



计算机应用
Journal of Computer Applications
ISSN 1001-9081, CN 51-1307/TP

《计算机应用》网络首发论文

题目：基于多维空间卷积信息增强的低质车牌信息超分辨率重建
作者：张睿，惠永科，张延军，潘理虎
收稿日期：2024-02-02
网络首发日期：2024-05-06
引用格式：张睿，惠永科，张延军，潘理虎. 基于多维空间卷积信息增强的低质车牌信息超分辨率重建[J/OL]. 计算机应用.
<https://link.cnki.net/urlid/51.1307.TP.20240428.2331.021>



网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

基于多维空间卷积信息增强的低质车牌信息超分辨率重建

张睿^{1*}, 惠永科¹, 张延军², 潘理虎¹

(1.太原科技大学 计算机科学与技术学院, 太原 030024;

2.太原科技大学 机械工程学院, 太原 030024)

(*通信作者电子邮箱 zhangrui@tyust.edu.cn)

摘要: 现有交通监控终端采集到的车辆影像往往存在远景低分辨率现象, 并伴随有强噪、模糊、过曝、欠曝等一些不确定性像素影响因素, 导致车牌信息智能识别的精度难以保证。针对上述问题, 提出基于多维空间卷积信息增强的低质车牌信息超分辨率网络(LL-SR)。首先, 通过卷积挖掘特征点在空间和通道上协同卷积计算的关联关系, 聚合浅层特征; 其次, 从不同感受野和不同维度挖掘特征图和特征图之间的关联关系, 恢复车牌信息的高频细节; 最后, 对得到的不同尺度特征进行跨通道像素级融合和矫正, 减少无用特征在上下文的传播, 从而实现低质车牌信息的超分辨率重建。通过在数据集太原车牌(LT)和美国车牌(LU)上开展的实验结果表明, 所提网络的峰值信噪比(PSNR)和结构相似性(SSIM)为 26.6824dB、0.8203 和 22.3567、0.7831, 相较于 CARN(fast, Accurate, and lightweight super-resolution with Cascading Residual Network)、NGramSwin(N-Gram in Swin transformers for efficient lightweight image super-resolution)提升了 0.2109~1.7361 dB、0.0057~0.033 和 0.4728~1.4192dB、0.0196~0.0399; 且重建后的车牌信息具有更好的视觉效果。

关键词: 低质车牌信息; 高质量重建; 卷积计算; 交通监控

中图分类号: TP391.4

文献标志码: A

Super-resolution reconstruction for low-quality license plate information based on multi-dimensional spatial convolutional information enhancement

ZHANG Rui^{1*}, HUI Yongke¹, ZHANG Yanjun², PAN Lihu¹

(1.School of Computer Science and Technology, Taiyuan University of Science and Technology, Taiyuan Shanxi, 030024, China;

2. School of Mechanical Engineering, Taiyuan University of Science and Technology, Taiyuan Shanxi, 030024, China)

Abstract: Vehicle images collected by existing traffic monitoring terminals often exhibit low resolution in long range, accompanied by uncertain pixel factors such as strong noise, blur, overexposure, and underexposure, making it difficult to ensure accuracy of intelligent recognition of license plate information. In response to above issues, Super-Resolution reconstruction for Low-quality License plate information based on multi-dimensional spatial convolutional information enhancement (LL-SR) was proposed. First, the correlation of feature points in space and channels mined by convolution was used to aggregate shallow feature. Second, correlation between feature maps and feature maps was mined from different receptive fields and different dimensions to recover high-frequency details of license plate information. Finally, the obtained features of different scales were fused and corrected at the pixel level across channels to reduce spread of useless features in context, thus achieving super-resolution reconstruction of low-quality license plate information. Through experiments conducted on datasets License plate of Tai yuan (LT) and License plates of the United states of america (LU), results showed that Peak Signal to Noise Ratio (PSNR) and Structural SIMilarity (SSIM) of proposed network is 26.68dB, 0.8203 and 22.3567dB, 0.7831, improved by 0.2109~1.7361dB, 0.0057~0.033, and 0.4728~1.4192dB, 0.0196~0.0399 compared to fast, Accurate, and lightweight super-resolution with Cascading Residual Network(CARN) and N-Gram in Swin transformers for efficient lightweight image super-resolution(NGramSwin). Moreover, reconstructed license plate information had better visual effects.

收稿日期: 2024-02-02; 修回日期: 2024-04-24; 录用日期: 2024-04-24。

基金项目: 山西省基础研究计划项目(20210302123216); 山西省产教融合研究生联合培养示范基地项目(2022JD11); 山西省机械产品质量司法鉴定中心企业委托项目(2023254, 2023069); 太原科技大学研究生联合培养示范基地项目(JD2022004)。

作者简介: 张睿(1987—), 男, 山西太原人, 副教授, 博士, CCF 高级会员, 主要研究方向: 智能信息处理、计算机视觉; 惠永科(1998—), 男, 河南南阳人, 硕士研究生, 主要研究方向: 计算机视觉; 张延军(1982—), 男, 山西太原人, 教授, 博士, 主要研究方向: 智能信息处理; 潘理虎(1974—), 男, 山西太原人, 教授, 博士, CCF 高级会员, 主要研究方向: 智能信息处理。

Keywords: low-quality license plate information; high quality reconstruction; convolution computation; traffic monitoring

0 引言

道路场景中低质车牌信息的精准检测是实现智能交通系统的基本要求和前提^[1],但目前采集到的车牌影像往往存在远景低分辨率问题,同时受环境影响还会伴有强噪、模糊、过曝、欠曝等一些不确定性像素影响因素,这些问题对道路场景下低质车牌信息的智能精准识别形成阻碍,使车牌信息的识别精度难以保障。完全通过改进硬件设备提高车牌影像分辨率解决车牌识别准确率的问题将花费高昂的经济成本,因此,如何构建高效的低质车牌信息超分辨率重建网络将道路场景下车牌低分辨率影像重建为高分辨率影像,并对不确定性像素影响因素进行有效消除,具有重要的应用价值和现实意义。

超分辨率重建网络主要为基于插值和基于深度学习的方法。基于插值^[2]的方法计算简单,运算速度快,但是进行超分辨率重建时,并不会消除模糊、强噪和恢复光照失常区域,还会使起车牌影像边缘失真,产生明显的马赛克和锯齿现象。随着深度学习的快速发展,基于深度学习的超分辨率重建方法已成为主流,即通过构建深度神经网络,自动学习低分辨率(Low Resolution, LR)影像到高分辨率(High Resolution, HR)影像的非线性映射,提升影像的高频细节,相较于基于插值的方法具有更好的重建效果,且具有更强的自适应能力^[3]。Liang 等^[4]将 Transformer 架构应用于影像复原任务中,这种创新使网络能够更好地捕捉影像中的全局依赖关系,有助于提高超分辨率性能;Kong 等^[5]构建了一个改进的特征提取器,可以有效地提取边缘和细节,并提出了一种新颖的多阶段热启动训练策略,加速网络收敛并提高精度,最终实现了先进的性能并保持良好的推理速度;Lu 等^[6]提出一种新结构,有效地增强影像中相似影像块的特征表达能力和长期依赖性,以较低的计算成本获得更好的性能;Choi 等^[7]发现相邻的局部窗口滑动具有窗口自注意力然后使其交互,扩大了恢复退化像素的区域,并采用分层编码器的多尺度输出,在保持高效结构的同时,实现了具有竞争力的性能。

有少数专家也开展了面向道路场景的车牌影像重建技术的研究,相较于其他领域,道路场景采集到的低质车牌影

像存在角度偏转的现象,且存在强噪和模糊的不确定性、过曝或者欠曝导致车牌字符不完整的研究难点。吴玮等^[8]提出基于核偏最小二乘回归的超分辨率网络,将影像块以此重建,然后拼接为高分辨率影像,这种方法相较于插值提升了车牌影像的重建效果;徐胜军等^[9]提出一种编码解码结构的车牌影像超分辨率网络,并且引入常用于语音和文本识别的损失函数,通过这种方式提升了低质车牌影像的重建效果和处理车牌关键信息缺失问题的鲁棒性;Lin 等^[10]提出基于生成对抗网络的车牌影像重建方法,该方法对输入影像进行预处理,利用残差密集网络提取影像特征并引入渐进式采样,使重建后的车牌影像质量和细节更高;Shal 等^[11]提出由生成对抗神经网络提供支持的车牌影像分析,通过有效改变 SRGAN(photo-realistic single image Super-Resolution using a Generative Adversarial Network)层数,引入新的正则化函数,并将字符分割和识别与超分辨率重建网络相结合,提高了低质车牌信息识别的效率。

综上所述,现有超分辨率重建研究多集中在开放世界领域,针对低质车牌信息超分辨率重建的相关研究在重建效果上取得了一定进展,但是还处于起步阶段,现有研究多是集中在解决某一问题或某些问题,在面对多种影响因素叠加的情况下部分网络重建效果并不理想,还需要进一步优化完善。因此,本文提出了基于多维空间卷积信息增强的低质车牌信息超分辨率网络(Super-Resolution reconstruction for Low-quality License plate information based on multi-dimensional spatial convolutional information enhancement, LL-SR)。

1 整体网络架构

本文所设计的 LL-SR 如图 1 所示,采用端到端的非线性映射模块作为主体,首先使用卷积对输入低质车牌影像特征升维,挖掘低质车牌影像的浅层特征信息,再输入多个串联的像素矫正模块(Pixel Calibration Module, PCM)进行深层特征提取,获得更丰富的语义信息,减轻由于强噪、模糊、过曝、欠曝带来的影响,最后通过亚像素卷积上采样^[12]对深层特征进行车牌信息重建。

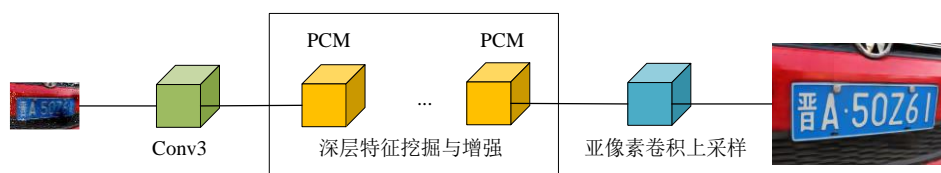


图1 LL-SR 结构示意图

Fig. 1 Structure diagram of LL-SR

2 PCM 设计

本文所提 LL-SR 中核心模块 PCM 的整体结构如图 2 所示, 该模块采用残差连接^[13], 将均噪模块(Average Noise Module, ANM), 像素信息增强模块(Pixel Information Enhancement Module, PIEM)、字符矫正(Character Correction Module, CCM)串联, 分别从相邻特征关联关系、不同感受野和不同维度、像素语义前后关联分析等角度对低质车牌信息进行特征提取和融合, 通过挖掘车牌影像底层信息局部特征之间的关联性, 提升低质车牌影像的高频细节, 从而减小在低质车牌信息超分辨率重建中, 由于强噪、模糊的不确定性和曝光造成车牌影像信息不完整所产生的影响。

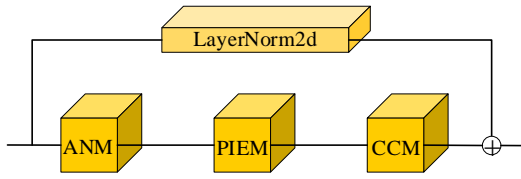


图2 PCM 结构示意图

Fig. 2 Structure diagram of PCM

2.1 ANM 设计

道路场景下采集到的低质车牌影像受不同环境影响往往伴有强噪问题, 使关键细节难以分辨, 影响车牌影像质量。为降低强噪影响, 本文设计了 ANM(如图 3 所示), 通过将强噪影响均分到具有相邻关系的特征点, 再减少强噪特征点在特征矩阵的占比, 从而减小不确定性强噪的影响。

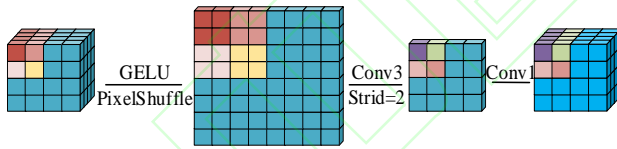


图3 ANM 结构示意图

Fig. 3 Structure diagram of ANM

ANM通过分析特征点在经过亚像素卷积上采样时的运算过程(如图 4 所示), 对于某一个特征点, 跨通道邻居(Cross Channel Neighbors, CCN)和空间邻居(Spatial Neighbors, SN)在经过亚像素卷积上采样生成 HR 车牌影像后, 在不同维度的 CCN 和 SN 会处于同一纬度, 且仍然和原特征点保持空间上的邻近关系。因此, 可以采用卷积模块聚合 CCN 和 SN 特征点, 挖掘特征点在空间和通道上协同卷积计算的关联关系, 将强噪特征点的影响均分到通道和空间两个维度的邻居像素。再使用卷积进行通道升维, 减小强噪特征点占比和恢复通道到原有维度, 从而在该位置减小强噪所产生的影响。ANM 数学模型可表示为式 1, 其中 $C_{1 \times 1}$ 和 $C_{3 \times 3}$ 分别代表 1×1 和 3×3 卷积, P 代表亚像素卷积上采样, F_1 代表输入特征, F_2 代表输出特征。

$$F_2 = C_{1 \times 1}(C_{3 \times 3}(P(F_1))) \quad (1)$$

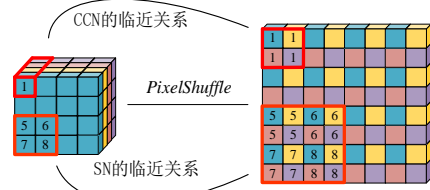


图4 CCN 和 SN 的空间临近关系

Fig. 4 Spatial proximity relationship of CCN and SN

ANM通过挖掘特征点在空间和通道上协同卷积计算的关联关系, 聚合 CCN 和 SN 特征点, 减小强噪特征点在特征矩阵占比, 从而减小强噪对车牌影像检测的影响。且由于作用的位置处于浅层特征, 此时并没有堆叠过多卷积, 感受野较小, 仍保留着车牌影像的细节信息, 可利用的细粒度特征也更多, 因此也可以起到聚拢车牌细节特征的作用, 使重建效果更好。

2.2 PIEM 设计

道路场景采集的车牌影像往往存在模糊的问题, 在一定程度上会丢失车牌影像的高频细节以及存在车牌像素的移位。为降低采集到的车牌信息受模糊问题的影响, 本文设计 PIEM 从不同感受野和不同维度挖掘特征图和特征图之间的关联关系, 提升车牌影像的高频细节改善和恢复车牌影像的清晰度。PIEM 结构如图 5 所示, 由空间像素信息增强模块(Spatial Pixel Information Enhancement Module, SPIEM)和通道像素信息增强模块(Channel Pixel Information Enhancement Module, CPIEM)构成, r_1 和 r_2 代表不同感受野大小(设置 $r_1 > r_2$), 通过具有小感受野 r_2 的 SPIEM 补充学习模糊字符的细节, 并且希望 CPIEM 在较大感受野 r_1 上去增强通道的关联性。该模块采用具有不同感受野的 SPIEM 从多尺度^[14]学习车牌字符特征, 并最终通过 CPIEM 加强空间和通道特征的关联性。

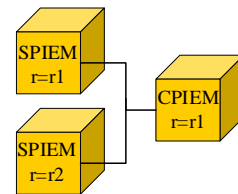


图5 PIEM 结构示意图

Fig. 5 Structure diagram of PIEM

SPIEM 和 CPIEM 的整体结构一样, 如图 6 所示, 输入特征 y 和输出特征 x 关系可以描述为:

$$y = f(x) + x \quad (2)$$

整体结构采用残差连接, 其中 $f(x)$ 函数展开为 LayerNorm2d 归一化计算跨通道的方差和具有反向传播的

通道权重动态参数以生成隐式编码全局上下文 *Attention* 的计算^[15-16], 具体可以描述为:

$$f(x) = \text{Attention}(\text{LayerNorm2d}(x)) \quad (3)$$

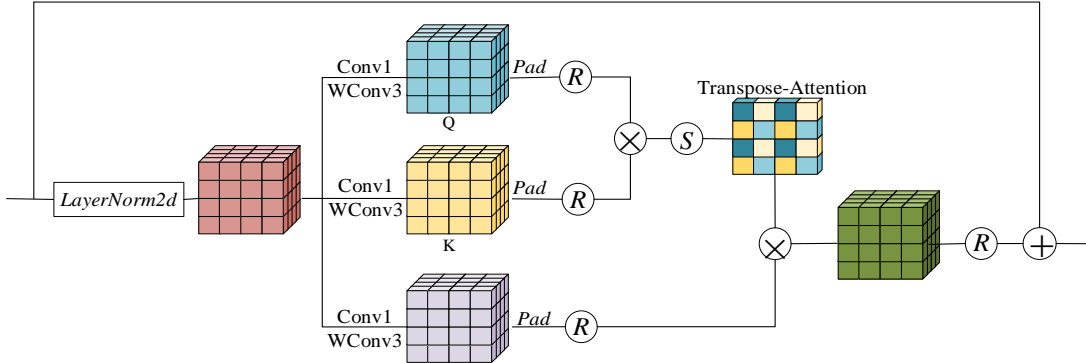


图6 SPIEM 和 CPIEM 整体结构示意图

Fig. 6 Schematic diagram of SPIEM and CPIEM overall structure

注意力模块对于输入的特征首先利用 1×1 卷积对特征图通道升维, 聚合像素级跨通道上下文, 然后用 3×3 逐通道卷积^[17]编码通道级空间上下文, 基于此生成查询 (Q)、索引 (K)、值 (V)。

$$(Q, K, V) = W_{3 \times 3}(C_{1 \times 1}(x)) \quad (4)$$

再对生成的 Q 、 K 、 V 进行 *Pad* 操作和 *R* 操作, 其中 *Pad* 操作是对特征图进行边缘填充使特征图满足 *R* 操作所需要的特征点个数, *R* 操作是利用 *Rearrange* 函数, 对张量进行重新排序:

$$(Q, K, V) = R(\text{Pad}(Q, K, V)) \quad (5)$$

通过不同 *R* 操作变化特征矩阵形状, 生成不同注意力图, 再进行不同注意力的计算:

$$\text{Attention}(Q, K, V) = V * \text{Softmax}(K * V / \delta) \quad (6)$$

SPIEM 和 CPIEM 对应着两种不同 *R* 操作, 这两种操作分别对应着进行空间和通道注意力的计算, 部分可视化过程分别如图 7~8 所示。对于特征矩阵 $M^{B \times C \times H \times W}$, 其中 B 、 C 、 H 、 W 分别为每批处理个数、通道、长、宽。对 M 进行填充使 $C = \text{head} \times d$; $H = a_1 \times h \times ph$; $W = a_2 \times w \times pw$, 其中 head 为多头个数, h 、 ph 、 w 、 pw 为实验所设置感受野大小 r 。

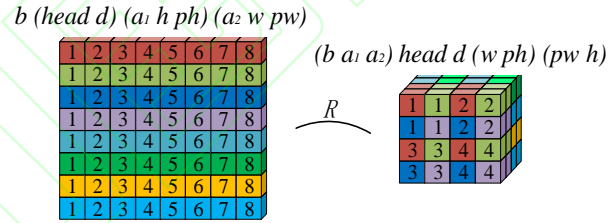


图7 空间注意力的 *R* 操作

Fig. 7 *R* operation of spatial attention

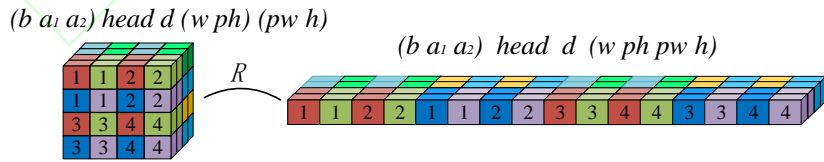


图8 通道注意力的 *R* 操作

Fig. 8 *R* operation of channel attention

空间注意力重新排列使某一通道内 $r \times r$ 的特征点作为一个整体, 改变为三维的立体块, 对于重新排列后的每个特征点获得了更大的感受野和更强的信息感知能力, 并且由于没有进行卷积和池化运算, 避免模糊车牌关键信息的丢失。空间注意力是对重新排列后的块与块之间做注意力的计算, 通道注意力对块的每个通道进行一维展开之后与通道进行注意力计算, 挖掘每个块的空间与通道之间的关联关系, 进而关注特征点在不同像素维度下的特征表达。

挖掘不同感受野空间上下文的 SPIEM 到通道上下文的 CPIEM, PIEM 从不同感受野和不同像素维度挖掘特征及特征间的关联关系, 使网络关注车牌信息在多维空间上的语义表达, 进而挖掘并增强模糊、曝光失常等不确定因素干扰下目标像素及像素之间的关联性, 丰富底层目标像素的特征表达。因此面对模糊的车牌影像, 网络可以学习关键像素和周围像素的关联性重建车牌信息, 从而恢复和提升模糊车牌影像的高频细节。

2.3 CCM 设计

当拍摄到的车牌影像受曝光影响会存在曝光失常的区域, 此时车牌部分字符会因为过曝或者欠曝使关键像素区

域被遮挡, 导致车牌信息不完整。为减小由于过曝或者欠曝对车牌信息重建产生的影响, 本文设计 CCM 学习字符的整体轮廓, 根据未受曝光影响区域的车牌字符信息去恢复被曝光影响区域的车牌字符信息。

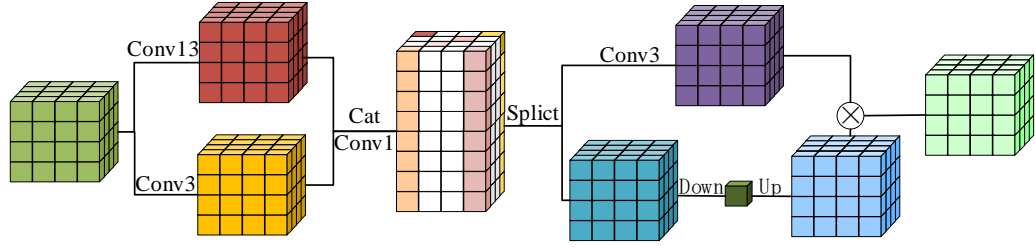


图9 CCM 结构示意图

Fig. 9 Schematic diagram of CCM structure

CCM 从不同尺度学习车牌信息的特征, 通过一个大卷积核 13×13 从大感受野学习车牌字符的整体轮廓, 减少由于曝光失常区域像素所占比重, 并行短接一个小卷积核 3×3 恢复车牌字符的细节^[18], 对于得到的特征, 本文并没有采用特征矩阵直接相乘或者相加的方式。为有效融合不同感受野所获取到的车牌字符特征, 对被曝光影响位置车牌字符信息更好的恢复, 本文对得到的特征先进行拼接和用 1×1 卷积核促进不同感受野特征的跨通道像素级融合。数学模型如式 7 所示, 其中 f_b 是经过大卷积核之后的特征, f_s 是经过小卷积核之后的特征, B 代表 *Batch Normalization* 归一化, $+$ 代表对通道的拼接操作, $C_{1 \times 1}$ 代表 1×1 的卷积运算。

$$f_c = C_{1 \times 1}(B(f_b) + B(f_s)) \quad (7)$$

将融合后的特征分离成两份相同的特征矩阵: 车牌信息矩阵 f_u 和车牌校准矩阵 f_v :

$$f_u, f_v = \text{Split}(f_c) \quad (8)$$

其中 f_u 为被校准的原始对象, 使用 $C_{3 \times 3}$ 卷积进行特征提取:

$$f_u = C_{3 \times 3}(f_u) \quad (9)$$

f_v 则采用卷积和池化下采样提取和精炼特征, 从而获得更丰富的语义信息和更大的感受野, 然后用卷积和插值上采样恢复到下采样之前的特征维度, 其中 *Down* 代表下采样, *Up* 代表上采样:

$$f_v = \text{Up}(\text{Down}(f_v)) \quad (10)$$

再使用 *Sigmoid* 激活函数获取特征校准权重校准 f_u , 得到最终特征 f_{out} :

$$f_{out} = f_u \times S(f_v) \quad (11)$$

由于该模块的整体处理仅考虑了特征点在空间尺度上的关联关系, 并没有考虑特征的上下文信息, 避免了曝光失常区域的特征在上下文的传播^[19], 因此可以减少由于曝光造成的干扰。

3 实验与结果分析

3.1 数据集

本文筛选自制数据集太原车牌(License plate of Tai yuan, LT)和公共数据集美国车牌(License plates of the United states of america, LU)作为车牌的 HR 影像, 通过做相同的退化处理生成 LR 车牌影像构成数据集对。为了提高网络的泛化能力, 本文对两个数据集进行了翻转、旋转、裁剪等数据增强处理, 处理后二者训练集个数分别为 18784 和 41952, LR 车牌影像大小均值为 83×167 和 124×184 。为了进一步提升本文网络的鲁棒能力和抗干扰能力, 选用双三次插值下采样^[20], 并考虑到影响低质车牌信息检测因素: 强噪、模糊、过曝、欠曝等, 对于下采样后的车牌影像, 随机选择添加运动模糊或高斯模糊, 并随机决定是否添加噪声, 从而完成 LR 车牌影像的生成。数学模型如式 12 所示, 其中 \downarrow 代表双三次下采样, R_c 为随机选取, \otimes 为点乘操作, k_1, k_2 为高斯模糊核和运动模糊核, p 代表添加干扰因素的概率, n_1, n_2, n_3 分别为高斯噪声, 椒盐噪声和改变曝光。

$$I_{LR} = I_{HR} \downarrow \otimes R_c(k_1, k_2) + \sum_{i=1}^3 p n_i \quad (12)$$

3.2 实验配置

实验平台搭载 NVIDIA RTX 3090, 显存 24GB, 深度学习框架为 Pytorch1.10.1, 编程语言为 Python3.8, 损失函数采用 L1 损失, 优化器采用 Adam, 输入网络的车牌影像大小裁剪为 64×64 , 每批次处理数为 32, 初始学习率为 5×10^{-4} , 实验设置训练总批次为 300, 学习率衰减为 0.5, 重建倍数为 4, 随机数种子 1234, 评价指标使用峰值信噪比(Peak Signal to Noise Ratio, PSNR)和结构相似性(Structural SIMilarity, SSIM)^[21]。

3.3 消融实验

为验证该文网络的有效性,在数据集 LT 和 LU 上开展实验,结果如表 1 所示。在 ANM 的基础上,添加了 PIEM 和 CCM 后,虽然参数分别提升了 369×10^3 和 406×10^3 ,但

是 PSNR 和 SSIM 提升 1.1552dB、0.0263、1.1122dB、0.05 和 1.6158dB、0.0363、1.1472dB、0.0433,提升效果明显。当采用最终网络时候,参数量为 934×10^3 ,网络性能最佳,为 26.6824dB、0.8203、22.3567dB、0.7813,实验结果验证了该文网络各个模块的有效性。

表1 消融实验

Tab.1 Ablation experiment

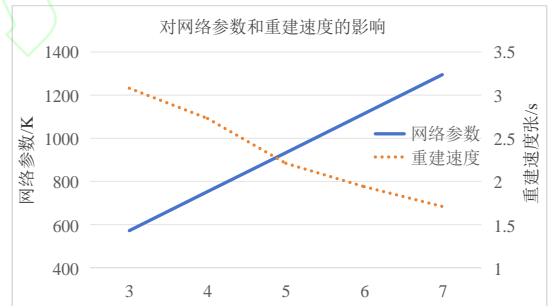
ANM	PIEM	CCM	参数量 / 10^3	LT		LU	
				PSNR/dB	SSIM	PSNR/dB	SSIM
√			159	24.7374	0.7779	20.6782	0.7116
√	√		528	25.8926	0.8042	21.7904	0.7616
√		√	565	26.3532	0.8142	21.8254	0.7549
√	√	√	934	26.6824	0.8203	22.3567	0.7813

本文对 PIEM 个数在数据集 LT 和 LU 上分别进行了验证,测试的个数为 3、4、5、6、7,测试指标为 PSNR 和 SSIM、网络参数以及重建速度,测试结果以自建数据集为例。如图 a 所示,PIEM 个数为 5 时,PSNR 和 SSIM 达到最高: 26.6824dB 和 0.8203,6 和 7 时略有下降,根据该文多次实验推测,大卷积核虽然可以有效提升感受野,但是,由于输入影像为 64×64 ,当单个特征点感受野超过车牌影像尺寸,再叠加过大卷积核会有可能造成重建影像的模糊,导致重

建车牌影像质量下降,重建效果对比如图 11 所示。参数个数和网络重建速度如图 b 所示,当 PIEM 个数为 5 时,此时参数为 934×10^3 ,对于分辨率均值为 83×167 的 LR 车牌影像,在 NVIDIA 3090 Ti 24GB 显卡上进行 4 倍超分辨率重建的速度为每秒 2.21 张。在数据集 LU 的测试结果与数据集 LT 类似,综合考虑在两个数据集上的测试结果,选取网络的 PIEM 个数为 5。



(a)对 PSNR 和 SSIM 的影响



(b)对网络参数和重建速度的影响

图10 PIEM 个数对网络的影响

Fig. 10 Influence of PIEM number on network



(a) 原图

(b) PIEM 个数为 5

(c) PIEM 个数为 6

(d) PIEM 个数为 7

图11 不同 PIEM 个数可视化对比

Fig. 11 Visual comparison of different PIEM number

3.4 客观评价

为了验证本文网络的泛化性和鲁棒性,本文在数据集 LT 和 LU 对比了 2018 年到 2023 年具有代表性的网络,实验数据如表 2 所示。实验结果表明,该文网络在参数和推理

速度上并不具备优势,但在输入为 64×64 大小的低质影像下,浮点运算量减少了 0.16×10^6 、PSNR 和 SSIM 提高了 0.8339dB、0.0179 和 0.9183dB、0.0344,无论是在数据集 LT 还是在 LU 上,本文网络在重建的客观指标上均有较大幅度的提升,验证了本文网络具有一定优势。

表2 不同网络在数据集 LT 和 LU 上的量化比较

Tab. 2 Quantitative comparison of different networks on dataset TL and LU

网络	参数量/ 10^3	LT		LU		浮点运算量 /GFLOPs	推理速度 /(张·秒 ⁻¹)
		PSNR/dB	SSIM	PSNR/dB	SSIM		
CARN ^[22]	1592	24.9463	0.7873	20.9375	0.7414	6.463	4.73
IMDN ^[23]	715	25.7147	0.8008	21.3724	0.7454	2.909	4.12
MAFFSRN ^[24]	830	25.9104	0.8009	20.9035	0.7251	3.163	1.81
SwinIR ^[4]	897	26.4662	0.8145	21.9975	0.7636	3.769	2.12
RLFN ^[5]	543	25.4384	0.7962	21.4635	0.7449	2.120	4.30
ESRT ^[6]	751	25.9914	0.8024	21.5102	0.7457	4.169	1.99
NGramSwin ^[7]	1201	26.4715	0.8146	21.8839	0.7617	3.826	0.94
LL-SR	934	26.6824	0.8203	22.3567	0.7813	3.614	2.21

3.5 可视化效果

为了验证该文网络重建低质车牌信息的可视化效果，该文针对选取了多张车牌影像进行可视化效果对比。如图 11 所示，图 a 和图 b 为 LT 数据集下的存在角度偏转、强噪以及曝光失常干扰因素的低质车牌影像，图 c、图 d、图 e 和图 f 为 LU 数据集下存在模糊和曝光失常的低质车牌影像。其中最左图为选取的低质车牌影像，框出部分和 LR 标签为截取展示对比效果的部分，HR 标签是高分辨率车牌影像，其他则为对比的网络重建结果。可以看到在 LR 影像中，

字符由于存在的问题而难以辨别，经过对比的网络重建之后，虽然可以比较好的恢复车牌背景，但是车牌字符信息存在字符轮廓粘连，字符缺失、伪影等问题，仍然难以准确识别车牌信息。本文网络重建后，存在的强噪、模糊、过曝、欠曝等问题有了明显的改善，重建后的车牌信息字符轮廓明显，基本可以辨别截取出的每个字符，效果要优于对比的网络，且本文网络对角度偏转下车牌影像的信息重建也具有较好的可视化效果，综合验证了本文网络对于低质车牌信息重建的有效性、鲁棒性和泛化性。





图12 低质车牌影像重建可视化效果对比

Fig. 12 Comparison of visual effects of low-quality license plate image reconstruction

3.6 真实场景车牌重建

为验证本文网络在真实场景下的重建效果，本文对真实拍摄的低质车牌影像进行重建，并与真实拍摄车牌进行对比，如表 3 所示，重建后车牌整体信息基本清晰准确，可视化效果较好，便于辨认识别，验证了本文网络的有效性。

表3 真实场景车牌重建效果对比

Tab. 3 Real scene license plate reconstruction

effect comparison

真实拍摄车牌	重建后车牌	真实车牌号
		晋 AJ2Z60
		晋 AD66454
		晋 AF11219
		晋 AD01997

4 结语

本文针对道路交通场景中低质车牌存在的强噪、模糊以及受曝光影响的问题，提出了基于多维空间卷积信息增强的低质车牌信息超分辨率网络，结合应用问题，通过网络及功能模块的综合设计，从相邻特征关联关系、不同感受野和不同维度、像素语义前后关联分析等维度空间进行车牌信息底层特征挖掘与增强，降低低质车牌影像中不确定性干扰因素的影响。实验结果表明，本文网络参数为 934×10^3 ，对分辨率均值为 83×167 的 LR 车牌影像，在 3090Ti24G 显卡上进行 4 倍超分辨率重建的速度为每秒 2.21 张，相较于对比的网络重建数据集 LT 和 LU 的 PSNR 和 SSIM 均值，本文网络提升了 0.8339dB、0.0179 和 0.9183dB、0.0344，在不同低质车牌影像重建可视化分析中可以看出本文网络具有更好的重建效果和较好的重建性能竞争力，综合验证了网络的有效性、鲁棒性和泛化性。当然，本文网络虽然取得了较好的效果，但在关键超参数的选取上依然存在主观性和随机性，下一步将围绕超分辨率网络自优化问题开展研究。

参考文献

- [1] 刘小宇,陈怀新,刘壁源,等.自适应置信度阈值的非限制场景车牌检测算法[J].计算机应用,2023,43(1):67-73.(LIU X Y, CHENG H X, LIU B Y, et al. License plate detection algorithm in unrestricted scenes based on adaptive confidence threshold[J]. Journal of Computer Applications, 2023, 43(1): 67-73.)
- [2] 张芳,赵东旭,肖志涛,等.单幅图像超分辨率重建技术研究进展[J].自动化学报,2022,48(11):2634-2654. (ZHANG F, ZHAO D X, XIAO Z T, et al. Research progress of single image super-resolution reconstruction technology[J]. ACTA AUTOMATICA SINICA, 2022, 48(11): 2634-2654.)
- [3] 缪永伟,张新杰,任瀚实,等.基于通道多尺度融合的场景深度图超分辨率网络[J].计算机辅助设计与图形学学报,2023,35(1):37-47. (LIAO Y W, ZHANG X J, REN H S, et al. A channel multi-scale fusion network for scene depth map super-resolution[J]. Journal of Computer-Aided Design & Computer Graphics, 2023, 35(1): 37-47.)
- [4] LIANG J, CAO J, SUN G, et al. SwinIR: image restoration using swin transformer[C]. ICCVW 2021: in 2021 IEEE/CVF International Conference on Computer Vision Workshops, Montreal, BC, Canada, 2021: 1833-1844.
- [5] KONG F Y, LI M X, LIU S W, et al. Residual local feature network for efficient super-resolution[C]. CVPRW 2022: in 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, New Orleans, LA, USA, 2022: 1061-1101.
- [6] LU Z S, LI J C, LIU H, et al. Transformer for single image super-resolution[C]. CVPRW 2022: in 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, New Orleans, LA, USA, 2022: 456-465.
- [7] CHOI H, LEE J, YANG J, et al. N-gram in swin transformers for efficient lightweight image super-resolution[C]. CVPR 2023: in 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, BC, Canada, 2023: 2071-2081.
- [8] 吴炜,杨晓敏,余艳梅,等.核偏最小二乘算法的图像超分辨率算法[J].电子科技大学学报,2011,40(1):105-110.(WU W, YANG X M, XU Y M, et al. Image super-resolution using KPLS[J]. Journal of University of Electronic Science and Technology of China, 2011, 40(1): 105-110.)
- [9] 徐胜军,邓博文,史亚,等.一种编解码结构的车牌图像超分辨率网络[J].西安交通大学学报,2022,56(10):101-110.(XU S J, DENG W B, SHI Y, et al. An encoder-decoder-based super resolution network for license plate images[J]. JOURNAL OF XI'AN JIAOTONG UNIVERSITY, 2022, 56(10): 101-110.)
- [10] LIN M F, LIU L X, LIU J C, et al. License plate image reconstruction based on generative adversarial networks[J]. Remote Sensing, 2022, 13(15): 3018-3019.
- [11] SHAL I H, FAHMY O M, ELATTAT M A, et al. License plate image analysis empowered by generative adversarial neural networks[J]. Network Daily News, 2022, 10(8): 30846-30857.
- [12] SHI W Z, CABALLERO J, HUSZ F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network[C]. in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016: 1874-1883.
- [13] HE K M, ZHANG X, REN S Q, et al. Deep residual learning for image recognition[C]. CVPR 2016: in 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 2016: 770-778.
- [14] MIN K, LEE G H, LEE S W. Attentional feature pyramid network for small object detection[J]. Neural Networks, 2022, 155(10): 439-450.
- [15] SYED W Z, AITYA A, SALMAN K, et al. Restormer: efficient transformer for high-resolution image restoration[C]. CVPR 2022: in 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 2022: 5718-5729.
- [16] WANG H, CHEN X H, NI B B, et al. Omni aggregation networks for lightweight image super-resolution[C]. CVPR 2023: in 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, BC, Canada, 2023: 22378-22387.
- [17] CHOLLET F. Xception: Deep learning with depthwise separable convolutions[C]. CVPR 2017: in 2017 IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 2017: 1800-1807.
- [18] DING X H, ZHANG X Y, ZHOU Y Z, et al. Scaling up your kernels to 31x31: revisiting large kernel design in cnns[C]. CVPR 2022: in 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 2022: 11953-11965.
- [19] LIU J J, HOU Q B, CHENG M M, et al. Improving convolutional networks with self-calibrated convolutions[C]. CVPR 2020: in 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Seattle, WA, USA, 2020: 10093-10102.
- [20] 李键红,吴亚榕,詹瑾.挖掘理想重建图像自相似性的超分辨率[J].湖南大学学报(自然科学版),2021,48(8):149-160.(LI J H, WU Y R, ZHAN J. Image super-resolution by exploiting self-similarity of ideal reconstruction[J]. Journal of Hunan University (Natural Sciences), 2021, 48(8): 149-160.)
- [21] 李颖,黄超,孙成栋等.真实复杂场景下基于残差收缩网络的单幅图像超分辨率方法[J].计算机应用,2023,43(12):3903-3910.(LI Y, HUANG C, SUN C D, et al. Single image super-resolution method based on residual shrinkage network in real complex scenes[J]. Journal of Computer Applications, 2023, 43(12): 3903-3910)
- [22] AHN N, KANG B K, SOHN S A. fast, accurate, and lightweight super-resolution with cascading residual network[OL]. [2018-3-23] <https://arxiv.org/abs/1803.08664>.
- [23] HUI Z, GAO X B, YANG Y C, et al. Lightweight image super-resolution with information multi-distillation network[OL]. [2019-9-26] <https://arxiv.org/pdf/1909.11856v1.pdf>.
- [24] ABDUL M, HWANG J W, Yang S B, et al. Multi-attention based ultra lightweight image super-resolution[OL]. [2020-9-21] <https://arxiv.org/abs/2008.12912>.

This work is partially supported by Supported by Fundamental Research Program of Shanxi Province (20210302123216), Shanxi Province Industry Education Integration Graduate Joint Training Demonstration Base Project (2022JD11), Shanxi Provincial Machinery Product Quality Judicial Appraisal Center Enterprise Entrusted Project (2023254, 2023069), Taiyuan University of Science and Technology Graduate Joint Training Demonstration Base Project (JD2022004).

ZHANG Rui, born in 1987, Associate Professor, PhD. His research interests include intelligent information processing.

HUI Yongke, born in 1998, born in Nanyang, Henan Province, master candidate. His research interests include computer vision.

ZHANG Yanjun, born in 1982, male, professor, PhD. His research interests include intelligent information processing.

PAN Lihu, born in 1974, Professor, PhD. His research interests include intelligent information processing.