面向智慧交通的双重特征融合图像曝光校正

杨佳钧[1,2)], 张金艺[1,2)*], 陈琪[3)]

[1)] (上海大学特种光纤与光接入网重点实验室 上海 200444)
[2)] (上海大学特种光纤与先进通信国际合作联合实验室 上海 200444)
[3)] (上海三思电子工程有限公司 上海 201100)
(zhangjinyi@shu.edu.cn)

摘 要: 在智慧交通领域中, 监控拍摄的车辆图像质量会受到场景光照条件的影响而造成关键信息损失, 导致后续交通执法缺乏可靠的图像依据, 而现有的图像曝光校正方法存在校正后图像颜色失真、细节模糊等问题. 为此, 提出一种面向智慧交通的双重特征融合图像曝光校正方法. 该方法通过全局子网与曝光子网提取相关特征, 全局子网引入特征注意力机制提取图像全局特征, 改善校正图像的整体亮度与颜色; 曝光子网基于U-Net结构, 引入多尺度注意力机制提取图像中不同尺度的曝光特征, 提升校正图像的细节纹理; 最后对双重特征进行融合, 将其映射至图像空间, 实现图像曝光校正. 在 MSEC 数据集上的实验结果表明, 所提方法的峰值信噪比达到 21.8015 dB, 结构相似度达到 0.867 7, 自然图像质量评价达到 11.711 8, 均优于对比的曝光校正方法, 能够有效地还原图像颜色信息与细节纹理, 证明该方法具有良好的曝光校正效果.

关键词: 图像曝光校正; 注意力机制; 双重特征; 智慧交通
中图法分类号: TP391.41 **DOI:** 10.3724/SP.J.1089.2023-00217

Image Exposure Correction Based on Dual Feature Fusion for Intelligent Transportation

Yang Jiajun[1,2)], Zhang Jinyi[1,2)*], and Chen Qi[3)]

[1)] (*Key Laboratory of Specialty Fiber Optics and Optical Access Networks, Shanghai University, Shanghai* 200444)
[2)] (*Joint International Research Laboratory of Specialty Fiber Optics and Advanced Communication, Shanghai University, Shanghai* 200444)
[3)] (*Shanghai Sansi Institute for System Integration, Shanghai* 201100)

Abstract: In the field of intelligent transportation, the quality of vehicle images captured by surveillance cameras can be affected by the lighting conditions of the scene, resulting in the loss of critical information, which in turn leads to a lack of reliable image evidence for subsequent traffic enforcement. However, existing image exposure correction methods have problems such as color distortion and blurry details in the corrected image. Therefore, a dual feature fusion image exposure correction method for intelligent transportation is proposed, which extracts relevant features through global subnet and exposure subnet, respectively. The global subnet introduces the feature attention mechanism to extract global features of the image, improving the overall brightness and color of the corrected image. The exposure subnet is based on the U-Net structure and introduces a multi-scale attention mechanism to extract exposure features at different scales in the image, improving the detailed texture of the corrected image. Finally, image exposure correction is achieved by fusing the dual features and mapping them to the image space. The experimental results on the MSEC dataset demonstrate that, compared to the existing exposure correction methods, the proposed method has a peak signal-to-noise ratio of

收稿日期: 2023-05-29; 修回日期: 2024-01-17. **基金项目**: 高等学校学科创新引智计划(111)项目(D20031); 上海市闵行区重大产业技术攻关计划(2022MH-ZD19). **杨佳钧**(1999—), 男, 硕士研究生, 主要研究方向为深度学习、图像曝光校正; **张金艺**(1965—), 男, 博士, 研究员, 论文通信作者, 主要研究方向为伴随性机器人、计算机视觉; **陈琪**(1980—), 男, 硕士, 工程师, 主要研究方向为智能图像识别、信号处理.

21.801 5 dB, a structural similarity index measure of 0.867 7, and a natural image quality evaluator of 11.711 8. The proposed method can effectively restore color information and detailed textures in images, proving its excellent exposure correction effect.

Key words: image exposure correction; attention mechanism; dual feature; intelligent transportation

随着图像拍摄设备的大规模部署，基于深度学习的监控图像处理对于智慧交通的发展与建设起着至关重要的作用. 然而在实际应用中，由于道路中场景信息与光照条件复杂多变，图像质量极易受到曝光错误的影响，造成车辆区域信息损失，影响对车牌区域、车内区域等原始信息的提取与分析. 因此，针对曝光错误的图像进行校正，还原车辆区域中应有的亮度、色彩、细节纹理等信息，对推动智慧交通相关技术的实际应用具有重要意义.

针对曝光错误导致图像质量下降的问题，早期，研究人员依赖图像直方图实现曝光校正[1-3]. 此类方法利用直方图中图像的统计信息，通过特定的变换操作对像素强度值进行调整，但是容易导致校正后图像局部区域信息丢失. 随着 Retinex 理论[4]在计算机视觉领域中的发展，部分图像曝光校正方法通过图像分解的方式预测反射图或照射图，最终得到校正结果[5-9]；但此类方法大多仅适用于欠曝图像校正[10]，并且容易造成图像偏色. 也有一些研究人员基于大气散射模型开展图像曝光校正[11]，将过曝区域视作正常曝光场景与非均匀雾度的叠加，但由于大气散射模型求解本身属于病态问题[12]，因此图像校正的结果易为次优解.

近年来，随着深度学习理论的发展，出现了大量通过神经网络实现欠曝图像或过曝图像校正的方法. 对于曝光不足的情况，Guo 等[13]将欠曝图像校正视作一种使用深度学习预测图像特定曲线的任务，使用无监督网络结构估计用于调整图像动态范围的像素级高阶曲线；Jiang 等[14]利用图像空间变换的光照信息和曝光错误产生的伪影校正欠曝图像；而对于曝光过度的情况，Cao 等[15]将场景信息与图像曝光信息解耦，旨在学习与曝光信息无关的图像信息，并使用恢复网络生成出更真实的校正结果：上述方法仅能单独针对曝光不足或曝光过度场景，无法有效地满足曝光不足与曝光过度均会出现的交通场景，难以恢复由曝光错误而丢失的信息. 为了解决图像曝光校正现有数据集在实际应用中的局限性，Afifi 等[10]提出包含欠曝图像与过曝图像的数据集 MSEC，首次通过深度学习的方法同时解决曝光不足和曝光过度问题，利用图像的拉普拉斯金字塔分解对输入图像逐层进行校正，在校正全局颜色信息的基础上逐步细化图像细节，使网络由粗到细地对图像进行颜色与细节上的增强，但是该方法容易产生模糊问题；Nsamp 等[16]基于特征解耦的思想，通过神经网络学习特征空间中的曝光不变的表征，增强网络输出图像的曝光一致性，但是使用该方法校正后的图像存在颜色失真；Cui 等[17]通过预测逆色调映射函数，将输入的 RGB 图像映射到其对应的 raw-RGB 空间中，通过预测简化的色彩矩阵和 Gamma 参数模拟图像信号处理流程，实现对图像曝光的校正，但是校正后的图像亮度存在偏差；Eyiokur 等[18]研究曝光错误对人像抠图任务的影响，通过端到端的图像曝光校正模型提升人像抠图的效果. 与一般图像的曝光校正相比，智慧交通场景下的车辆区域通常受到强光和低光等极端照明条件的影响，导致车牌区域、车内区域等交通监控中的重要信息丢失，对于该场景下曝光校正后的图像中的亮度、色彩、细节纹理等信息更应当符合真实场景，以支持智慧交通领域的各种应用.

综上所述，针对现有方法较难有效地解决智慧交通场景下曝光错误引起的问题，本文提出一种面向智慧交通的双重特征融合图像曝光校正方法. 通过引入特征注意力机制的全局子网提取图像全局特征，改善校正图像的整体亮度与颜色；在基于 U-Net 结构的曝光子网中引入多尺度注意力机制，提取图像中不同尺度的曝光特征，提升校正图像的细节纹理；最后融合全局子网与曝光子网提取到的双重特征，将其映射至图像空间，实现图像曝光校正.

1　本文方法

本文提出面向智慧交通的双重特征融合图像曝光校正方法，其结构如图 1 所示. 其中，全局子网引入特征注意力机制，从曝光有误的图像中提取全局特征；曝光子网通过引入多尺度注意力机

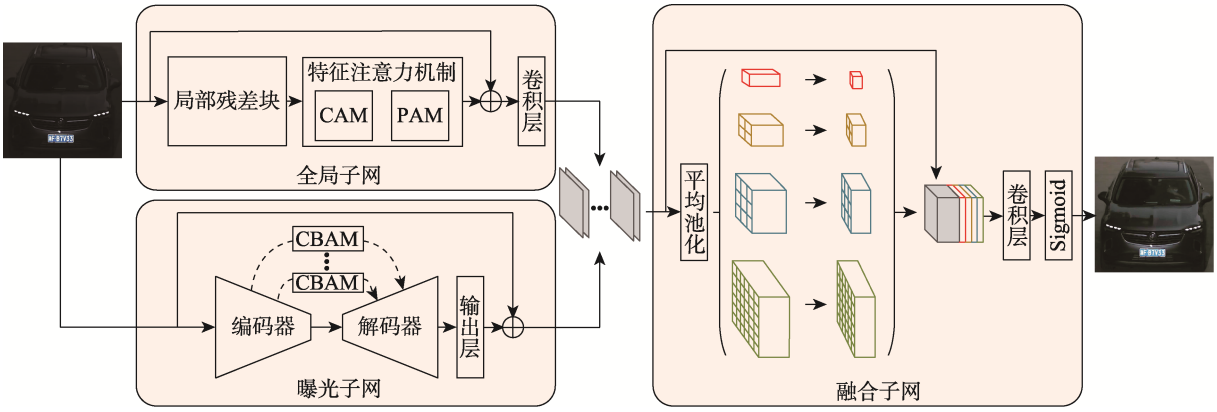


图 1 本文方法结构

制的 U-Net 结构实现曝光特征的提取，并提升图像纹理细节；最后通过基于金字塔池化结构的融合子网进行双重特征融合，并将其映射至图像空间，得到曝光校正后的输出结果.

1.1 特征注意力机制引导的全局特征提取

在智慧交通场景中，监控相机曝光错误往往使得捕获的车辆区域图像亮度偏暗或偏亮，导致图像颜色失真，因此需要对容易产生过曝的区域(如车牌)以及容易产生欠曝的区域(如车内)进行颜色与亮度上的校正，然而现有图像曝光校正方法较难还原车辆区域应有亮度与颜色信息. 在计算机视觉任务[19-20]中，注意力机制[21-22]能够有效地引导网络关注于感兴趣的区域，因此本文在全局子网中引入特征注意力机制，引导网络对输入图像的颜色分布、平均亮度等全局特征进行提取，提升图像曝光校正后的视觉效果. 全局子网的网络结构如图 2 所示.

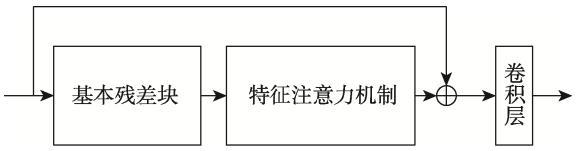


图 2 全局子网结构

全局子网首先通过局部残差块提取图像的全局空间信息，在局部残差块中使用卷积层和 ReLU 激活函数将输入图像由 RGB 图像空间映射至特征空间，再通过 3×3 的卷积层后得到全局空间信息特征图；然后对全局空间信息特征图进行特征注意力机制计算，特征注意力机制由通道注意力模块(channel attention module, CAM)和像素注意力模块(pixel attention module, PAM)这 2 部分组成，其结构图如图 3 所示，旨在关注曝光错误图像中的平均亮度与颜色分布.

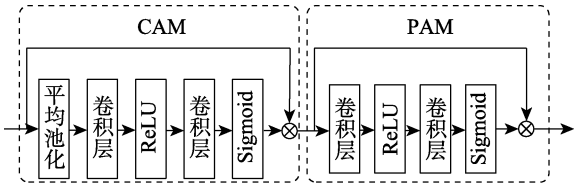


图 3 特征注意力机制结构

监控相机补光灯导致图像中车辆区域各部分存在不同程度的曝光错误，应当给予图像不同通道不同的权重信息. CAM 旨在自适应校正全局空间信息特征图中各个通道之间的权重，补偿图像中亮度较高的通道与亮度较低的通道. 该模块首先使用全局池化将全局空间信息特征图转化为通道描述符，图像尺寸由 $C\times H\times W$ 变为 $C\times 1\times 1$；随后通过 2 个卷积层、ReLU 激活函数和 Sigmoid 激活函数后得到不同通道的权重，并与输入空间信息特征图相乘，学习曝光错误的输入图像与曝光正确的参考图像通道映射之间的关系，使网络关注不同通道中的有效信息. 其计算过程表示为

$$I_{\mathrm{CAM}}=I\cdot\sigma\left(f\left(\gamma\left(f\left(\frac{1}{H\times W}\sum_{x}^{H}\sum_{y}^{W}I(x,y)\right)\right)\right)\right).$$

其中，I_{CAM} 表示 CAM 的输出结果；I 表示输入空间信息特征图；$I(x,y)$ 表示输入空间信息特征图在 (x,y) 处的值；H 和 W 分别表示输入空间信息特征图的高和宽；f 表示卷积操作；σ 表示 Sigmoid 激活函数；γ 表示 ReLU 激活函数.

传统的基于卷积神经网络(convolutional neural network, CNN)的图像曝光校正方法通常是平等处理空间信息特征. 然而由于在曝光错误的图像中其错误程度并非均匀分布，在对空间信息特征图进行卷积时，针对存在不同程度曝光错误的区域其权重应当有所差异. 因此，本文引入的 PAM

对不同空间区域的像素进行加权赋值，旨在自适应校正空间像素权重，以关注全局像素信息．PAM 的计算流程与 CAM 相似，在经过 2 个卷积层、ReLU 激活函数和 Sigmoid 激活函数后得到像素权重，最终与CAM的输出相乘得到PAM的输出结果

$$I_{\mathrm{PAM}} = I_{\mathrm{CAM}} \cdot \sigma\Big(f\big(\gamma\big(f\big(I_{\mathrm{CAM}}\big)\big)\big)\Big).$$

为了有效地防止训练网络时梯度的消失，使用残差连接的方式，将特征注意力机制加权的全局空间信息特征图与输入连接，再经过卷积层后得到全局子网的输出特征 G_{out}．在道路监控场景中，由于补光灯的作用，车辆区域易受到分布不均匀的光照影响，导致曝光不足与曝光过度的现象共存，不仅使得监控拍摄图像整体亮度失真，还会造成图像中不同区域存在不同程度的颜色失真．引入特征注意力机制，能够有效地捕获到输入的空间信息特征图中的亮度信息与颜色分布，有助于更准确地还原出道路监控图像中的真实信息．

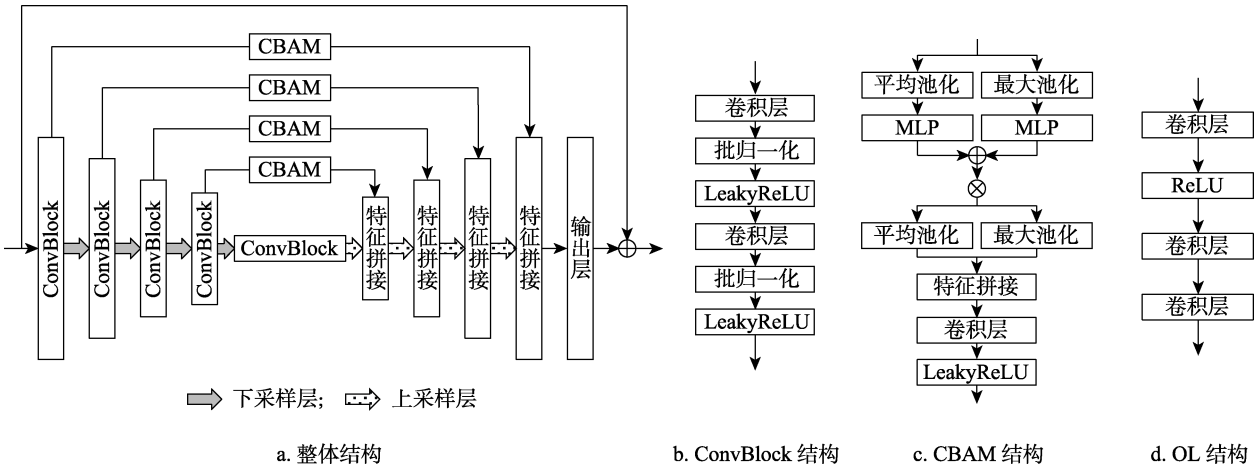
1.2　基于 U-Net 结构的曝光特征提取

在智慧交通场景中，光照条件复杂多变，往往在一幅监控图像的车辆区域中出现部分区域曝光不足，而部分区域曝光过度的情况．为了更准确地捕获图像中不同区域的曝光特征，将能够有效地提取图像中层次化上下文特征的 U-Net 结构作为曝光子网的基础，在此基础上引入多尺度注意力机制．曝光子网由编码器、解码器和输出层(output layer, OL)这 3 个部分构成，整体结构如图 4a 所示．编码器中，每层的输出分别对应不同尺度的曝光特征，在每层中通过 ConvBlock 进行特征提取，其结构如图 4b 所示，并使用下采样层扩大感受野，供下一层提取更深层次的曝光特征．其中，ConvBlock 由连续的卷积层、批归一化和激活函数构成，激活函数采用带泄露的线性修正单元 Leaky ReLU，提高神经网络各层之间的非线性关系，有效地缓解网络训练过程中梯度消失的现象．解码器中，每层对上一层的输出特征进行上采样，通过卷积实现降维，将其与编码器中对应层的曝光特征进行拼接．在编码器与解码器之间特征拼接的过程中，本文引入卷积块注意力模块(convolution block attention module, CBAM)，其结构如图 4c 所示；最终，通过由卷积层和 ReLU 激活函数组成的 OL 实现曝光特征 E_{out} 的提取，如图 4d 所示．

考虑编码器在下采样过程中容易造成空间细节信息的丢失，传统的U-Net结构中将编码器不同层级的特征通过跳跃连接的方式，实现浅层特征与深层特征之间的融合，旨在补充输入解码器中的细节特征．但是，针对场景信息复杂多样的交通监控车辆区域，通过跳跃连接的方式过于简单，无法有效地对图像中的曝光特征进行提取．为了解决这一问题，本文在特征拼接的过程中引入CBAM，首先将输入的特征图分别进行平均值池化和最大值池化；然后将经过2个共享权重的多层感知机(multilayer perceptron, MLP)进行升维，并与降维后的特征图进行相加，得到通道特征权重 ρ_1．对于空间特征权重，首先将通道特征加权后的输入特征图作为输入，经过平均池化(average pooling, AP)和最大池化(max pooling, MP)后进行特征拼接；然后使用卷积将其特征维度降为 1，经过 Sigmoid 激活函数后，得到空间特征权重 ρ_2．CBAM 的计算过程可以表示为 $F_{\mathrm{out}} = X_{\mathrm{en}} \cdot \rho_1 \cdot \rho_2$．其中，

$$\rho_1 = F \cdot \Big(\mathrm{MLP}\big(\mathrm{AP}\big(F\big)\big) + \mathrm{MLP}\big(\mathrm{MP}\big(F\big)\big)\Big);$$

$$\rho_2 = \sigma\Big(f\Big(\big\{\mathrm{AP}\big(F \cdot \rho_1\big); \mathrm{MP}\big(F \cdot \rho_1\big)\big\}\Big)\Big);$$



a. 整体结构　　b. ConvBlock 结构　　c. CBAM 结构　　d. OL 结构

图 4　曝光子网结构

F_{out} 和 X_{en} 分别表示 CBAM 的输出特征和输入特征. 最终, 曝光子网输出的曝光特征可表示为

$$E_{\text{out}}=\text{OL}\left(X_{\text{de}}^{1}\right)+I_{\text{in}}.$$

其中, I_{in} 表示输入图像. 第 i 层解码器表示为

$$X_{\text{de}}^{i}=\begin{cases}\left\{F_{\text{out}}^{i},u_i\left(\text{CB}\left(d_i\left(X_{\text{en}}^{i}\right)\right)\right)\right\}, & i=4\\ \left\{F_{\text{out}}^{i},u_i\left(X_{\text{de}}^{i+1}\right)\right\}, & i=1,2,3\end{cases}.$$

其中, F_{out}^{i} 和 X_{en}^{i} 分别为第 i 层编码器、CBAM 的输出; CB 表示 ConvBlock; d_i 表示第 i 层编码器的下采样层, 由卷积层、批量归一化和 Leaky ReLU 激活函数构成; u_i 表示第 i 层解码器的上采样层, 由双线性插值与卷积层构成. 解码器的输出结果经过 OL 后与输入残差连接, 得到曝光子网提取的曝光特征 E_{out}, 其可视化结果如图 5 所示. 其中, 蓝色区域为图像中的欠曝部分, 橙黄色区域为图像中的过曝部分, 白色区域为图像中曝光良好的部分.



图 5 曝光子网提取的曝光特征可视化

在基于 U-Net 结构的曝光特征提取中引入 CBAM, 可以充分地利用多尺度特征增强相关的特征响应. 一方面, 充分地保留智慧交通场景中图像的原始特征信息, 提升浅层特征图中的空间细节信息, 解决下采样过程中的细节丢失; 另一方面, 使得深层特征图中能够关注智慧交通场景的图像中过曝区域和欠曝区域的信息, 实现图像曝光校正.

1.3 融合双重特征实现图像曝光校正

在智慧交通场景下, 由于车辆不同区域可能受到不同程度曝光错误的影响, 因此车辆图像曝光校正往往难以兼顾整体信息与局部信息的还原. 本文将图像整体亮度、颜色特征的提取与图像曝光特征的提取视为 2 个独立的任务, 需要充分地融合全局子网与曝光子网输出的特征, 得到最终的图像曝光校正结果. 鉴于金字塔池化结构可以有效地将不同的特征嵌入到最终的结果中[23], 本文通过基于金字塔池化的融合子网对提取到的双重特征进行自适应加权融合, 充分地利用双重特征中的上下文信息, 生成具有不同大小感受野的特征图以在不同尺度上重建图像, 融合子网的结构如图 6 所示.

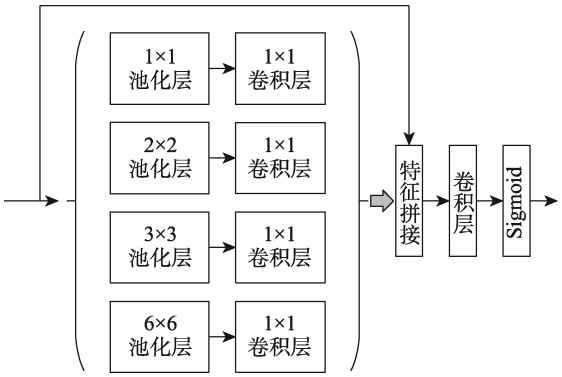


图 6 融合子网的结构

首先将来自全局子网与曝光子网的双重特征拼接得到 $\{E_{\text{out}},G_{\text{out}}\}$; 随后分别对其进行尺寸为 1×1, 2×2, 3×3 和 6×6 的自适应平均池化操作, 得到多个尺度的特征图 K_i, 建立一个多尺度的金字塔池化结构. 金字塔池化结构的每层中, 通过 1×1 卷积层对特征图 K_i 进行重建, 其中, 1×1 卷积层隐含了自适应加权的通道注意力机制, 引导网络恢复场景信息; 之后, 将金字塔每层重建后的特征图上采样至与 $\{E_{\text{out}},G_{\text{out}}\}$ 的空间尺寸相同, 得到金字塔每层的输出特征图 P_i, 并将其与 $\{E_{\text{out}},G_{\text{out}}\}$ 在通道维度上相连接, 实现全局信息与局部信息的融合. 对连接后的特征图依次经过 3×3 卷积层和 Sigmoid 激活函数, 得到最终的曝光校正后的输出图像 $R_{\text{out}}=\sigma\left(f\left(\left\{\left\{E_{\text{out}},G_{\text{out}}\right\},P_1,P_2,P_3,P_4\right\}\right)\right)$. 其中, K_i 表示金字塔池化结构中第 i 层金字塔的特征图; up 表示上采样操作; P_i 表示金字塔每层的输出特征图, 即 $P_i=\text{up}_i\left(f\left(K_i\right)\right),i=1,2,3,4$.

融合子网通过全局平均池化构造具有不同大小感受野的特征图, 以重建曝光校正后的图像, 保证全局信息与局部信息的有效融合, 尽可能地保留并增强图像中原有的细节纹理, 并还原场景中应有的颜色分布与亮度信息.

1.4 损失函数

在智慧交通场景中, 曝光错误的车辆区域图像信息通常复杂多变, 使用回归任务中常用的 L_1 或 L_2 损失难以引导网络恢复出真实场景. 为此, 本文设计一种联合损失函数, 其由内容损失 L_{c} 和感知损失 L_{p} 构成. L_{c} 的目的是在图像空间中最小

化曝光校正后图像与参考图像之间的差异，定义为 $L_{\mathrm{c}}=\sum_{i=1}^{N}\left\|y_i-x_i\right\|_1$．其中，$x_i$ 表示曝光校正后图像的像素；y_i 表示参考图像的图像像素：N 表示图像中像素的个数．L_{p} 基于预训练的 VGGNet，提取出其前 i 层 φ_i，并将 φ_i 的输出视作感知特征．与 L_{c} 相比，L_{p} 旨在缩小曝光校正后图像与参考图像之间的感知特征差异，本文使用均方误差(mean square error, MSE)损失函数优化感知特征差异．L_{p} 定义为 $L_{\mathrm{p}}=\frac{1}{C_iH_iW_i}\left\|\varphi_i(y)-\varphi_i(x)\right\|_2$．其中，$x$ 和 y 分别表示曝光校正后的图像和参考图像；C 为图像的通道数，H 和 W 分别表示图像的高和宽．最终，面向智慧交通的双重特征融合图像曝光校正网络的损失函数为 $L_{\mathrm{total}}=\lambda_1L_{\mathrm{p}}+\lambda_2L_{\mathrm{c}}$．其中，$\lambda_1$ 和 λ_2 分别表示为平衡 L_{p} 和 L_{c} 的权重，本文设定 λ_1=0.1，λ_2=1.0.

2　实验结果与分析

本文使用在图像曝光校正的研究中有较高学术价值的 MSEC 数据集进行训练与测试，以体现本文方法在非智慧交通场景下的通用性．通过对实际道路监控拍摄的车辆图像进行实验，验证本文方法在智慧交通场景中的有效性．使用峰值信噪比(peak signal-to-noise ratio, PSNR)、结构相似度(structural similarity index measure, SSIM)和自然图像质量评价(natural image quality evaluator, NIQE)这 3 个指标，对本文方法与现有主流图像曝光校正方法 MSEC[10]，ECCM[16]，IAT[17]和 ECMEIQ[18]进行实验．此外，本文将对不同方法的校正结果进行定性分析，比较其主观视觉效果．

本文实验的硬件环境为 AMD 5600X CPU，32 GB 内存，NVIDIA RTX 3070Ti 显卡；软件环境为 CUDA 11.5, PyTroch 1.11.0, Python 3.10.0．训练参数设置如下：为便于网络进行训练，本文将数据集图像尺寸调整为 256×256 像素；使用 Adam 优化器进行模型优化，初始学习率设置为 0.001，在训练过程中线性下降为 0, batchsize 设置为 4，共训练 100 个 epoch.

2.1　定量对比实验

将本文方法与现有主流图像曝光校正方法进行定量对比，结果如表 1 所示．测试模型均来自相应文献中作者公开的源代码，使用 MSEC 数据集中的测试集进行测试，该数据集由 MIT Adobe FiveK 数据集渲染而成；对于每个原始 RGB 图像，使用 5 种不同的曝光参数渲染以模拟真实的曝光误差，即通过调整相对曝光值分别为–1.5，–1，+0，+1 和+1.5，分别渲染具有曝光不足和曝光过度的图像；其中，训练集包含 17 675 幅图像，验证集包含 750 幅图像，测试集包含 5 905 幅图像．实验中，将完整测试集划分为曝光不足部分和曝光过度部分分别进行对比实验；为了验证本文方法的有效性，还从相对曝光值为+0 的数据中筛选出部分同时存在曝光过度与曝光不足的图像进行测试．为了便于不同方法之间的公平比较，将图像尺寸按比例缩放至最大边为 512 像素．

从表 1 可以看出，在完整测试集和曝光不足部分中，本文方法的 PSNR 分别达到 21.801 5 dB 和 21.676 5 dB，SSIM 分别达到了 0.867 7 和 0.860 7，均超越主流方法；在曝光过度部分中，本文方法的 PSNR 达到 21.884 9 dB，略低于 ECCM 方法[16]，但 SSIM 达到 0.872 3，超越了其他方法；对于 NIQE 指标，本文方法在所有场景中均取得最佳结果，体现出该方法的优越性．

表 1　与主流图像曝光校正方法定量结果对比

方法	曝光不足部分			曝光过度部分			曝光过度与曝光不足共存部分			完整测试集		
	PSNR/dB	SSIM	NIQE	PSNR/dB	SSIM	NIQE	PSNR/dB	SSIM	NIQE	PSNR/dB	SSIM	NIQE
MSEC w/o L_{adv} [10]	20.386 2	0.832 8	12.155 1	20.081 6	0.842 2	11.494 5	20.296 4	0.846 4	11.741 9	20.203 4	0.838 4	11.759 6
MSEC w/ L_{adv} [10]	20.068 7	0.835 8	12.363 8	20.030 5	0.840 9	11.697 5	20.248 2	0.844 9	11.931 6	20.045 8	0.838 8	11.964 1
IAT[17]	18.182 9	0.759 8	16.121 1	19.506 5	0.798 0	15.922 5	19.127 6	0.788 2	16.151 5	18.977 0	0.782 7	16.001 9
ECCM[16]	21.212 9	0.843 5	13.386 2	**21.952 5**	0.870 7	12.909 9	21.726 1	0.867 6	13.072 0	21.656 6	0.859 8	13.100 4
ECMEIQ[18]	20.724 7	0.833 0	12.365 0	21.003 5	0.845 8	11.966 0	20.945 6	0.844 1	12.018 4	20.892 0	0.840 7	12.125 6
本文	**21.676 5**	**0.860 7**	**12.103 4**	21.884 9	**0.872 3**	**11.450 8**	**21.758 9**	**0.871 8**	**11.624 9**	**21.801 5**	**0.867 7**	**11.711 8**

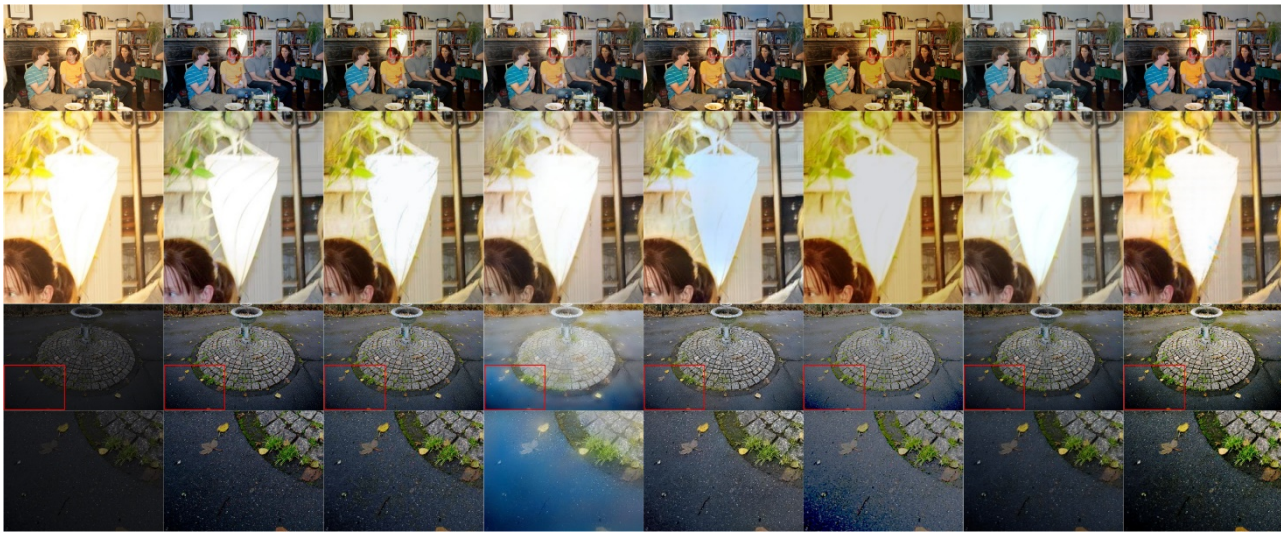
注．加粗数据表示最优指标.

2.2 定性对比实验

为了体现本文方法的泛化能力，分别对非智慧交通场景图像与智慧交通场景中的车辆图像进行定性对比实验，以验证该方法在真实交通场景中的有效性.

2.2.1 非智慧交通场景定性对比实验

在非智慧交通场景的定性对比实验中，使用 MSEC 数据集进行测试，验证本文方法的通用性. 6 种方法的主观视觉效果对比如图 7 所示. 从整体图像视觉效果和局部细节纹理方面与参考图像之间的差异程度来看，本文提出的双重特征融合图像曝光校正方法明显优于其他方法. 可以看出，针对曝光不足图像的处理过程中，MSEC w/o L_{adv} 方法[10]产生了严重的颜色失真；引入对抗损失函数后，MSEC w/ L_{adv} 方法[10]在一定程度上解决了整体图像颜色失真的问题，但是在局部区域的色彩还原度仍然表现不佳；在过度曝光的场景中，IAT 方法[17]校正后的图像无法对产生色彩失真的区域做出处理，并且由于其模型大小的限制，在不同的曝光错误场景中校正后的图像均难以达到良好的视觉效果；ECCM 方法[16]虽然能够达到合适的曝光程度，但是在局部色彩方面与参考图像之间存在一定差异；ECMEIQ 方法[18]虽然能够还原整体亮度信息，但是曝光校正后容易产生颜色失真. 此外，上述对比方法在曝光校正过后均产生了一定的局部细节模糊或无法还原出局部的细节纹理.



a. 输入图像 b. 参考图像 c. 本文 d. MSEC w/o L_{adv}[10] e. MSEC w/ L_{adv}[10] f. IAT[17] g. ECCM[16] h. ECMEIQ[18]

图 7 6 种方法在非智慧交通场景中的视觉效果对比

与所有对比方法相比，本文方法能够在不产生颜色失真与细节模糊的情况下，使得图像曝光水平达到合适的程度，取得了令人满意的视觉效果. 例如，第 1 行场景中，灯光过度曝光，导致缠绕在灯具上的藤蔓在输入图像中几乎不可见，同时造成了图像整体偏黄，本文方法在解决图像颜色失真的同时，对藤蔓部分做出了细节增强，而其他方法在曝光校正过后均难以看清藤蔓部分；第 2 行场景中，本文方法能够在曝光不足的环境中还原出场景应有的亮度信息，对草的颜色做出准确的校正，同时有效地保留了地上的石头颗粒等细节信息. 实验结果表明，本文引入的特征注意力机制能够有效地提取全局特征，改善图像整体的颜色分布与亮度信息，使得图像曝光水平达到正常. 验证了在提取曝光特征的 U-Net 结构中引入多尺度注意力机制能够有效地校正场景中局部区域的曝光错误，而最终的融合子网实现了将所有特征细节嵌入到最终的输出图像中，使其呈现出更多的细节纹理，更加贴合参考图像. 此外，结合内容损失与感知损失能够使得校正后的图像相较于其他方法展现出更自然的视觉效果.

2.2.2 智慧交通场景定性对比实验

为了验证本文方法在智慧交通中的有效性，对于交通监控场景中的车辆区域图像进行针对性对比实验，6 种方法的主观视觉效果对比如图 8 所示. 从车牌、车内等区域的颜色、细节等校正程度可知，本文方法明显优于其他对比方法. 可以看出，第 1 行场景中，MSEC w/o L_{adv} 方法[10]对于车

内区域的亮度有所提升，但是使得过度曝光的车牌区域产生严重模糊，而在第 2 行场景中导致整幅图像严重失真；第 1 行场景中，MSEC w/ L_{adv} 方法[10]产生了颜色失真，车牌区域也较模糊，而在第 2 行场景中，原本曝光正常的车牌区域周围出现了失真的色块；第 1 行场景中，IAT 方法[17]同样出现了颜色失真，对曝光不足的车内区域未能有效校正，而在第 2 行场景中出现了严重的虚轮廓现象；第 1 行场景中，ECCM 方法[16]不仅没有对车头过度曝光区域合理校正，反而进一步强化了由于强光产生的光斑，而在第 2 行场景中，车内区域产生了较为明显的模糊现象；第 1 行场景中，ECMEIQ 方法[18]车内区域产生颜色失真，而在第 2 行场景中，模糊了车牌区域．与所有对比方法相比，本文方法在第 1 行场景中能够有效地对车牌区域的曝光过度现象进行校正，使得文字部分更为清晰，同时能够改善车内的曝光不足现象；而在第 2 行场景中，本文方法能够有效地提升车内区域的亮度，并且保留了曝光正常的车牌区域信息．实验结果表明，本文方法能够有效地恢复交通监控场景中车辆区域由曝光错误而丢失的信息，更有利于执法人员对交通违法行为的判定．



a. 输入图像　b. 本文　c. MSEC w/o L_{adv}[10]　d. MSEC w/ L_{adv}[10]　e. IAT[17]　f. ECCM[16]　g. ECMEIQ[18]

图 8　6 种方法在交通监控场景中的视觉效果对比

2.3　消融实验

通过消融实验，进一步验证本文方法中注意力机制和损失函数的有效性．使用的测试集为 MSEC 数据集，并将测试集图像尺寸按比例缩放至最大边为 512 像素．

本文对曝光子网与全局子网中所使用的注意力机制进行消融实验，其量化结果如表 2 所示．可以看出，与仅引入特征注意力和仅引入卷积块注意力相比，本文方法的 PSNR 分别提升 0.484 4 dB 和 0.835 8 dB, SSIM 分别提升 0.005 3 和 0.008 8，证明了该方法全局子网中引入特征注意力以及在曝光子网中引入卷积块注意力，对于模型提升是有效的．

对损失函数进行消融实验，验证本文设计的损失函数的合理性，其量化结果如表 3 所示．可以看出，与仅使用 L_c 作为损失函数相比，本文方法的 PSNR 提升 0.572 2 dB, SSIM 提升 0.013 0，证明了该方法将 L_c 与 L_p 进行加权结合作为损失函数对于模型提升是有效的．

表 2　注意力机制消融量化结果

注意力机制	PSNR/dB	SSIM
仅特征注意力	21.317 1	0.862 4
仅卷积块注意力	20.965 7	0.858 9
特征注意力+卷积块注意力	**21.801 5**	**0.867 7**

注．加粗数据表示最优指标．

表 3　损失函数消融量化结果

损失函数	PSNR/dB	SSIM
L_c	21.229 3	0.854 7
$L_c + L_p$	**21.801 5**	**0.867 7**

注．加粗数据表示最优指标．

损失函数消融实验结果的视觉效果对比如图 9 所示．可以看出，在智慧交通场景车辆区域中，仅使用 L_c 作为损失函数时，校正后的图像中车身颜色出现失真，由于强光产生的光斑更为明显，因此其不符合真实场景信息，如图 9b 所示；而使用结合 $L_c + L_p$ 的损失函数时，校正后的图像视觉效果更佳．综合客观指标与主观效果，本文设计的损失函数能够有效地提升模型性能．



a. 输入图像 b. L_c c. $L_c + L_p$

图 9 损失函数消融实验结果的视觉效果对比

3 结 语

本文提出一种面向智慧交通的双重特征融合图像曝光校正方法. 该方法在全局子网中引入特征注意力机制以提取整体亮度、颜色等图像全局特征; 在曝光子网中引入多尺度注意力机制提取图像中不同尺度的曝光特征, 同时增强校正图像的细节纹理; 最后通过融合子网, 将双重特征映射至图像空间, 实现图像曝光校正. 定量对比实验与定性对比实验的结果表明, 本文方法均优于对比方法, 在智慧交通应用中能够有效地还原出真实场景中亮度、色彩、细节纹理等信息. 然而, 本文方法在处理极端欠曝或极端过曝的图像方面效果欠佳, 未来将进一步优化模型结构, 引入生成对抗网络, 使得曝光校正后图像的视觉效果更佳, 在交通监控场景中更具有实用性.

参考文献(References):

[1] Pizer S M, Amburn E P, Austin J D, *et al*. Adaptive histogram equalization and its variations[J]. Computer Vision, Graphics, and Image Processing, 1987, 39(3): 355-368

[2] Celik T, Tjahjadi T. Contextual and variational contrast enhancement[J]. IEEE Transactions on Image Processing, 2011, 20(12): 3431-3441

[3] Lee C, Lee C, Kim C S. Contrast enhancement based on layered difference representation of 2D histograms[J]. IEEE Transactions on Image Processing, 2013, 22(12): 5372-5384

[4] Land E H, McCann J J. Lightness and retinex theory[J]. Journal of the Optical Society of America, 1971, 61(1): 1-11

[5] Guo X J, Li Y, Ling H B. LIME: low-light image enhancement via illumination map estimation[J]. IEEE Transactions on Image Processing, 2017, 26(2): 982-993

[6] Wang S H, Zheng J, Hu H M, *et al*. Naturalness preserved enhancement algorithm for non-uniform illumination images[J]. IEEE Transactions on Image Processing, 2013, 22(9): 3538-3548

[7] Zhang Q, Yuan G Z, Xiao C X, *et al*. High-quality exposure correction of underexposed photos[C] //Proceedings of the 26th ACM International Conference on Multimedia. New York: ACM Press, 2018: 582-590

[8] Wei C, Wang W J, Yang W H, *et al*. Deep retinex decomposition for low-light enhancement[C] //Proceedings of the 29th British Machine Vision Conference. London: BMVA Press, 2018: 155

[9] Wang R X, Zhang Q, Fu C W, *et al*. Underexposed photo enhancement using deep illumination estimation[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2019: 6849-6857

[10] Afifi M, Derpanis K G, Ommer B, *et al*. Learning multi-scale photo exposure correction[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2021: 9157-9167

[11] Zheng Jian, Liu Hao, Yu Xiangchun, *et al*. Image exposure correction method based on inversion fusion framework[J]. Laser & Optoelectronics Progress, 2022, 59(12): Article No.1210002(in Chinese)
(郑剑, 刘豪, 于祥春, 等. 一种基于反转融合框架的图像曝光校正方法[J]. 激光与光电子学进展, 2022, 59(12): Article No.1210002)

[12] Zhao D, Li J, Li H Y, *et al*. Complementary feature enhanced network with vision transformer for image Dehazing[OL]. [2023-05-29]. https://arxiv.org/abs/2109.07100

[13] Guo C L, Li C Y, Guo J C, *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. Los Alamitos: IEEE Computer Society Press, 2020: 1780-1789

[14] Jiang Y F, Gong X Y, Liu D, *et al*. EnlightenGAN: deep light enhancement without paired supervision[J]. IEEE Transactions on Image Processing, 2021, 30: 2340-2349

[15] Cao Y H, Ren Y R, Li T H, *et al*. Over-exposure correction via exposure and scene information disentanglement[C] //Proceedings of the 15th Asian Conference on Computer Vision. Heidelberg: Springer, 2021: 407-422

[16] Nsamp N E, Hu Z Y, Wang Q. Learning exposure correction via consistency modeling[C] //Proceedings of the 32nd British Machine Vision Conference. London: BMVA Press, 2021: 12

[17] Cui Z T, Li K C, Gu L, *et al*. Illumination adaptive transformer[OL]. [2023-05-29]. https://arxiv.org/abs/2205.14871v2

[18] Eyiokur F, Yaman D, Ekenel H K, *et al*. Exposure correction model to enhance image quality[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Rec-

ognition Workshops. Los Alamitos: IEEE Computer Society Press, 2022: 676-686

[19] Qin X, Wang Z L, Bai Y C, *et al*. FFA-Net: feature fusion attention network for single image dehazing[C] //Proceedings of the 34th AAAI Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2020: 11908-11915

[20] Zhang K H, Li D X, Luo W H, *et al*. Dual attention-in-attention model for joint rain streak and raindrop removal[J]. IEEE Transactions on Image Processing, 2021, 30: 7608-7619

[21] Woo S, Park J, Lee J Y, *et al*. CBAM: convolutional block attention module[C] //Proceedings of the 15th European Conference on Computer Vision. Heidelberg: Springer, 2018: 3-19

[22] Li X, Wang W H, Hu X L, *et al*. Selective kernel networks[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2019: 510-519

[23] Qu Y Y, Chen Y Z, Huang J Y, *et al*. Enhanced pix2pix Dehazing network[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2019: 8152-8160