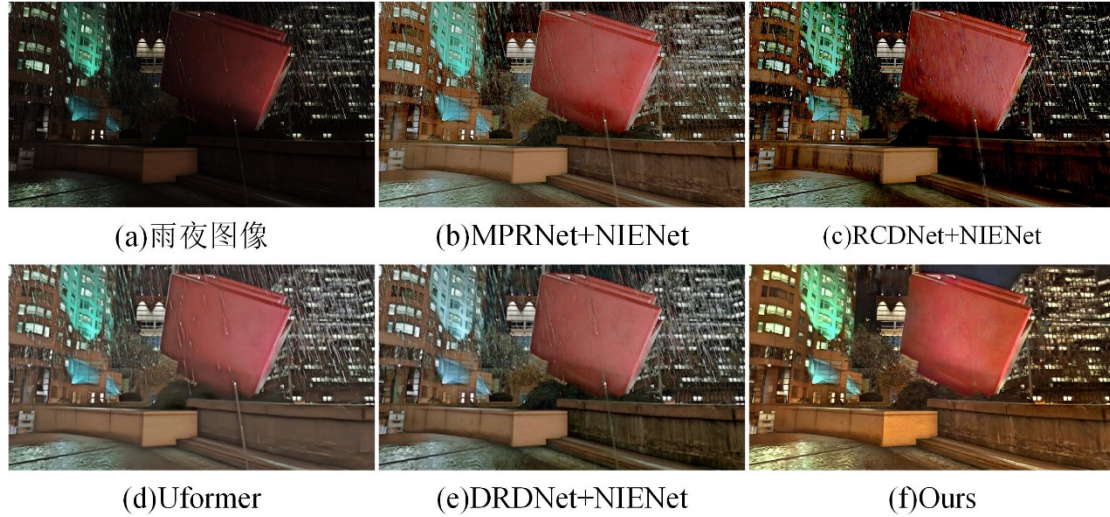


图 4-14 真实雨夜场景二增强结果图

表 4-2 各算法在真实雨夜的定量比较 (NIQE)

Method	NIQE	Time(s)
MPRNet+NIENet	3.843	0.711
RCDNet+NIENet	4.071	0.364
Uformer	4.270	0.214
DRDNet+NIENet	4.545	0.848
Ours	3.275	0.105

4.5.4 夜间雨天数据集上的车道线检测

在夜间雨天环境下，车道线检测受到两部分因素影响，一是雨条纹会遮挡车道线，导致车道线的检测准确率下降；二是夜间环境，图像整体亮度较低，车道线部分对比度不高，检测网络无法识别出车道线。为了验证夜间雨天图像增强网络的必要性，本章使用第二章训练好的车道线检测网络对不同方法增强的车道线图像进行检测，检测结果如图 4-15 所示。从图中可以看出，对于原始输入的夜间雨天图像，车道线检测网络只能检测出一点车道线，部分图像甚至无法检测出其中的车道线。组合形式的网络在两次增强后即使不可避免的引入了更多的噪声，且受到去雨网络去雨不彻底的影响，车道线检测网络只能检测到部分车道线，如图 4-15(b)和(e)所示，未去除的雨条纹遮挡了部分车道线，导致检测网络只能检测到一点车道线。本章网络属于端到端的网络，去雨效果较

好，车道线检测网络可以精准的检测出其中的车道线。

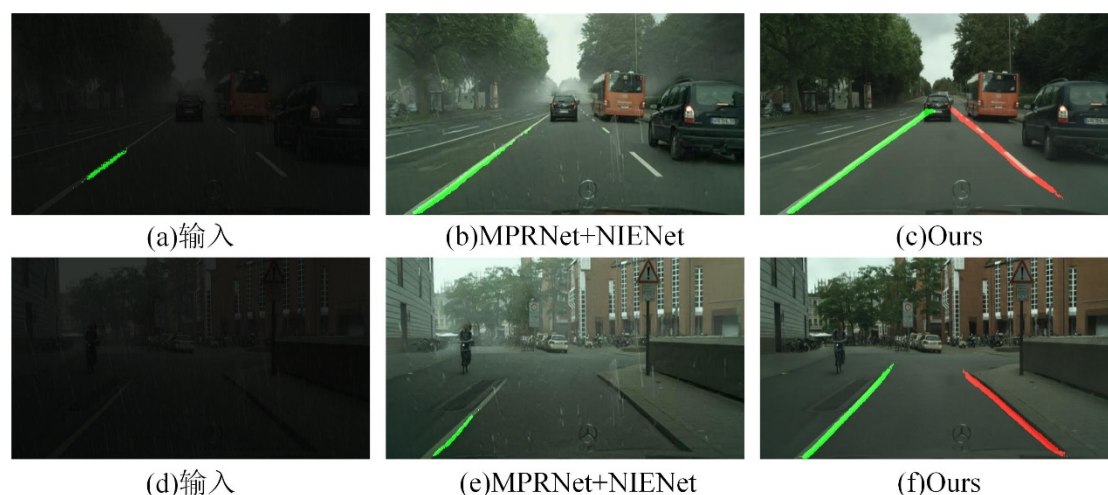


图 4-15 雨夜车道线检测结果图

4.6 本章小结

本章针对夜间雨天环境下很难检测出图像中的车道线，设计一个端到端的夜间雨天增强网络。首先使用以 PReNet 网络为基础的网络对夜间图像中的雨条纹进行位置检测，为后续复原网络提供雨条纹位置先验知识，然后使用 Uformer 网络根据雨条纹位置先验对图像进行增强。为了更好的训练网络，在 Cityscapes 数据集的基础上构建了 Night-Rain-Cityscapes 数据集。最后将本章网络和其他组合型网络进行定量和定性分析，本章增强网络可以很好的恢复图像，同时提升车道线检测网络的检测效果。

第 5 章 总结与展望

5.1 总结

无人驾驶技术中，车道线检测是其关键技术之一。恶劣环境下能见度下降、车道线模糊不清，严重影响车道线检测的精度。不同恶劣环境对车道线的影响也不同，夜间环境下车道线被黑暗所掩盖，车道线特征不明显，复杂的光照条件也会干扰检测结果；雾天环境下，车道线和浓雾融为一体，车道线检测网络无法从图像中将其与背景区分开；雨夜环境下，受低光照的影响，去雨算法无法去除隐藏在黑暗部分的雨条纹，间接导致未去除的雨条纹遮挡图像中的车道线部分，严重降低检测效果。针对以上问题，论文的主要工作与总结如下：

(1) 针对夜间场景下车道线图像整体呈现低光照、光照不均匀的特征，直接导致车道线检测准确度低，漏检率和误检率较高的问题，设计两个网络，分别是夜间图像增强网络和车道线检测网络。对于夜间图像增强网络，在 Retinex 理论的基础上设计反射率估计网络，通过在编码层引入注意力机制提升网络对夜间低光照区域的关注，同时引入深度可分离卷积提升网络的运行速度。通过实验结果可以看出，改进的夜间图像增强网络可以在保留图像颜色信息的基础上提升低光照部分的亮度，并且不破坏图像的细节信息。对于车道线检测网络，在 ERF-Net 网络的编码层引入自注意力蒸馏，提升网络对车道线信息的认知能力，从实验结果中可以看出网络在保证实时性的同时提升了检测精度。

(2) 针对雾天车道线图像容易受到雾气遮掩的问题，使用无监督型去雾网络对雾天图像进行增强。为了保证网络可以适用于真实雾天环境，使用 CycleGAN 网络作为去雾网络的骨干网络，在生成器中引入多层特征融合模块提升不同编码层特征信息的利用率，通过重新设计多个残差块结构以降低网络的计算量，使网络可以适用于无人驾驶场景。此外使用判别器从全局和局部两个角度分别判断生成的图像，使网络可以达到更好的博弈效果。在不同数据上的实验结果表明改进的去雾网络可以有效的去除不同浓度的雾，同时提升车道线检测网络的检测效果。

(3) 针对夜间雨天环境下车道线图像整体亮度较低，去雨网络无法去除隐藏

在黑暗中的雨条纹，雨条纹会遮挡车道线信息等问题。设计一个端到端的夜间雨天图像增强网络，通过使用 PReNet 网络和坐标注意力预先检测夜间图像中雨条纹的位置，将雨条纹位置先验和夜间雨天图像输入到图像恢复网络对图像进行增强。从合成和真实的夜间雨天数据上的实验结果可以看出，端到端的增强网络可以去除不同光照下的雨条纹，同时可以提升车道线检测网络的检测效果。

5.2 展望

本文主要研究了不同恶劣环境下的图像增强方法和车道线检测方法。虽然对于图像增强和车道线检测取得了一些研究成果，但是由于自身研究能力有限，本文的工作还有进一步研究的空间。主要内容如下：

(1) 目前的车道线数据集中大多是白天状态下的车道线图像，因此训练出来的车道线检测网络对于恶劣环境的车道线检测能力较差。可以通过扩充不同环境下的车道线数据集进行研究；

(2) 目前许多基于 Transformer 的网络模型对于图像复原领域起到了很好的作用，后续可以尝试以 Transformer 网络为基础对夜间图像和雾天图像进行增强。

参考文献

- [1] 张鹏博, 陈仁祥, 邵毅明, 等. 纯电动汽车电驱动系统故障诊断研究进展[J]. 汽车工程, 2024, 46(01): 61-74.
- [2] 李洪涛. 基于 RBF 神经网络的车辆纵向避撞预警模型研究[D]. 哈尔滨: 东北林业大学, 2021.
- [3] 刘国豪. 基于卷积神经网络的车道线和交通目标检测算法研究[D]. 北京: 北京交通大学, 2022.
- [4] 任柯燕, 谷美颖, 袁正谦, 等. 自动驾驶 3D 目标检测研究综述[J]. 控制与决策, 2023, 38(04): 865-889.
- [5] 郭心悦, 韩星宇, 习超, 等. 基于非对称卷积的多车道线检测方法[J]. 计算机工程与设计, 2024, 45(02): 428-435.
- [6] 常振廷, 肖智豪, 张文军, 等. 基于网格分类与纵横向注意力的城市道路车道线检测方法[J]. 交通信息与安全, 2023, 41(03): 92-102+110.
- [7] Reza A M. Realization of the Contrast Limited Adaptive Histogram Equalization (CLAHE) for Real-Time Image Enhancement[J]. Journal of VLSI Signal Processing Systems for Signal, Image and Video Technology, 2004, 38: 35-44.
- [8] 涂毅晗, 汪普庆. 基于多尺度局部直方图均衡化的矿井图像增强方法[J]. 工矿自动化, 2023, 49(08): 94-99.
- [9] 赵丽. 基于直方图均衡的夜间图像增强算法研究[D]. 兰州: 兰州大学, 2018.
- [10] Land E H, McCann J J. Lightness and Retinex Theory[J]. Journal of the Optical Society of America, 1971, 61(1): 1-11.
- [11] Rahman Z, Jobson D J, Woodell G A. Retinex Processing for Automatic Image Enhancement[J]. Journal of Electronic Imaging, 2004, 13(1): 100-110.
- [12] Lore K G, Akintayo A, Sarkar S. LLNet: A Deep Autoencoder Approach to Natural Low-Light Image Enhancement[J]. Pattern Recognition, 2017, 61: 650-662.
- [13] Shen L, Yue Z, Feng F, et al. MSR-net: Low-Light Image Enhancement Using Deep Convolutional Network[J]. arXiv preprint arXiv:1711.02488, 2017.

- [14] Wei C, Wang W, Yang W, et al. Deep Retinex Decomposition for Low-Light Enhancement[J]. arXiv preprint arXiv:1808.04560, 2018.
- [15] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative Adversarial Nets[J]. Advances in Neural Information Processing Systems, 2014, 27.
- [16] Jiang Y, Gong X, Liu D, et al. Enlighten: Deep Light Enhancement without Paired Supervision[J]. IEEE Transactions on Image Processing, 2021, 30: 2340-2349.
- [17] 吴佳奇, 张文琪, 陈伟, 等. 基于改进 CycleGAN 的煤矿井下低照度图像增强方法[J]. 华中科技大学学报(自然科学版), 2023, 51(05): 40-46.
- [18] Anoosheh A, Sattler T, Timofte R, et al. Night-to-Day Image Translation for Retrieval-based Localization[C]. Proceedings of the 2019 International Conference on Robotics and Automation (ICRA). IEEE, 2019: 5958-5964.
- [19] Guo C, Li C, Guo J, et al. Zero-Reference Deep Curve Estimation for Low-Light Image Enhancement[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020: 1780-1789.
- [20] Nayar S K, Narasimhan S G. Vision in Bad Weather[C]. Proceedings of the Seventh IEEE International Conference on Computer Vision. IEEE, 1999, 2: 820-827.
- [21] He K, Sun J, Tang X. Single Image Haze Removal Using Dark Channel Prior[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 33(12): 2341-2353.
- [22] Berman D, Avidan S. Non-Local Image Dehazing[C]. Proceeding of the IEEE Conference on Computer Vision and Pattern Recognition. 2016: 1674-1682.
- [23] 高涛, 刘梦尼, 陈婷, 等. 结合暗亮通道先验的远近景融合去雾算法[J]. 西安交通大学学报, 2021, 55(10): 78-86.
- [24] 杨燕, 陈阳. 基于双支特征联合映射的端到端图像去雾算法[J/OL]. 湖南大学学报(自然科学版), <http://kns.cnki.net/kcms/detail/43.1061.N.20231213.1545.004.html>, 2023-12-13.

- [25] 唐剑, 车文刚, 高盛祥. 融入注意力机制的多尺度卷积图像去雾方法[J]. 计算机工程与科学, 2023, 45(08): 1453-1462.
- [26] Cai B, Xu X, Jia K, et al. Dehazenet: An End-to-End System for Single Image Haze Removal[J]. IEEE Transactions on Image Processing, 2016, 25(11): 5187-5198.
- [27] Ren W, Liu S, Zhang H, et al. Single Image Dehazing via Multi-Scale Convolutional Neural Networks[C]. Proceedings of the Computer Vision–ECCV. 2016: 11-14.
- [28] Li B, Peng X, Wang Z, et al. Aod-Net: All-in-One Dehazing Network[C]. Proceeding of the IEEE International Conference on Computer Vision. 2017: 4770-4778.
- [29] Chen D, He M, Fan Q, et al. Gated Context Aggregation Network for Image Dehazing and Deraining[C]. Proceedings of the 2019 IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE, 2019: 1375-1383.
- [30] Qin X, Wang Z, Bai Y, et al. FFA-Net: Feature Fusion Attention Network for Single Image Dehazing[C]. Proceeding of the AAAI Conference on Artificial Intelligence. 2020, 34(07): 11908-11915.
- [31] Engin D, Genç A, Kemal Ekenel H. Cycle-Dehaze: Enhanced CycleGAN for Single Image Dehazing[C]. Proceeding of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2018: 825-833.
- [32] Sun Z, Zhang Y, Bao F, et al. ICycleGAN: Single Image Dehazing Based on Iterative Dehazing Model and CycleGAN[J]. Computer Vision and Image Understanding, 2021, 203: 103133.
- [33] 张成, 潘明阳, 高翊然, 等. 融合暗通道先验与循环生成对抗网络的航海图像去雾模型[J]. 大连海事大学学报, 2022, 48(04): 84-93.
- [34] 但志平, 方帅领, 孙航, 等. 基于双判别器异构 CycleGAN 框架下多阶通道注意力校准的室外图像去雾[J]. 电子学报, 2023, 51(09): 2558-2571.
- [35] Chaitanya B, Mukherjee S. Single Image Dehazing Using Improved CycleGAN[J]. Journal of Visual Communication and Image Representation, 2021, 74: 103014.

- [36] Kang L W, Lin C W, Fu Y H. Automatic Single-Image-based Rain Streaks Removal via Image Decomposition[J]. IEEE Transactions on Image Processing, 2011, 21(4): 1742-1755.
- [37] 汤红忠, 王翔, 张小刚, 等. 面向单幅图像去雨的非相干字典学习及其稀疏表示研究[J]. 通信学报, 2017, 38(07): 28-35.
- [38] Fu X, Huang J, Zeng D, et al. Removing Rain from Single Images via A Deep Detail Network[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017: 3855-3863.
- [39] Yang W, Tan R T, Feng J, et al. Deep Joint Rain Detection and Removal from A Single Image[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017: 1357-1366.
- [40] Zhang H, Patel V M. Density-Aware Single Image De-raining Using A Multi-Stream Dense Network[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 695-704.
- [41] Ren D, Zuo W, Hu Q, et al. Progressive Image Deraining Networks: A Better and Simpler Baseline[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 3937-3946.
- [42] Deng S, Wei M, Wang J, et al. DRD-Net: Detail-Recovery Image Deraining via Context Aggregation Networks[J]. arXiv preprint arXiv:1908.10267, 2019.
- [43] Zamir S W, Arora A, Khan S, et al. Multi-Stage Progressive Image Restoration[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021: 14821-14831.
- [44] 王美华, 何海君, 李超. 自适应卷积的残差修正单幅图像去雨[J]. 中国图象图形学报, 2020, 25(12): 2484-2493.
- [45] Wang H, Xie Q, Zhao Q, et al. A Model-Driven Deep Neural Network for Single Image Rain Removal[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020: 3103-3112.
- [46] Wang Z, Cun X, Bao J, et al. Uformer: A general U-Shaped Transformer for Image Restoration[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022: 17683-17693.

- [47] Cai H, Hu Z, Huang G, et al. Robust Road Lane Detection from Shape and Color Feature Fusion for Vehicle Self-Localization[C]. Proceedings of the 2017 4th International Conference on Transportation Information and Safety (ICTIS). IEEE, 2017: 1009-1014.
- [48] 宋琴琴, 杨国平. 基于 HSV 颜色阈值分割的车道线检测[J]. 计算机与数字工程, 2021, 49(09): 1895-1898.
- [49] 化春键, 潘瑞, 陈莹. 一种基于边缘方向连续性的车道线快速检测算法[J]. 传感技术学报, 2021, 34(08): 1082-1088.
- [50] Kim J, Lee M. Robust Lane Detection based on Convolutional Neural Network and Random Sample Consensus[C]. Proceedings of the International Conference on Neural Information Processing. 2014: 454-461.
- [51] He B, Ai R, Yan Y, et al. Accurate and Robust Lane Detection based on Dual-View Convolutional Neural Network[C]. Proceedings of the 2016 IEEE Intelligent Vehicles Symposium (IV). IEEE, 2016: 1041-1046.
- [52] Neven D, De Brabandere B, Georgoulis S, et al. Towards End-to-End Lane Detection: An Instance Segmentation Approach[C]. Proceedings of the 2018 IEEE Intelligent Vehicles Symposium (IV). IEEE, 2018: 286-291.
- [53] Pan X, Shi J, Luo P, et al. Spatial as Deep: Spatial Cnn for Traffic Scene Understanding[C]. Proceedings of the AAAI Conference on Artificial Intelligence. 2018, 32(1).
- [54] Qin Z, Wang H, Li X. Ultra Fast Structure-Aware Deep Lane Detection[C]. Proceedings of the 16th European Conference on Computer Vision. 2020. 276-291.
- [55] 杜黎, 吕毅斌, 武德安, 等. 复杂场景中的快速车道线检测方法[J]. 计算机工程与应用, 2023, 59(13): 178-185.
- [56] Romera E, Alvarez J M, Bergasa L M, et al. Erfnet: Efficient Residual Factorized Convnet for Real-Time Semantic Segmentation[J]. IEEE Transactions on Intelligent Transportation Systems, 2017, 19(1): 263-272.

- [57] 吕伏, 崔向燕, 刘铁. 亮度均衡与边缘强化的多尺度低照度图像增强[J/OL]. 激光与光电子学进展, <http://kns.cnki.net/kcms/detail/31.1690.tn.20240220.1123.078.html>, 2024-2-22.
- [58] Chollet F. Xception: Deep Learning with Depthwise Separable Convolutions[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017: 1251-1258.
- [59] Woo S, Park J, Lee J Y, et al. Cbam: Convolutional Block Attention Module[C]. Proceedings of the European Conference on Computer Vision (ECCV). 2018: 3-19.
- [60] Paszke A, Chaurasia A, Kim S, et al. Enet: A Deep Neural Network Architecture for Real-Time Semantic Segmentation[J]. arXiv preprint arXiv:1606.02147, 2016.
- [61] Hou Y, Ma Z, Liu C, et al. Learning Lightweight Lane Detection Cnns by Self Attention Distillation[C]. Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019: 1013-1021.
- [62] Cai J, Gu S, Zhang L. Learning A Deep Single Image Contrast Enhancer from Multi-Exposure Images[J]. IEEE Transactions on Image Processing, 2018, 27(4): 2049-2062.
- [63] Yu F, Xian W, Chen Y, et al. Bdd100k: A Diverse Driving Video Database with Scalable Annotation Tooling[J]. arXiv preprint arXiv:1805.04687, 2018, 2(5): 6.
- [64] Wang Z, Bovik A C, Sheikh H R, et al. Image Quality Assessment: from Error Visibility to Structural Similarity[J]. IEEE Transactions on Image Processing, 2004, 13(4): 600-612.
- [65] Mittal A, Soundararajan R, Bovik A C. Making A “Completely Blind” Image Quality Analyzer[J]. IEEE Signal Processing Letters, 2012, 20(3): 209-212.
- [66] Zhu J Y, Park T, Isola P, et al. Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks[C]. Proceedings of the IEEE International Conference on Computer Vision. 2017: 2223-2232.
- [67] Qin Z, Zhang P, Wu F, et al. Fcanet: Frequency Channel Attention Networks[C]. Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021: 783-792.

- [68] Iizuka S, Simo-Serra E, Ishikawa H. Globally and Locally Consistent Image Completion[J]. ACM Transactions on Graphics (ToG), 2017, 36(4): 1-14.
- [69] Isola P, Zhu J Y, Zhou T, et al. Image-to-Image Translation with Conditional Adversarial Networks[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017: 1125-1134.
- [70] Miyato T, Kataoka T, Koyama M, et al. Spectral Normalization for Generative Adversarial Networks[J]. arXiv preprint arXiv:1802.05957, 2018.
- [71] Versteegen R, Gimel'farb G, Riddle P. Texture Modelling with Nested High-Order Markov-Gibbs Random Fields[J]. Computer Vision and Image Understanding, 2016, 143: 120-134.
- [72] Choi L K, You J, Bovik A C. Referenceless Prediction of Perceptual Fog Density and Perceptual Image Defogging[J]. IEEE Transactions on Image Processing, 2015, 24(11): 3888-3901.
- [73] Mittal A, Moorthy A K, Bovik A C. No-Reference Image Quality Assessment in the Spatial Domain[J]. IEEE Transactions on Image Processing, 2012, 21(12): 4695-4708.
- [74] Li B, Ren W, Fu D, et al. Benchmarking Single-Image Dehazing and Beyond[J]. IEEE Transactions on Image Processing, 2018, 28(1): 492-505.
- [75] Qu Y, Chen Y, Huang J, et al. Enhanced Pix2pix Dehazing Network[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 8160-8168.
- [76] Yang W, Tan R T, Feng J, et al. Deep Joint Rain Detection and Removal from A Single Image[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017: 1357-1366.
- [77] Hu X, Fu C W, Zhu L, et al. Depth-Attentional Features for Single-Image Rain Removal[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 8022-8031.
- [78] Zhang Y, Di X, Zhang B, et al. Self-Supervised Image Enhancement Network: Training with Low Light Images Only[J]. arXiv preprint arXiv:2002.11300, 2020.

- [79] Jiang Z, Li H, Liu L, et al. A Switched View of Retinex: Deep Self-Regularized Low-Light Image Enhancement[J]. Neurocomputing, 2021, 454: 361-372.

攻读学位期间发表的学术论文

一、 读硕期间发表的学术论文

- [1]. Wang J, Ge Y, Han C, et al. Nighttime Lane Detection Based on Retinex Theory and Self-Attention Distillation[C]. Proceedings of the 2023 35th Chinese Control and Decision Conference (CCDC). IEEE, 2023: 1801-1806.
- [2]. Wang J, Ge Y, Zhao J, et al. Multi-Feature Fusion Dehazing Based on CycleGAN [J]. AI Communications, 2024. In press.

致谢

三年的时间转瞬即逝，在论文完成之际，感激之情溢于言表。

首先十分感谢我的导师葛愿教授，您的严谨治学态度为我树立了学术典范，在研究过程中您的悉心指导使我得以构建起扎实的研究基础。每一次组会，您都耐心教导我，并给我指明正确的前进方向。您的言传身教，使我受益匪浅，也让我明白科研不止是对未知的探寻，更是对自我的不断提升，这将是我不受用不尽的宝贵财富。

其次感谢我的父母，他们在我的人生旅途中扮演了无可替代的角色，在我遇到困难，停滞不前时，他们就如同坚固的避风港湾，给予我源源不断的力量。谨以此页，深情怀抱养育之恩。在求学岁月里，父母以无言的爱为我筑梦导航。父亲的坚强背影，母亲的温暖笑容，是我勇往直前的力量。每一项成就背后，都凝聚着他们默默的奉献与支持。感谢你们，不仅给予生命，更教会我如何去生活，去爱，去探索。

最后，我要衷心地感谢我在读研期间朝夕相处的室友以及同门伙伴们（黄跃祖、李雅莉、昌婷婷、汪鹏程、王洁、程洪德、苏世林、申国文）。这段共同度过的时光，无疑是我人生历程中极其宝贵的一部分。正是有了他们的陪伴与支持，我在学术研究的道路上才会坚定前行。

尚德敏学 唯实惟新

