**上海大学 2023 ～ 2024 学年 夏 季学期**

**人工智能应用联合大作业**  **课程报告**

课 程 名 称： 人工智能应用联合大作业 课程编号： 08B2A005

报 告 题 目： 逼真的车牌去模糊数据集和模型

学生学号姓名： 王妍骅(21122892),浦扬茗(21122893),吴睿敏(21122894),蒋鑫怡(21122895),徐蓓蓓(21122896),王子棠(21122897)

课程报告评语：

课程报告成绩：

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 文献总结（20分） | | 算法实验（50分） | | 报告综合评价（30分） | | | 总分  （100分） |
| 广度 | 深度 | 工作量 | 效果 | 规范性 | 质量 | 创新 |
| 10分 | 10分 | 30分 | 20分 | 10分 | 10分 | 10分 | 100 |
|  |  |  |  |  |  |  |  |

任课教师:

评阅日期: 年 月 日

逼真的车牌去模糊数据集和模型

王妍骅(21122892),浦扬茗(21122893),吴睿敏(21122894),蒋鑫怡(21122895),徐蓓蓓(21122896),王子棠(21122897)

**摘要：**

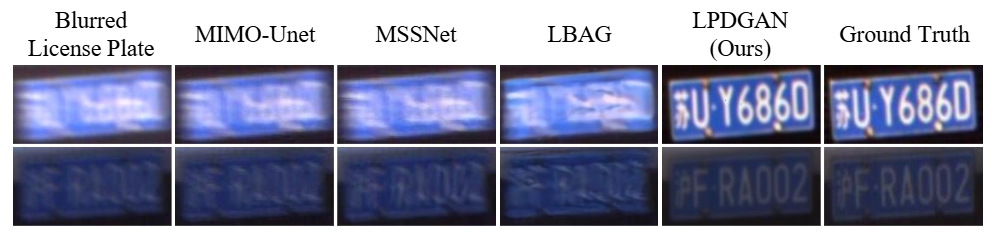
车牌识别是智能交通管理系统中的一项关键任务。然而，由于快速行驶车辆的运动模糊，实现准确识别的挑战仍然存在。尽管图像合成方法在现有的去模糊和识别算法中得到了广泛使用，但它们在现实场景中的有效性仍未得到证实。为了解决这个问题，引入了第一个名为车牌模糊（LP-Blur）的大规模车牌去模糊数据集，该数据集由双摄像头系统捕获，并通过后处理管道进行处理，以避免错位问题。然后，提出了一个车牌去模糊生成对抗网络（LPDGAN）来解决车牌去模糊问题：

1. 一个特征融合模块来集成多尺度潜在代码；
2. 一个文本重建模块，通过文本模态恢复结构；
3. 一个分区判别器模块，以增强模型对每个字母中细节的感知。

广泛的实验验证了LPBlur数据集在模型训练和测试中的可靠性，表明提出的模型在真实的车牌去模糊场景中优于其他最先进的运动去模糊方法。

**1 引言**  
 有效识别车牌对于智能交通管理系统至关重要，然而，由于运动模糊，现实场景通常会带来重大挑战。这种模糊使车牌无法读取，尤其是在涉及高速车辆或弱光条件的情况下。这些问题在夜间或恶劣天气下会加剧，导致捕获图像中相当大的运动模糊。为了解决这些问题，我们的研究引入了一个全面的数据集和一个为现实车牌去模糊量身定制的新模型。  
 图像去模糊是计算机视觉中的一项关键任务，专注于将模糊的图像恢复为清晰的图像，以进行准确的观察和识别。该领域的进展在很大程度上取决于相关数据集的开发。当前创建去模糊数据集的方法分为三大类：

1. 使用模糊内核进行合成模糊，这导致在这些合成图像上训练的模型在应用于真实世界的图像时缺乏泛化能力。
2. 通过平均或融合从清晰的帧生成模糊图像，它不能完全模仿真实世界的过曝光异常值[Chang et al.，2021]。

最后，分束系统通过相机抖动捕捉锐利和模糊的图像对，在颜色精度和对齐方面存在潜在问题。每种方法都有助于该领域，但也有影响数据集真实性和实用性的固有限制。  
 随着深度学习的出现，众多基于卷积神经网络的方法浮出水面，在运动去模糊任务中发挥着至关重要的作用。最近，通用对抗网络（GAN）的提出也深刻地影响了图像去模糊领域。尽管取得了这些进步，车牌图像去模糊仍然是一个重大挑战，主要是由于缺乏大规模、定制的数据集。车牌模糊的复杂性，与标准运动模糊相比，其退化更严重，带来了额外的挑战。为了更好地证明现有图像去模糊算法在真实世界模糊车牌图像上的性能，我们评估了几种最先进的模糊车牌图像去模糊算法的性能。  
  
 图1 几个最先进模型的和我们的模型的去模糊车牌图像  
 如图1所示，我们可以得出结论，所有这些方法都不能很好地完成这项任务。这强调了进一步研究专门针对真实世界车牌去模糊的必要性。  
 为了应对这些挑战，一个由大规模配对车牌数据集和专用车牌去模糊模型组成的综合解决方案被提出。数据收集采用具有不同快门速度的双摄像头设置，以同时捕捉锐利和模糊的图像，消除颜色偏差并实现后处理对齐。我们创新的端到端模型利用编码器和潜在融合模块来处理多尺度拉帐篷代码，采用Swin转换器块进行有效的远程建模。为了增强字母重建和文本易读性，我们引入了评估每个字母清晰度的分区鉴别器。使用我们的LPBlur数据集进行了广泛的实验，包括L1损耗、峰值信噪比（PSNR）、结构相似度指数（SSIM）、感知损耗（PerL）和文本Levenshtein距离（TLD）[Levenshtein等人，1966年] 等指标，验证了其是否适合训练和测试。我们提出的模型在真实世界的车牌去模糊场景中优于最先进的运动去模糊方法。  
 总之，研究的主要贡献如下：

1. 提出了一个名为LPBlur的真实世界清晰模糊车牌数据集。该数据集由10,288张配对图像组成，使用我们设计的双摄像头系统在不同的真实道路场景下精心收集，并通过后处理管道进行校正。
2. 介绍了一种新的车牌模糊模型LPDGAN，该模型利用多尺度潜在代码作为参考。它结合了分区鉴别器和文本重建技术，分别通过空间架构和文本信息增强模型生成高质量车牌图像的能力。
3. 大量实验表明，数据集LP-Blur对于模型训练和评估非常有效。与其他最先进的（SOTA）去模糊模型相比，我们提出的LPDGAN可以实现21.24%的车牌识别准确率提高。

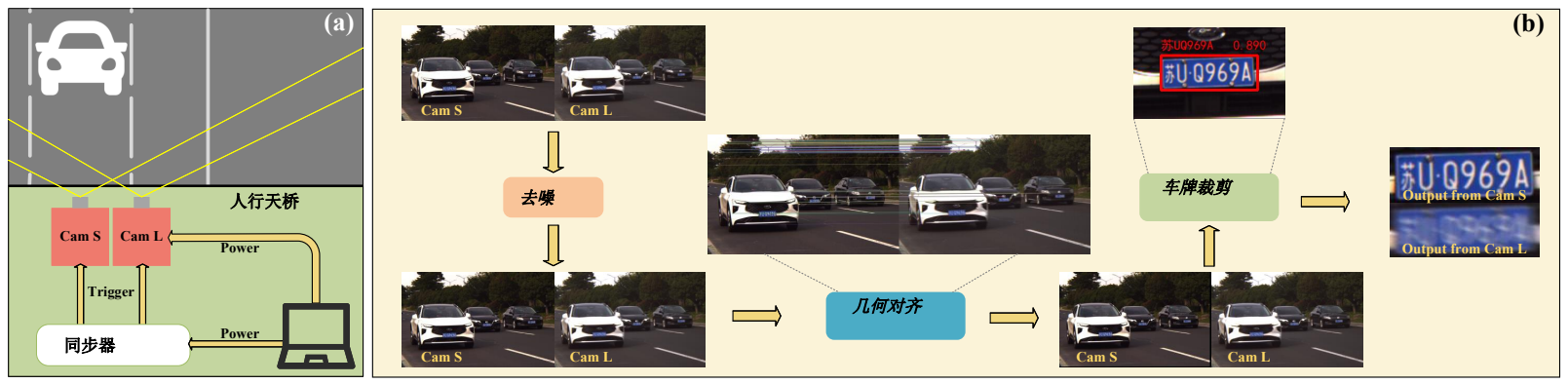
**2 文献总结**

图2 (a）在人行天桥中采集数据的配对图像采集系统示意图（b）配对图像后处理的流水线  
**2.1 模糊图像数据集** 图像去模糊依赖于成对的锐利模糊图像数据集。传统上，模糊图像是通过使用均匀或非均匀模糊内核对锐利图像进行卷积来生成的。因此，一些研究人员试图在振动相机时捕捉锐利帧序列，将这些帧序列平均或融合成相应的运动模糊图像。HIDE数据集由平均11个连续帧创建，中央帧作为锐利图像。Blur-DVS数据集和MSCNN（WILD）数据集的收集中采用了相同的策略。然而，当模型在使用上述方法生成的此类合成数据集上进行训练时，它们缺乏对现实世界场景的通用性。最近，某些研究人员收集了使用分束系统的真实锐利模糊图像对。他们将两个摄像头以固定的角度放置，以确保两个图像共享相同的视野，正如Rim等人的作品中所描述的那样。和Li等人…然而，由于分束系统的固有问题，这种方法可能会导致配对图像中的偏色差异。

大多数传统方法使用自然图像的先验来估计潜在图像或模糊内核。然而，上述技术通过预测均匀图像模糊的假设具有一定的局限性。为了解决这个问题，一些方法在像素级别估计模糊内核，从而适应更复杂的模糊情况。随着深度学习技术的出现，图像去模糊取得了长足的进步，应用深度学习来预测模糊内核或潜在图像以获得清晰的图像。在的工作中，提出了一种基于CNN的方法来预测块级运动模糊的概率分布。龚等人介绍了一种直接估计模糊图像的运动流的方法，从估计的运动流中恢复非模糊图像。MIMO-UNet部署了一个多输入多输出的单Unet网络来模拟多级Unet，以实现不同图像尺度的降噪。MSSNet通过使用反映模糊尺度的阶段配置、跨尺度信息传播方案和基于像素洗牌的多尺度方案来增强去模糊网络的性能。XYDeblur通过在解码器中使用旋转和共享内核，进一步提高了网络效率和去模糊性能。

**2.2 基于GAN的去模糊** 近年来，随着GAN的出现，它们在图像去模糊领域的应用取得了显著的成功。Deblagan首先提出了一种运动去模糊的端到端学习方法，同时也引入了一种新的模糊生成方法。Deblagan-v2引入了基于相对条件GAN框架的双尺度鉴别器，并将特征金字塔纳入去模糊过程，从而允许骨干网络的灵活替换。MSG-GAN通过允许从鉴别器到生成器的多尺度梯度网络，解决了GAN中从鉴别器到生成器的传输过程中真假支持分布之间重叠不足的问题。FCL-GAN设计了轻量级域转换单元（LDCU）和无参数频域对比单元（PFCU）以实现轻量级性能和性能优势。

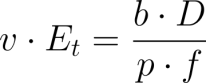
**2.1 改进的LPBlur数据集**

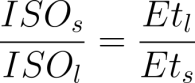
**运动模糊的原因**

运动模糊是指在捕捉物体运动时观察到的可感知的条纹效果。在捕捉过程中，进入感光组件的光量与相机基本参数之间的相关性满足以下方程：  
 ， (1)

其中La表示相机接收到的光子数量，SL表示场景的光强度，ISO表示相机的ISO值，Et是曝光时间，Ap表示相机光圈大小。相机会根据照明情况在一定范围内自动调整这些光圈。例如，相机在低照明设置下增加光圈和曝光时间以捕捉足够的光线。在延长的曝光时间内，快速移动的物体会在一帧内留下轨迹，从而产生运动模糊效果。

**配对图像采集** 为了收集成对的锐利模糊图像，我们使用了两台相同的科学相机，它们设置了不同的曝光时间。如图2a所示，相机S和相机L被固定在三脚架上平行，以保持与地面的全向距离。具体来说，相机S被设置为极短的曝光时间Ets，用于收集锐利的图像，而相机L被设置为相对较长的曝光时间Etl，用于获取模糊的图像。这两个相机都通过同步器与计算机连接，这确保了曝光开始的同步，并且两个相机都捕捉相同的场景。

场景是在不同的位置拍摄的，包括道路的上方、右侧和左侧，以保证数据集的多样性。此外，根据道路和照明条件，我们根据后续方程动态调节摄像机曝光时间：  
 ， (2)  
 其中v表示车辆速度，Et表示曝光时间，b表示模糊像素，D是车辆和相机之间的距离，f表示相机焦距，p是传感器上的像素边缘长度。鉴于单个车辆的实际速度是不确定的，我们将高速路段的图像捕捉标准化，监管限速为70公里/小时，以确保D保持在10-20米的范围内。

此外，为了确保两台相机之间的曝光相等，我们使它们的曝光时间和ISO值满足以下等式：  
 ， (3)  
 其中ISO是相机S和ISOl相机L的ISO值。但是，ISO值的变化会导致图像噪声水平的变化，在后处理中，我们为锐利的图像加入了去噪步骤。  
**2.2 数据后处理** 如图2(b)所示，配对图像后处理包括去噪、几何对齐和车牌裁剪。

**去噪**

由于两个相机的ISO设置不同，相机S和相机L捕获的图像具有不同的噪声水平。因此，在将RAW图像转换为RGB格式的过程中，在白平衡、颜色映射和伽马校正之后采用小波去噪。

**几何对齐**

相机S和L捕捉锐利和模糊的图像对，即使它们从左到右紧密对齐以最小化差异。为了完美地对齐这些图像对，我们首先将没有任何移动车辆的静态图像对作为每个场景的参考对。然后，采用增强相关系数最大化来估计参考图像对锐利和模糊之间的几何变换。最后，将估计的几何变换应用于同一场景中的图像对。

**车牌裁剪**  
 一个预训练活在当下v5[Jocher，2020]和一个预训练CRNN[shi et al.，2016]模型，便于在标准条件下车牌的检测和识别，两个模型都在CCPD[徐et al.，2018]车牌数据集上预训练。按照几何对齐，预训练活在当下v5和CRNN检测和识别配对图像中每个锐利车牌的边界框，然后使用相同的检测坐标裁剪锐利和模糊的图像。总之，我们收集了10,288个图像对，原始图像大小为1920×1220。后处理裁剪图像大小为224×112，模糊大小为20-50像素。其中，在正常光照条件下捕获5672对，在弱光条件下捕获4616对，其中在雨天环境条件下捕获1000对。

**3 研究内容**

**3.1 理论概述**

该论文使用设计的图像到图像转换框架（称为LPDGAN）来提高车牌图像的清晰度。首先构建了一个多尺度特征提取和融合模块，旨在对输入的模糊图像进行有效编码。随后，使用图像解码器生成清晰和高质量的图像。为了进一步增强整体画质，集成了全局鉴别器和分区鉴别器用于训练。此外，还结合了一个文本重建模块来丰富嵌入在生成的车牌图像中的语义信息。

**3.2 多尺度特征提取与潜在融合模块**

**3.2.1 特征提取**

在现实世界的场景中，受运动模糊影响的车牌图像经常表现出复杂的降级，包括噪声、低分辨率和重影效应。特征编码器E专门用于解决这些降级，提取必要的图像特征以供后续处理。特别是，Swin transformer[Liu et al.，2021]因其能够通过自我注意力机制捕获全局信息而被选中。这对于解决运动模糊车牌图像中经常出现的细长重影伪影至关重要。为了适应由于不同捕获距离而导致的车牌图像大小的变化，采用了多尺度特征提取策略。如图3所示，这种方法有助于在每个尺度上对特征进行编码，对于i=1,2,3表示为。

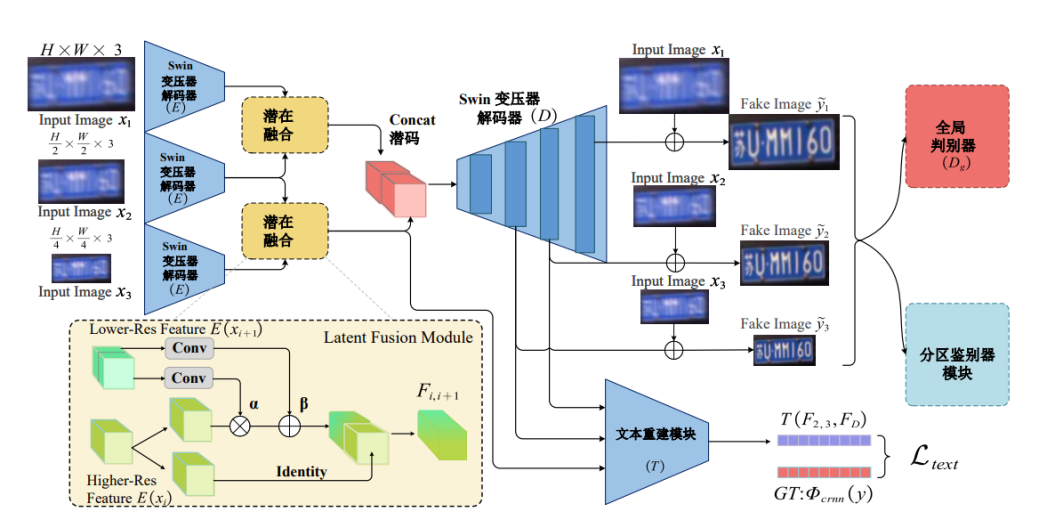


图3 车牌去模糊生成对抗网络框架图

**3.2.2 潜在融合**

基于空间特征变换（SFT）[Wang et al.，2018]，进一步提出了一个潜在融合模块F。该模块基于仿射变换设计，以有效地集成获得的多尺度特征。具体来说，对于和的融合，首先沿着通道维度拆分。然后通过一系列卷积层对每个部分进行处理，以导出融合参数α和β。这些参数被用于通过缩放和移位操作来调制。此外，使用跳过连接重新整合原始的，然后沿着通道维度与修改后的组合。这种融合过程也应用于和之间。

其中和表示沿通道维度划分为的两部分。

**3.2.3 解码器和判别器**

解码器D由一系列Swin transformer和补丁扩展操作组成。与编码器类似，解码器被设计为以多尺度方式生成清晰的图像，输出图像分别表示为，和。

实验设计了两个判别器：

1）全局判别器（Dg）增强了恢复图像中的整体空间和颜色信息；

2）分区判别器（Dp）通过检查车牌图像中随机选择的n个字母分区，专门关注细化字符信息。Dp的结构如图3所示。它识别和定位真实图像y和生成图像中的字母位置。在此之后，随机选择n个分区图像供部分判别器评估。在训练的早期阶段，当模型产生清晰图像的能力仍在发展时，生成的图像可能无法以高精度识别。为了解决这个问题，对y和应用了平均分区方法，最初将n设置为7。随着训练的进行，预训练YOLO V5模型用于精确的字母检测，分区数n设置为3。



分区鉴别器模块

分区判别器

随机选择

Fake

Partitions

Fake Image

P

Real Image

Real

Partitions

图4 分区判别器模块的架构

在实验中，使用WGAN-GP来训练模型。特别是，全局判别器的对抗损失可以公式如下：

同样，对于分区判别器Dp，公式如下：

其中P是分区运算，λgp1和λgp2是用于控制梯度惩罚的加权参数，将两个判别器应用于三种不同尺度的输出。除了对抗性损失，还包含重建损失rec，定义如下：

其中表示预训练的VGG-19网络[Si-monyan和Zisserman，2014]，实验使用来自第8、15和22个ReLU层的特征图来比较真实图像和生成图像之间的浅纹理和深特征。

**3.3 文本重建模块**

实验还结合了一个文本重建模块T，该模块专门设计用于增强模型准确解释车牌图像上字符的能力。T将解码器的中间特征FD与融合潜在代码F2,3沿通道维度合并。然后，该组合特征遍历一系列卷积和线性层，重新封装在表示识别文本的向量中。目前，预训练的CRNN模型从真实图像中提取真实文本向量。论文计算文本重建模块T的输出向量和真实文本向量之间的L1损失。

其中FD是从Decoder的中间层获得的特征图，ψcrnn代表预训练的CRNN模型。

**3.4 完整目标**

实验的完整目标是：

其中、和控制不同目标的相对重要性。实验的目标是解决：

**4 实验结果与分析**

**4.1 实验数据**

本次实验提出的LPBlur数据集是针对车辆车牌去模糊任务而设计的，它是首个大规模的、专为现实世界车牌去模糊场景打造的数据集。

**4.1.1 数据集特点**

LPBlur数据集包含10,288对清晰与模糊的车牌图像对，这些图像都是在多样化的实际道路环境下通过特制的双摄像头系统精心收集的。每一对图像都对应一个清晰的车牌图像及其对应的运动模糊版本，确保了数据集能够反映真实环境下的车牌模糊情况。同时，数据采集利用了双摄像头系统，两个摄像头设置不同的快门速度，以同时捕捉同一视场下的清晰图像和模糊图像。这种方法有效避免了颜色偏差，并且通过后期处理流程对齐图像，解决了因相机运动或对象移动造成的图像错位问题。此数据集特别关注了高速行驶车辆和低光条件下的车牌识别难题，这些条件下产生的运动模糊极大地影响了车牌的可读性，特别是在夜间或恶劣天气情况下。

**4.1.2 数据集划分和评估**

数据集被划分为训练集和测试集，其中训练集包含9,288对图像，测试集有1,000对图像，进一步细分为正常光照条件下的500对和低光照条件下的500对，以全面评估模型在不同条件下的表现。同时，为了评估模型性能，采用了一系列量化的评估指标，包括峰值信噪比(PSNR)、结构相似度指数(SSIM)、感知损失(Perceptual Loss)以及文本编辑距离(Text Levenshtein Distance，TLD)。特别是TLD用来衡量生成的车牌图像上的文字识别准确率，直接关联到车牌识别系统的最终效果。

**4.2 实验方案**

**4.2.1 数据预处理**

在论文中，为了更好地识别车牌号，将数据集分成三部分：清晰车牌图片，模糊车牌图片以及车牌号标签文件，其中车牌号标签文件存放了每张车牌图片的文件名和车牌号。由于作者并未提供车牌号标签文件，为了训练LPDGAN网络，需要我们自己利用数据集中的清晰车牌图片生成标签文件。

我们选择了基于深度学习的文字识别模型CRNN [ B Shi, X Bai and C Yao (2017)]，该模型将CNN与RNN结合，将文本识别转化为时序依赖的序列学习问题。为了让CRNN从车牌图片中识别出车牌号，我们使用中科大开源的CCPD [ Xu, Z. et al. (2018)] 中国城市车牌开源数据集对CRNN进行训练，训练100轮，batchsize设置为5，学习率0.001。训练完毕后使用预训练好的CRNN模型对清晰车牌图片进行文字识别，输出标签文件。

**4.2.2 模型训练**

训练设置与论文设置一样，使用Adam优化器 [Diederik P Kingma and Jimmy Ba (2014)]，batchsize为7，训练200轮，学习率为0.0001。损失函数采用对抗损失、L1损失和字符识别损失进行优化。所有实验均在NVIDIA GeForce RTX 3060上进行，显存12GB，使用PyTorch深度学习框架实现。

**4.2.3 训练过程**

1. 输入模糊图像：将模糊车牌图像输入生成器，生成对应的清晰车牌图像。

2. 计算对抗损失：判别器判断生成图像和真实清晰图像的真假，生成器通过对抗损失提升生成图像的真实性。

3. 计算L1损失：生成图像与真实清晰图像之间的L1损失，用于优化图像的整体清晰度。

4. 计算字符识别损失：利用车牌号标签，计算生成图像的字符识别损失，指导生成器优化生成的车牌字符，提高识别准确性。

5. 更新生成器和判别器：通过梯度下降算法，分别更新生成器和判别器的参数。

**4.2.4 模型评估**

为了评估模型的性能，我们采用了峰值信噪比（PSNR）和结构相似性指数（SSIM）作为主要评价指标。PSNR用于衡量图像的复原质量，数值越高表明复原效果越好；SSIM用于衡量复原图像与清晰图像在结构上的相似程度，数值越接近1表明图像质量越好。

**4.3 结果分析**

**4.3.1 训练结果**

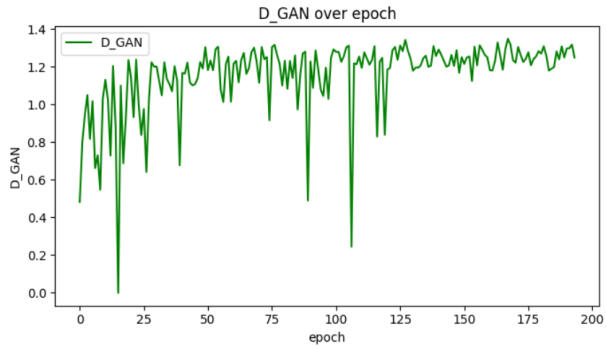
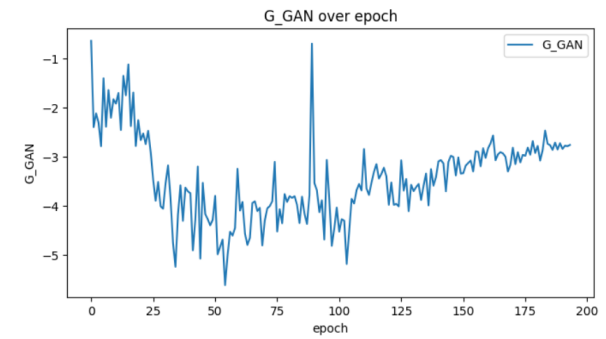
训练200个epochs后我们得到了5个模型：latest\_net\_D\_smallblock，latest\_net\_G.pth，latest\_net\_D，latest\_net\_D1.pth和latest\_net\_D2.pth。其中，latest\_net\_G.pth是我们最终用于测试数据集的模型，存储了生成器网络训练到目前为止的最优权重配置。另外四个模型都是判别器的权重，latest\_net\_D\_smallblock的设计更为简单、专注于像素级的特征鉴别，用于捕捉生成图像中的细微失真和局部不一致性。即这个模型专注于提高生成图像在细节上的真实度。相比之下，latest\_net\_D 保存的是主要判别器的权重，关注的是图像的整体结构和语义内容，而不仅仅是像素级别的匹配。Loss log为损失函数的日志文件，用于记录训练过程中的损失函数值随时间的变化情况。

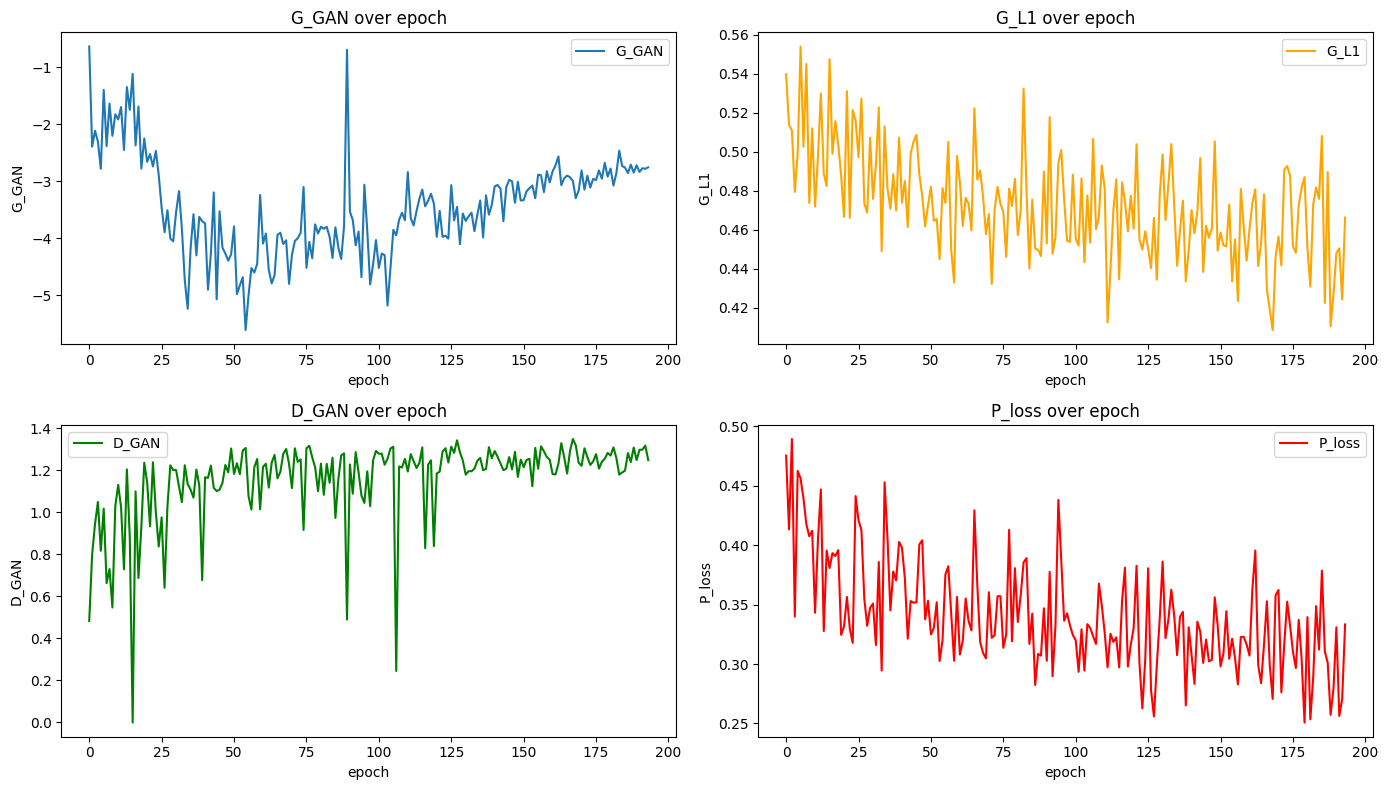
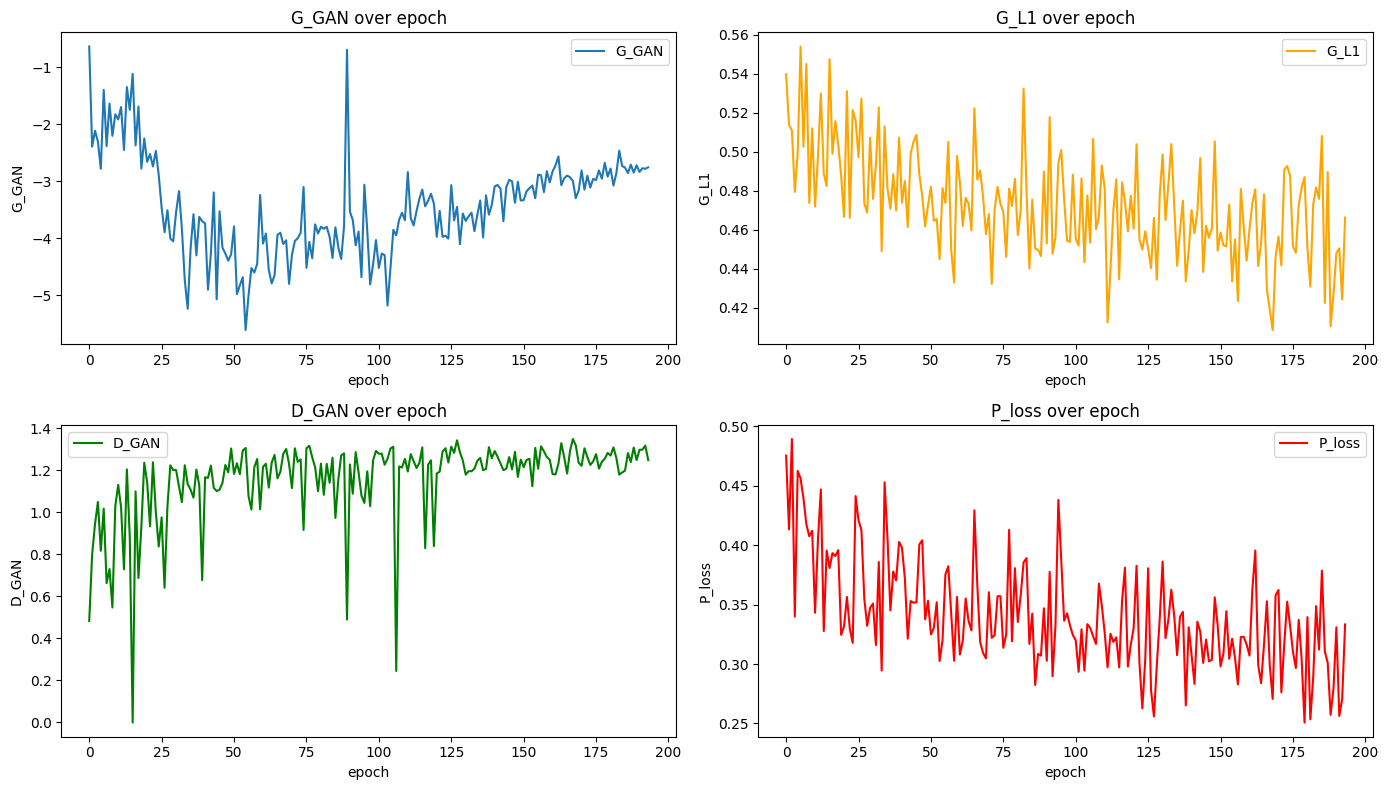
**4.3.2 损失函数可视化**

G\_GAN衡量生成器欺骗判别器的能力，通过计算判别器误以为生成的图像为真的概率与期望之间的差距，促进生成器生成更接近真实数据的样本。在训练的初始阶段，生成器生成的图像质量通常不高，难以欺骗判别器。因此，判别器对合成图像的预测很可能是负的或偏向于非真实的预测，导致生成器对抗损失较高，反映出生成器性能的不足。随着训练的进行，生成器通过反向传播和梯度更新逐步学习如何生成更接近真实图像的内容。这意味着判别器开始错误地认为生成的图像为真，导致loss\_G\_GAN值减小。此阶段，G\_GAN的值出现了波动，因为生成器和判别器之间的竞争不断升级，生成器的进步会促使判别器增强其区分能力，反之亦然，形成一个动态的优化过程。

D\_GAN是判别器（Discriminator, D）的总对抗性损失，是多个损失项的组合，包括来自不同样本集（loss\_D\_fake, loss\_D\_real）的损失。最初，判别器不够强大，难以准确区分真实图像和生成的假图像，导致loss\_D\_fake较大（判别器误以为假图像为真）而loss\_D\_real相对较小（未能完全确认真图像）。随着训练开始，这两个值都会有所调整，但总体loss\_D\_GAN会先有上升趋势，因为判别器在学习区分过程中。随着判别器性能提升，它会更好地识别真假图像，导致loss\_D\_fake减少（成功识别假图像）和loss\_D\_real增加（确信地识别真图像）。这一过程中，损失值可能会经历波动，生成器和判别器达到一种动态平衡状态。

感知损失（Perceptual Loss）有助于提升生成图像的视觉质量和真实感，因为该损失考虑了图像的整体结构和纹理。而L1损失是一种衡量生成图像与目标图像像素级差异的指标，关注于直接减小像素值之间的绝对差。相比于L2损失，L1损失更侧重于边缘信息的保持，能够促进生成图像的清晰度。G\_L1低表明生成图像与目标图像在结构和细节上更为接近。由图可知，感知损失和L1损失明显呈下降趋势，整体来说是呈收敛状态的。



  
图5 损失函数可视化

**4.3.3 测试结果**

（1）定性分析

使用训练好的模型对测试集进行测试，得到包含三类图片的图像对输出：fake\_B.png、real\_A.png和real\_B.png。其中，fake\_B.png为生成的去模糊车牌图像；real\_A.png为模糊的车牌原始图像；real\_B.png为清晰的车牌图像，用于作为理想输出的参考。

通过分析得到的去模糊结果表1，可以看到在绝大部分情况下，该算法成功地将模糊的车牌变得更为清晰，使得车牌号码更容易识别(案例3）。然而，也有一些案例中，虽然图像整体变清晰了，但个别字符的识别仍然存在问题。案例1中，“8”这一数字并为完全呈现；案例2中，整体边界线较为模糊，“Q”尤其明显；案例4中，“皖”字模糊不清，其余数字字母相对清晰；案例5中，“A”被识别为“U”，字母识别不清；案例6中“7688D”被误认为“T888D”，数字识别错误。

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 弱光 | | | 常光 | | |
|  | 案例1 | 案例2 | 案例3 | 案例4 | 案例5 | 案例6 |
| **Blurred Image** |  |  |  |  |  |  |
| **Ground Truth** |  |  |  |  |  |  |
| **LPDGAN** |  |  |  |  |  |  |

表1 模糊的、真实的、去模糊的车牌图像

1. 定量分析

为了评估模型性能，采用了一系列量化的评估指标，包括峰值信噪比(PSNR)、结构相似度指数(SSIM)、感知损失(Perceptual Loss)以及文本编辑距离(Text Levenshtein Distance，TLD)，下面是对各指标的进一步介绍：

PSNR是一个广泛用于衡量图像或视频重建质量的客观指标。它通过比较原始图像与处理后的图像之间的均方误差（MSE），然后转换为分贝（dB）来表示。PSNR值越高，表明失真越小，图像质量越好。但是，PSNR主要关注像素级的误差，并不能很好地反映人眼对图像质量的感知。

SSIM是一种更加高级的图像质量评价指标，旨在克服PSNR只考虑像素间差异而忽略结构信息的缺点。SSIM同时考虑了亮度、对比度、结构三方面的相似性，能够更准确地反映人眼对图像质量的主观感受。SSIM的值范围从-1到1，值越接近1表示两张图像的结构信息越相似，图像质量越高。

PerL是一个在深度学习领域中常用的损失函数，特别是在图像生成、超分辨率和风格迁移等任务中。它基于预训练的卷积神经网络的中间层特征来计算损失，目的是使生成的图像在视觉上更接近真实图像，不仅考虑像素级的差异，还考虑了高层特征的相似性。

TLD是一种衡量两个字符串之间差异程度的指标。它定义为通过插入、删除、替换操作将一个字符串转换为另一个字符串所需的最少单字符编辑操作次数。它可以用来衡量生成的车牌图像上的文字识别准确率，直接关联到车牌识别系统的最终效果。

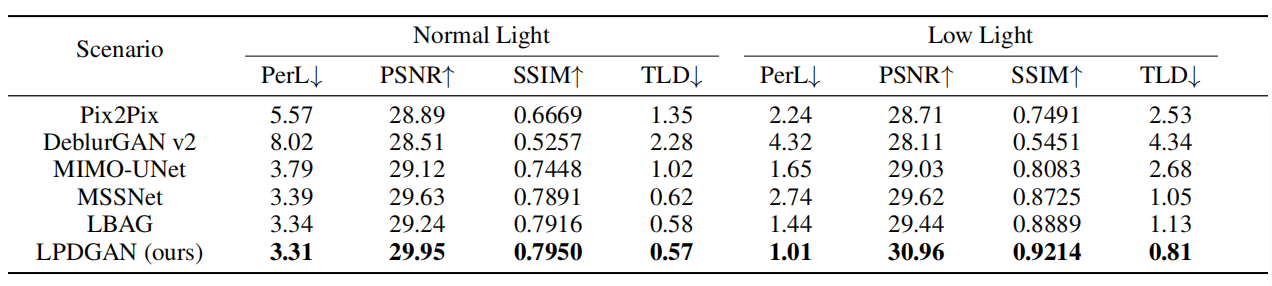


表2 LPDGAN在LPBlur上的消融实验

通过实验结果表2，可以对比发现在去模糊效果（前三个指标）方面，LPDGAN在常光和弱光条件下都优于所有其他模型。在常光条件下，LPDGAN实现了PerL为3.31，PSNR为29.95，SSIM为0.795，优于最新的去模糊方法LBAG。在处理弱光图像时，性能差距变得更加明显，与LBAG和MSSNet相比，LPDGAN在PerL、PSNR和SSIM中分别表现出29.8%、4.5%和3.7%的改进。

TLD可以评估车牌文本识别精度，通过对比可以看到，在正常光照条件下，LPDGAN在TLD方面表现出与LBAG相当的性能，但优于所有其他模型。在弱光条件下，LPDGAN是唯一一个TLD低于1的模型。这意味着预训练的CRNN模型在用于识别由LPDGAN生成的去模糊车牌图像时，将获得每个实例少于一个字符的平均误差。因此，LPDGAN总体上具有最佳性能，在两种场景下展示了对车牌进行去模糊的稳健能力。

**5 结论**

在本次实训中，我们研究了运动车牌去模糊问题，复现了论文源码。这项研究引入了第一个大规模车牌去模糊数据集，并通过适当的数据采集方法和后处理解决了颜色偏差和失调问题。此外，鉴于车辆运动引起的模糊程度大大超过了相机抖动引起的模糊程度，论文提出了一种基于多尺度输入和输出的车牌去模糊模型。这包括一个潜在的融合模块、一个文本模态信息的监督模块和一个分区判别器模块。实验结果表明，与当前最先进的去模糊算法相比，此论文的模型表现良好。通过本次复现作业，我们不仅加深了对图像去模糊技术的理解，也提高了对深度学习模型训练和评估的实践能力，这为我们今后在相关领域的研究和应用奠定了坚实的基础。

**6 课程总结**

在本次夏季学期实训课程中，我们小组对运动车牌去模糊问题进行了深入的研究，并且复现了论文中的源码。为实现任务我们进行了不同的分工：分析数据集；对论文中的算法和理论进行理解学习；理解和修正、调试源代码，在合适的服务器上运行；并且对测试结果做分析，优化程序。在各个小组成员的合作下，我们顺利完成了本次任务。

同时，马丽艳老师对我们的帮助是至关重要的。在课程期间遇到问题，老师会耐心指导我们如何利用所学知识解决，帮助我们克服困难，提升实践能力。马老师还会定期组织实时评估和项目展示，及时给予反馈和建议，帮助我们发现并弥补自身的不足之处。通过老师的指导和鼓励，我们在实训课程中不仅在技术水平上有所提升，也在团队协作能力和问题解决能力上有所突破。老师的辛勤付出和悉心教导让我深受启发，并且对算法和模型的理解更深。

除了老师的指导之外，其他小组也对我们的学习和进步起到了重要的帮助。通过老师每周组织的汇报，在聆听别的小组讲述自己新颖的思路时，我们会积极和他们进行探讨，学以致用进一步优化我们的模型和代码，在这样相互竞争相互合作的情况下也进一步增强了团队协作能力和沟通能力。

综上，本次夏季实训课程提高了我们对深度学习模型和理论算法的理解和使用，也强化了小组合作沟通的能力，使我们受益匪浅。

参考文献：

[CRNN] B Shi, X Bai and C Yao (2017). An End-to-End Trainable Neural Network for Image-Based Sequence Recognition and Its Application to Scene Text Recognition, in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 11, pp. 2298-2304.

[CCPD] Xu, Z. et al. (2018). Towards End-to-End License Plate Detection and Recognition: A Large Dataset and Baseline. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (eds) Computer Vision – ECCV 2018. ECCV 2018. Lecture Notes in Computer Science(), vol 11217. Springer, Cham.

[Adam] Diederik P Kingma and Jimmy Ba (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

[Chang et al., 2021] Meng Chang, Chenwei Yang, HuajunFeng, Zhihai Xu, and Qi Li. Beyond camera motion blur removing: How to handle outliers in deblurring. IEEE Transactions on Computational Imaging, 7:463– 474, 2021.

[Cho and Lee, 2009] Sunghyun Cho and Seungyong Lee.Fast motion deblurring. In ACM SIGGRAPH Asia 2009 papers, pages 1–8. 2009.

[Cho et al., 2021] Sung-Jin Cho,Seo-Won Ji, Jun-Pyo Hong, Seung-Won Jung, and Sung-Jea Ko. Rethinking coarse-to-fine approach in single image deblurring. In Proceedings of the IEEE/CVF international conference on computer vision, pages 4641–4650, 2021.

[Evangelidis and Psarakis, 2008] Georgios D Evangelidis and Emmanouil Z Psarakis.Parametric image alignment using enhanced correlation coefficient maximization. IEEE transactions on pattern analysis and machine intel-ligence, 30(10):1858–1865, 2008.

[Fergus et al., 2006] Rob Fergus, Barun Singh, Aaron Hertzmann, Sam T Roweis, and William T Freeman. Removing camera shake from a single photograph. In Acm Siggraph 2006 Papers, pages 787–794. 2006.

[Gong et al., 2017] Dong Gong, Jie Yang, Lingqiao Liu,Yanning Zhang, Ian Reid, Chunhua Shen, Anton Van Den Hengel, and Qinfeng Shi. From motion blur to mo-tion flow: A deep learning solution for removing heteroge-neous motion blur. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 2319– 2328, 2017.

[Hyun Kim et al., 2013] Tae Hyun Kim, Byeongjoo Ahn,and Kyoung Mu Lee. Dynamic scene deblurring. In Pro-ceedings of the IEEE international conference on com-puter vision, pages 3160–3167, 2013.

[Isola et al., 2017] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 1125–1134, 2017.  
[Ji et al., 2022] Seo-Won Ji, Jeongmin Lee, Seung-Wook Kim, Jun-Pyo Hong, Seung-Jin Baek, Seung-Won Jung, and Sung-Jea Ko. Xydeblur: divide and conquer for sin-gle image deblurring. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 17421–17430, 2022.

[Jiang et al., 2020] Zhe Jiang, Yu Zhang, Dongqing Zou,Jimmy Ren, Jiancheng Lv, and Yebin Liu. Learning eventbased motion deblurring. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 3320–3329, 2020.

[Jocher, 2020] Glenn Jocher. YOLOv5 by Ultralytics, May2020.

[Karnewar and Wang, 2020] Animesh Karnewar and OliverWang. Msg-gan: Multi-scale gradients for generative adversarial networks. In Proceedings of the IEEE/CVF con-ference on computer vision and pattern recognition, pages 7799–7808, 2020.

[Kimet al., 2022] Kiyeon Kim, Seungyong Lee, and Sunghyun Cho. Mssnet: Multi-scale-stage network for single image deblurring. In European Conference on Computer Vision, pages 524–539. Springer, 2022.

[Kingma and Ba, 2014] Diederik P Kingma and Jimmy Ba.Adam: A method for stochastic optimization.arXivpreprint arXiv:1412.6980, 2014.

[Kohleret al., 2012] Rolf Kohler, Michael Hirsch, Betty¨Mohler, Bernhard Scholkopf, and Stefan Harmeling.¨ Recording and playback of camera shake: Benchmark-ing blind deconvolution with a real-world database. In Computer Vision–ECCV 2012: 12th European Conference on Computer Vision, Florence, Italy, October 7-13, 2012, Proceedings, Part VII 12, pages 27–40. Springer, 2012.

[Kupyn et al., 2018] Orest Kupyn, Volodymyr Budzan,Mykola Mykhailych, Dmytro Mishkin, and Jiˇr´ı Matas. Deblurgan: Blind motion deblurring using conditional adversarial networks. In Proceedings of the IEEE confer-ence on computer vision and pattern recognition, pages 8183–8192, 2018.

[Kupyn et al., 2019] Orest Kupyn, Tetiana Martyniuk, JunruWu, and Zhangyang Wang. Deblurgan-v2: Deblurring(orders-of-magnitude) faster and better. In Proceedings of the IEEE/CVF international conference on computer vi-sion, pages 8878–8887, 2019.

[Lai et al., 2016] Wei-Sheng Lai, Jia-Bin Huang, Zhe Hu,Narendra Ahuja, and Ming-Hsuan Yang. A comparative study for single image blind deblurring. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 1701–1709, 2016.

[Levenshtein and others, 1966] Vladimir I Levenshtein et al.Binary codes capable of correcting deletions, insertions,and reversals. In Soviet physics doklady, volume 10, pages 707–710. Soviet Union, 1966.

[Levin et al., 2009] Anat Levin, Yair Weiss, Fredo Durand,and William T Freeman. Understanding and evaluating

blind deconvolution algorithms. In 2009 IEEE conference on computer vision and pattern recognition, pages 1964– 1971. IEEE, 2009.

[Li et al., 2023] Haoying Li, Ziran Zhang, Tingting Jiang,Peng Luo, Huajun Feng, and Zhihai Xu. Real-world deep local motion deblurring. In Proceedings of the AAAI Con-ference on Artificial Intelligence, volume 37, pages 1314– 1322, 2023.

[Liu et al., 2020] Wei Liu, Qiong Yan, and Yuzhi Zhao.Densely self-guided wavelet network for image denoising.In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, pages 432– 433, 2020.

[Liu et al., 2021] Ze Liu, Yutong Lin, Yue Cao, Han Hu,Yixuan Wei, Zheng Zhang, Stephen Lin, and Baining Guo.Swin transformer: Hierarchical vision transformer using shifted windows. In Proceedings of the IEEE/CVF in-ternational conference on computer vision, pages 10012– 10022, 2021.

[Nah et al., 2017] Seungjun Nah, Tae Hyun Kim, and Kyoung Mu Lee. Deep multi-scale convolutional neural network for dynamic scene deblurring. In Proceedings of the IEEE conference on computer vision and pattern recogni-tion, pages 3883–3891, 2017.

[Noroozi et al., 2017] Mehdi Noroozi, Paramanand Chandramouli, and Paolo Favaro. Motion deblurring in the wild. In Pattern Recognition: 39th German Conference, GCPR 2017, Basel, Switzerland, September 12–15, 2017, Proceedings 39, pages 65–77. Springer, 2017.

[Ramakrishnan et al., 2017] Sainandan Ramakrishnan,Shubham Pachori, Aalok Gangopadhyay, and Shanmuganathan Raman. Deep generative filter for motion deblurring. In Proceedings of the IEEE international con-ference on computer vision workshops, pages 2993–3000, 2017.

[Ren et al., 2017] Dongwei Ren, Wangmeng Zuo, David Zhang, Jun Xu, and Lei Zhang. Partial deconvolution with inaccurate blur kernel. IEEE transactions on image pro-cessing, 27(1):511–524, 2017.

[Ren et al., 2018] Wenqi Ren, Jiawei Zhang, Lin Ma, Jinshan Pan, Xiaochun Cao, Wangmeng Zuo, Wei Liu, and Ming-Hsuan Yang. Deep non-blind deconvolution via gen-eralized low-rank approximation. Advances in neural in-formation processing systems, 31, 2018.

[Rim et al., 2020] Jaesung Rim, Haeyun Lee, Jucheol Won,and Sunghyun Cho. Real-world blur dataset for learning and benchmarking deblurring algorithms. In Computer Vision–ECCV 2020: 16th European Conference, Glas-gow, UK, August 23–28, 2020, Proceedings, Part XXV 16, pages 184–201. Springer, 2020.

[Shan et al., 2008] Qi Shan, Jiaya Jia, and Aseem Agarwala.High-quality motion deblurring from a single image. Acm transactions on graphics (tog), 27(3):1–10, 2008.

[Shen et al., 2019] Ziyi Shen, Wenguan Wang, Xiankai Lu,Jianbing Shen, Haibin Ling, Tingfa Xu, and Ling Shao.Human-aware motion deblurring. In Proceedings of the IEEE/CVF International Conference on Computer Vision, pages 5572–5581, 2019.

[Shi et al., 2016] Baoguang Shi, Xiang Bai, and Cong Yao.An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition. IEEE transactions on pattern analysis and machine intelligence, 39(11):2298–2304, 2016.

[Simonyan and Zisserman, 2014] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition.arXiv preprintarXiv:1409.1556, 2014. [Sun et al., 2013] Libin Sun, Sunghyun Cho, Jue Wang, and James Hays. Edge-based blur kernel estimation using patch priors. In IEEE international conference on com-putational photography (ICCP), pages 1–8. IEEE, 2013.

[Sun et al., 2015] Jian Sun, Wenfei Cao, Zongben Xu, and Jean Ponce. Learning a convolutional neural network for non-uniform motion blur removal. In Proceedings of the IEEE conference on computer vision and pattern recogni-tion, pages 769–777, 2015.

[Tao et al., 2018] Xin Tao, Hongyun Gao, Xiaoyong Shen, Jue Wang, and Jiaya Jia. Scale-recurrent network for deep image deblurring. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 8174– 8182, 2018.

[Wang et al., 2018] Xintao Wang, Ke Yu, Chao Dong, and Chen Change Loy. Recovering realistic texture in image super-resolution by deep spatial feature transform. In Pro-ceedings of the IEEE conference on computer vision and pattern recognition, pages 606–615, 2018.

[Whyte et al., 2012] Oliver Whyte, Josef Sivic, Andrew Zisserman, and Jean Ponce. Non-uniform deblurring for shaken images. International journal of computer vision, 98:168–186, 2012.

[Xu et al., 2018] Zhenbo Xu, Wei Yang, Ajin Meng, Nanxue Lu, Huan Huang, Changchun Ying, and Liusheng Huang. Towards end-to-end license plate detection and recogni-tion: A large dataset and baseline. In Proceedings of the European conference on computer vision (ECCV), pages 255–271, 2018.

[Zhang et al., 2023] Xiang Zhang, Lei Yu, Wen Yang, Jianzhuang Liu, and Gui-Song Xia. Generalizing eventbased motion deblurring in real-world scenarios. In Pro-ceedings of the IEEE/CVF International Conference on Computer Vision, pages 10734–10744, 2023.

[Zhao et al., 2022] Suiyi Zhao, Zhao Zhang, Richang Hong,Mingliang Xu, Yi Yang, and Meng Wang. Fcl-gan: A lightweight and real-time baseline for unsupervised blind image deblurring. In Proceedings of the 30th ACM In-ternational Conference on Multimedia, pages 6220–6229, 2022.