



(12) 发明专利

(10) 授权公告号 CN 114581560 B

(45) 授权公告日 2024. 04. 16

(21) 申请号 202210199669.2

G06N 3/08 (2023.01)

(22) 申请日 2022.03.01

(56) 对比文件

(65) 同一申请的已公布的文献号

CN 113642634 A, 2021.11.12

申请公布号 CN 114581560 A

CN 114067018 A, 2022.02.18

CN 114092330 A, 2022.02.25

(43) 申请公布日 2022.06.03

CN 108062569 A, 2018.05.22

(73) 专利权人 西安交通大学

CN 113283529 A, 2021.08.20

地址 710049 陕西省西安市咸宁西路28号

JP 6830707 B1, 2021.02.17

(72) 发明人 汪航 孙宏滨 程成 张旭翀

董猛; 吴戈; 曹洪玉; 景文博; 于洪洋. 基于注意力残差卷积网络的视频超分辨率重构. 长春理工大学学报(自然科学版). 2020, (01), 全文.

(74) 专利代理机构 西安通大专利代理有限责任公司 61200

专利代理师 闵岳峰

审查员 孟圆

(51) Int. Cl.

G06T 11/40 (2006.01)

G06N 3/0464 (2023.01)

G06N 3/0455 (2023.01)

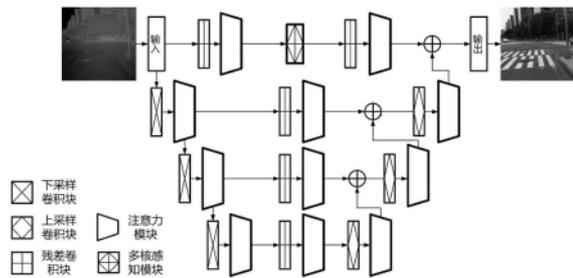
权利要求书3页 说明书6页 附图2页

(54) 发明名称

基于注意力机制的多尺度神经网络红外图像彩色化方法

(57) 摘要

本发明公开了一种基于注意力机制的多尺度神经网络红外图像彩色化方法,该方法首先利用二维卷积神经网络在不同分辨率尺度下对输入的红外图像对进行特征提取,然后通过注意力机制对提取出的高维特征信息进行提炼处理,最终将多尺度信息进行融合处理得到预测的彩色化红外图像。与现有的红外图像彩色化网络相比,本发明基于注意力机制与多尺度层次化结构构建神经网络算法模型,通过采用改进的空间注意力与多维特征连结机制,能够提高网络模型特征提取能力的同时有效降低模型复杂度,通过设计像素损失、边缘损失和感知损失的复合损失函数,进一步提高彩色化红外图像的质量。



1. 基于注意力机制的多尺度神经网络红外图像彩色化方法, 其特征在于, 包括以下步骤:

1) 从数据库中获取成对的红外图像和彩色图像, 进行图像预处理后分为训练集和测试集;

2) 构建基于注意力机制的多尺度神经网络模型, 将训练集输入该基于注意力机制的多尺度神经网络模型进行训练; 具体实现方法如下:

所述基于注意力机制的多尺度神经网络模型是从输入端到输出端的映射函数 $F(X^{N \times H \times W}; \theta)$, $\theta = \{\theta_1, \theta_2, \dots, \theta_n\}$, θ_1 表示第一层的参数, n 表示多尺度神经网络的总层数, $X^{N \times H \times W}$ 表示输入图像, N, H, W 分别为输入图像的维度, 高度和宽度, 损失函数为:

$$\begin{aligned} L(x, y) &= L_{pixel} + L_{edge} + L_{perceptual} \\ &= \sum_{n=1}^N \sum_{h=1}^H \sum_{w=1}^W (x - y) \\ &\quad + \sum_{n=1}^N \sum_{h=1}^H \sum_{w=1}^W \sqrt{(x' - y')^2 + \epsilon^2} + \sum_{t=1}^3 \sum_{n=1}^N \sum_{h=1}^H \sum_{w=1}^W (t_i(x) - t_i(y)) \end{aligned}$$

其中, x, y 分别表示输入红外图像和对应的彩色图, $L_{pixel}, L_{edge}, L_{perceptual}$ 分别为像素损失、边缘损失和感知损失, x', y' 分别表示利用拉普拉斯边缘算子得到的红外图像边缘图和对应的彩色图边缘图, $t_i(x)$ 和 $t_i(y)$ 分别表示红外图像和彩色图通过预训练的VGG19模型得到的特征图, ϵ 为常数, 且 $\epsilon = 10^{-3}$;

基于注意力机制的多尺度神经网络模型的训练是以取得损失函数 $L(x, y)$ 的最优值来估计映射函数 F 中的参数 θ 的最优值 θ' ;

取得损失函数 $L(x, y)$ 的最优值来估计映射函数 F 中的参数 θ 的最优值 θ' 具体为:

$$\theta_l^{i+1} = \theta_l^i - \eta \frac{\partial L}{\partial \theta_l^i}$$

其中, l 和 i 分别是卷积层的索引和迭代次数, η 是学习率, $\frac{\partial L}{\partial \theta_l^i}$ 是在第 i 次迭代时损失函数 $L(x, y)$ 对第 l 层的偏导数, 经过多尺度神经网络模型中参数多次的迭代更新, 损失函数达到最小, 此时模型中的参数就是映射函数 F 中参数 θ 的最优值 θ' ;

所述基于注意力机制的多尺度神经网络模型包括特征提取和特征融合;

特征提取包括多个卷积操作以对输入图像进行下采样并使得通道维度加大, 同时对图像特征进行提取得到特征图像, 公式为:

$$U_1^{N' \times W' \times H'} = F_1(X^{N \times H \times W}; \theta_1)$$

其中, N, H, W 分别为输入图像的维度, 高度和宽度; $U_1^{N' \times W' \times H'}$ 表示通过特征提取得到的图像, N', W', H' 表示输出图像的维度, 高度和宽度; θ_1 表示特征提取过程中的参数;

特征融合是对特征图像进行上采样并逐步使得通道的维度减少, 最后生成一个三维的

彩色化图像,公式为:

$$Y_1^{3 \times W^o \times H^o} = F_2(U_1^{N' \times W' \times H'}; \theta_2)$$

其中, θ_2 为在特征融合中的参数, Y_1 为生成的彩色化图像, W^o, H^o 分别表示输出图像的高度和宽度;

通过使用自适应矩估计的梯度优化方法,提高基于注意力机制的多尺度神经网络的收敛速度,给定超参数 $0 \leq \beta_1, \beta_2 \leq 1$, 给定时间步为 t , 动量 v_t 即小批量随机梯度 g_t 的指数移动平均:

$$m_t = \eta [\beta_1 m_{t-1} + (1 - \beta_1) g_t]$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) \text{diag}(g_t^2)$$

其中, η 表示学习率, m_t 和 v_t 分别表示梯度的一阶矩和二阶矩, 在迭代的阶段, m_t 和 v_t 的偏移矫正公式为:

$$m'_t = \frac{m_t}{1 - \beta_1^t}$$

$$v'_t = \frac{v_t}{1 - \beta_2^t}$$

并根据如上的公式对各个参数 μ_t 进行更新:

$$\mu_{t+1} = \mu_t - \frac{1}{\sqrt{v'_t + \epsilon}} m'_t$$

其中 β_1, β_2 和 ϵ 为预设的参数, μ 表示模型中的参数, μ_t 为第 t 个步长中 μ 的值, m'_t 和 v'_t 分别为经过偏移矫正后的梯度的一阶矩和二阶矩的估计值;

基于注意力机制的多尺度神经网络中的注意力机制,包括如下:

通道注意力机制:

$$F_c^{n \times h \times w} = M_c(F) \otimes F$$

$$M_c(F) = \sigma \left(\text{MLP} \left(\text{Maxpool}(F) \right) \right) = \sigma \left(W_1 \left(W_0 \left(F_{\max}^{1 \times 1 \times c} \right) \right) \right)$$

其中, σ 表示 Sigmoid 激活函数, F 为通道注意力模块的输入, \otimes 表示点乘操作, Maxpool 为最大池化操作, $F_{\max}^{1 \times 1 \times c}$ 为最大池化层的输出, $W_1 \in \mathbb{R}^{r/c \times c}$, $W_0 \in \mathbb{R}^{c \times c/r}$ 是两个权重矩阵的参数;

空间注意力机制:

$$F_s^{n \times h \times w} = M_s(F) \otimes F$$

$$M_s(F) = \sigma \left(f^{7 \times 7} \left(\text{Avgpool}(F), \text{Maxpool}(F) \right) \right)$$

其中 $f^{7 \times 7}$ 为卷积核大小的卷积操作, $F_s^{n \times h \times w}$ 为空间注意力模型的输出;

3) 将测试集输入训练好的基于注意力机制的多尺度神经网络模型进行测试,得到基于注意力机制的多尺度神经网络模型的新能和客观指标评估;具体实现方法如下:

进行测试得到的基于注意力机制的多尺度神经网络模型的性能和客观指标评价,包括

如下：

$$\text{PSNR} = 20 \log \left(\frac{1}{\text{MSE}(x, y)} \right)$$

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

其中 μ_x, μ_y 分别表示图像x和y的均值和方差, σ_x^2, σ_y^2 分别表示图像x和y的标准差, σ_{xy} 表示图像x和y的协方差, C_1 和 C_2 为常数；

4) 将待处理的红外图像输入测试通过的基于注意力机制的多尺度神经网络模型, 则输出彩色化图像。

基于注意力机制的多尺度神经网络红外图像彩色化方法

技术领域

[0001] 本发明属于计算机视觉技术领域,具体涉及一种基于注意力机制的多尺度神经网络红外图像彩色化方法。

背景技术

[0002] 红外图像彩色化是计算机视觉领域一个备受关注的研究问题,在安防监控、无人系统、军用装备等各类系统中具有广泛的应用前景。特别地,如何实现高质量、高分辨率、多细节的彩色红外图像是该技术亟待突破的挑战。近年来,基于深度学习技术的人工智能发展迅速,在目标检测、图像分类、语音识别等领域取得了突破性进展。红外图像彩色化作为计算机视觉的应用热点之一,同样得益于神经网络技术的不断创新,彩色化效果得到显著提升。

[0003] 常见的可见光灰度图像彩色化方法一般采用生成对抗网络实现彩色化功能。其生成器一般由U-net网络结构或者编解码网络结构(Encoder-Decoder)构成,这种结合了多尺度信息的生成器相比平坦网络可以获取更丰富的全局和局部信息。这样的网络设计在可见光灰度图像彩色化问题上取得了很好的结果。但由于红外图像和可见光图像在成像原理上的差异,直接将可见光图像彩色化方法用于红外图像彩色化中,会产生较为模糊的边缘以及细节信息,导致成像质量较差。

发明内容

[0004] 本发明的目的在于提供一种基于注意力机制的多尺度神经网络红外图像彩色化方法,以获得高质量的彩色红外图像。本发明提出的网络结构以多尺度网络作为骨干,设计创新的注意力机制与特征连结方法,以增强网络的特征提取与利用能力。此外,本发明提出了一种针对红外图像彩色化网络的复合损失函数,有效地将像素损失、边缘损失和感知损失结合起来。基于网络结构和损失函数上的创新,本发明在红外图像着色结果上达到了目前最好的效果。

[0005] 为达到上述目的,本发明采用如下技术方案来实现的:

[0006] 基于注意力机制的多尺度神经网络红外图像彩色化方法,包括以下步骤:

[0007] 1) 从数据库中获取成对的红外图像和彩色图像,进行图像预处理后分为训练集和测试集;

[0008] 2) 构建基于注意力机制的多尺度神经网络模型,将训练集输入该基于注意力机制的多尺度神经网络模型进行训练;

[0009] 3) 将测试集输入训练好的基于注意力机制的多尺度神经网络模型进行测试,得到基于注意力机制的多尺度神经网络模型的新能和客观指标评估;

[0010] 4) 将待处理的红外图像输入测试通过的基于注意力机制的多尺度神经网络模型,则输出彩色化图像。

[0011] 本发明进一步的改进在于,步骤2)的具体实现方法如下:

[0012] 所述基于注意力机制的多尺度神经网络模型是从输入端到输出端的映射函数 $F(X^{N \times H \times W}; \theta)$, $\theta = \{\theta_1, \theta_2, \dots, \theta_n\}$, θ_1 表示第一层的参数, n 表示多尺度神经网络的总层数, $X^{N \times H \times W}$ 表示输入图像, N, H, W 分别为输入图像的维度, 高度和宽度, 损失函数为:

$$\begin{aligned}
 L(x, y) &= L_{\text{pixel}} + L_{\text{edge}} + L_{\text{perceptual}} \\
 &= \sum_{n=1}^N \sum_{h=1}^H \sum_{w=1}^W (x - y) \\
 &+ \sum_{n=1}^N \sum_{h=1}^H \sum_{w=1}^W \sqrt{(x' - y')^2 + \epsilon^2} + \sum_{t=1}^3 \sum_{n=1}^N \sum_{h=1}^H \sum_{w=1}^W (t_i(x) - t_i(y))
 \end{aligned}$$

[0014] 其中, x, y 分别表示输入红外图像和对应的彩色图, $L_{\text{pixel}}, L_{\text{edge}}, L_{\text{perceptual}}$ 分别为像素损失、边缘损失和感知损失, x', y' 分别表示利用拉普拉斯边缘算子得到的红外图像边缘图和对应的彩色图边缘图, $t_i(x)$ 和 $t_i(y)$ 分别表示红外图像和彩色图通过预训练的VGG19模型得到的特征图, ϵ 为常数;

[0015] 基于注意力机制的多尺度神经网络模型的训练是以取得损失函数 $L(x, y)$ 的最优值来估计映射函数 F 中的参数 θ 的最优值 θ' ;

[0016] 取得损失函数 $L(x, y)$ 的最优值来估计映射函数 F 中的参数 θ 的最优值 θ' 具体为:

$$\theta_l^{i+1} = \theta_l^i - \eta \frac{\partial L}{\partial \theta_l^i}$$

[0018] 其中, l 和 i 分别是卷积层的索引和迭代次数, η 是学习率, $\frac{\partial L}{\partial \theta_l^i}$ 是在第 i 次迭代时损失函数 $L(x, y)$ 对第 l 层的偏导数, 经过多尺度神经网络模型中参数多次的迭代更新, 损失函数达到最小, 此时模型中的参数就是映射函数 F 中参数 θ 的最优值 θ' ;

[0019] 所述基于注意力机制的多尺度神经网络模型包括特征提取和特征融合, 特征提取包括多个卷积操作以对输入图像进行下采样并使得通道维度加大, 同时对图像特征进行提取得到特征图像, 公式为:

$$U_1^{N' \times W' \times H'} = F_1(X^{N \times H \times W}; \theta_1)$$

[0021] 其中, N, H, W 分别为输入图像的维度, 高度和宽度; $U_1^{N' \times W' \times H'}$ 表示通过特征提取得到的图像, N', W', H' 表示输出图像的维度, 高度和宽度; θ_1 表示特征提取过程中的参数;

[0022] 特征融合是对特征图像进行上采样并逐步使得通道的维度减少, 最后生成一个三维的彩色化图像, 公式为:

$$Y_1^{3 \times W^0 \times H^0} = F_2(U_1^{N' \times W' \times H'}; \theta_2)$$

[0024] 其中, θ_2 为在特征融合中的参数, Y_1 为生成的彩色化图像, W^0, H^0 分别表示输出图像的高度和宽度。

[0025] 本发明进一步的改进在于, $\epsilon = 10^{-3}$ 。

[0026] 本发明进一步的改进在于, 通过使用自适应矩估计的梯度优化方法, 提高基于注意力机制的多尺度神经网络的收敛速度, 给定超参数 $0 \leq \beta_1, \beta_2 \leq 1$, 给定时间步为 t , 动量 v_t

即小批量随机梯度 g_t 的指数移动平均:

$$[0027] \quad m_t = \eta[\beta_1 m_{t-1} + (1 - \beta_1) g_t]$$

$$[0028] \quad v_t = \beta_2 v_{t-1} + (1 - \beta_2) \text{diag}(g_t^2)$$

[0029] 其中, η 表示学习率, m_t 和 v_t 分别表示梯度的一阶矩和二阶矩,在迭代的阶段, m_t 和 v_t 的偏移矫正公式为:

$$[0030] \quad m'_t = \frac{m_t}{1 - \beta_1^t}$$

$$[0031] \quad v'_t = \frac{v_t}{1 - \beta_2^t}$$

[0032] 并根据如上的公式对各个参数 μ_t 进行更新:

$$[0033] \quad \mu_{t+1} = \mu_t - \frac{1}{\sqrt{v'_t + \epsilon}} m'_t$$

[0034] 其中 β_1, β_2 和 ϵ 为预设的参数, μ 表示模型中的参数, μ_t 为第 t 个步长中 μ 的值, m'_t 和 v'_t 分别为经过偏移矫正后的梯度的一阶矩和二阶矩的估计值。

[0035] 本发明进一步的改进在于,基于注意力机制的多尺度神经网络中的注意力机制,包括如下:

[0036] 通道注意力机制:

$$[0037] \quad F_c^{n \times h \times w} = M_c(F) \otimes F$$

$$[0038] \quad M_c(F) = \sigma \left(\text{MLP}(\text{Maxpool}(F)) \right) = \sigma \left(W_1(W_0(F_{max}^{1 \times 1 \times c})) \right)$$

[0039] 其中, σ 表示Sigmoid激活函数, F 为通道注意力模块的输入, \otimes 表示点乘操作,Maxpool为最大池化操作, $F_{max}^{1 \times 1 \times c}$ 为最大池化层的输出, $W_1 \in \mathbb{R}^{r/c \times c}$, $W_0 \in \mathbb{R}^{c \times c/r}$ 是两个权重矩阵的参数;

[0040] 空间注意力机制:

$$[0041] \quad F_s^{n \times h \times w} = M_s(F) \otimes F$$

$$[0042] \quad M_s(F) = \sigma(f^{7 \times 7}(\text{Avgpool}(F), \text{Maxpool}(F)))$$

[0043] 其中 $f^{7 \times 7}$ 为卷积核大小的卷积操作, $F_s^{n \times h \times w}$ 为空间注意力模型的输出。

[0044] 本发明进一步的改进在于,步骤3)的具体实现方法如下:

[0045] 进行测试得到的基于注意力机制的多尺度神经网络模型的性能和客观指标评价,包括如下:

$$[0046] \quad \text{PSNR} = 20 \log \left(\frac{1}{\text{MSE}(x, y)} \right)$$

$$[0047] \quad \text{SSIM} = \frac{(2\mu_x \mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}$$

[0048] 其中 μ_x, μ_y 分别表示图像 x 和 y 的均值和方差, σ_x^2, σ_y^2 分别表示图像 x 和 y 的标准差, σ_{xy} 表示图像 x 和 y 的协方差, C_1 和 C_2 为常数。

[0049] 与现有技术相比,本发明至少具有如下有益的技术效果:

[0050] 本发明提供的基于注意力机制的多尺度神经网络红外图像彩色化方法,利用注意力机制等适于红外图像特征的创新设计,有效解决了已有技术存在的特征提取能力不足、视觉感知质量不高等问题。

[0051] 进一步,本发明提出的针对红外处理的注意力机制可以从多层级的特征图中提取多尺度的上下文信息,并且借助于高分辨率的特征信息,能够在上采样过程中有效保留更多的细节纹理,提高网络的处理精度。

[0052] 进一步,本发明提出针对红外图像彩色化的复合损失函数,利用边缘损失和感知损失在保证红外图像彩色化质量的基础上提高结构信息质量和视觉感知效果。

附图说明

[0053] 图1为本发明基于注意力机制的多尺度神经网络红外图像彩色化方法的整体框架;

[0054] 图2为本发明注意力模块的示意图;

[0055] 图3为本发明下采样卷积模块的示意图;

[0056] 图4为本发明上采样卷积模块的示意图;

[0057] 图5为本发明残差卷积模块的示意图;

[0058] 图6为本发明多核感知模块的示意图。

具体实施方式

[0059] 下面将参照附图更详细地描述本公开的示例性实施例。虽然附图中显示了本公开的示例性实施例,然而应当理解,可以以各种形式实现本公开而不应被这里阐述的实施例所限制。相反,提供这些实施例是为了能够更透彻地理解本公开,并且能够将本公开的范围完整的传达给本领域的技术人员。需要说明的是,在不冲突的情况下,本发明中的实施例及实施例中的特征可以相互组合。下面将参考附图并结合实施例来详细说明本发明。

[0060] 本发明提供的基于注意力机制与多尺度神经网络的红外图像彩色化方法,其主要创新点在网络结构和损失函数两方面:

[0061] 1) 网络结构:首先,在多尺度网络中每一层对应的编码(Encoder)、解码(Decoder)卷积层后添加由通道注意力和空间注意力组成的注意力模块来增强网络的特征提取能力。通道注意力用于使彩色化网络专注于特征图(Feature maps)的某一层,空间注意力则使网络专注于特征图中的某一区域,二者是不同维度上的注意力机制。通道注意力和空间注意力的有效结合可以大幅提高彩色化网络的特征提取能力。其次,在多尺度网络结构中,跳级连结是常见的特征图融合方法,这样的跳级连结可以加速网络收敛。本发明在跳级连结中对于前级的特征图先进行提炼处理,再与后级的特征图完成连结操作。该设计的原因在于:较低层级中,由于特征图只经过了少量的卷积层,这些特征图包含的信息远少于较高层级。如先将较低层级中的特征图优先通过注意力机制和若干卷积层进行提炼处理,可以有效提高网络的彩色化能力。最后,考虑到多尺度网络具有在不同分辨率尺度下的特征提取能力并生成相应分辨率的高质量彩色化结果,本发明设计的红外图像彩色化网络在训练阶段将同时输出多种分辨率尺度的彩色化输出。

[0062] 2) 损失函数: 本发明使用的损失函数由像素损失 L_{pixel} 、边缘损失 L_{edge} 以及感知损失 $L_{\text{perceptual}}$ 三部分组成。首先, 像素损失 L_{pixel} 由视觉任务中常用的L1损失函数计算, 在训练前期还是训练后期均能提供有价值的梯度信息。其次, 本发明提出了一种新颖的边缘损失 L_{edge} , 首先利用拉普拉斯边缘算子提取生成的彩色化图像和基准图像的边缘信息, 随后用Charbonnier损失函数来计算二者之间的差异, 能够在不破坏原始红外图像内容的基础上有效增强彩色化图像的结构信息。最后, 本发明采用预训练的VGG19模型作为特征提取工具来计算感知损失 $L_{\text{perceptual}}$, 使彩色化图像与基准图像在高维度特征空间里具有更高的一致性, 从而提升彩色化图像的视觉感知质量。

[0063] 其中, 网络结构中, 在多层级网络的下采样阶段, 利用最大池化代替卷积操作。考虑到红外图像像素值普遍偏低, 卷积操作是某一区域内像素值的加权和, 而最大池化返回某一区域内像素的最大值。在像素值普遍偏低的情况下, 最大值往往代表着最具有信息的部分, 与卷积相比最大池化有着更好特征提取能力。

[0064] 网络结构中, 传统的多尺度网络通常会最大化下采样次数, 以更好的采集到多层级信息, 导致更多的参数量和更长的训练时间。本发明提出的多层级网络只对输入图进行了两次的下采样操作, 对比目前最优的彩色化网络有20%~30%的参数量减少和25%左右的速度提升。

[0065] 在空间注意力机制中, 采用最大池化进行特征的提取。空间注意力机制一般采用最大池化和平均池化来进行特征的提取, 然而平均池化对于红外图像的特征提取能力有限, 反而会导致模型复杂度的增加。单独使用基于最大池化的空间注意力机制即可在红外图像彩色化应用中取得较好的处理效果。

[0066] 在损失函数中, 添加边缘损失以增强彩色化红外图像的结构信息; 添加感知损失增强彩色化红外图像的视觉感质量。

[0067] 实施例

[0068] 参考图1-5, 本发明提出的基于注意力机制的红外图像着色方法主要包括特征提取、特征融合和生成着色图像等三个步骤:

[0069] 1) 图1是本发明的整体框架示意图。完成红外图像着色任务的神经网络模型输入是红外图像 I_{in} , 输出是着色图像 I_{out} 。在训练过程中, 输入红外图像与真实彩色图像是对准的。该网络将学习一个函数(模型) f 满足下列关系:

$$[0070] \quad f(I_{\text{in}}) = I_{\text{out}}$$

[0071] 具体地, 网络首先经过三次下采样卷积模块从原始输入红外图像 I_{in} 中提取四个不同分辨率的高维特征信息 F_1, F_2, F_3 和 F_4 , 然后通过注意力模块将 F_1, F_2, F_3 和 F_4 改善后与上采样模块中对应分辨率的高维特征信息连结起来, 最终利用连结后的高维特征信息预测出着色图像。如图1所示, 本发明的整体模型主要包括特征提取 f_1 、特征融合 f_2 和生成彩色图像 f_3 等三个模块。

[0072] 2) 特征提取 f_1 : f_1 采用一系列下采样卷积模块学习 I_{in} 的高维特征表示 F_1, F_2, F_3 和 F_4 , 可表达为 $F_1, F_2, F_3, F_4 = f_1(I_{\text{in}})$ 。首先, 输入红外图像的大小 $1 \times H \times W$, 经过一个预处理模块后得到一个大小为 $64 \times H \times W$ 的特征图 F_1 ; 此后该特征图经过一个下采样卷积模块后, 得到第一个高维特征信息 F_2 , 其大小为 $128 \times H/2 \times W/2$; 接着 F_2 经过第二个下采样卷积模块得到 F_3 , 其大小为 $256 \times H/4 \times W/4$; 最终, F_3 通过第三个下采样卷积模块得到 F_4 , 大小为 $512 \times H/$

$8 \times W/8$ 。每个特征提取过程的运算如下：

$$[0073] \quad F_1 = g_{\text{pre}}(I_{\text{in}})$$

$$[0074] \quad F_2 = g_1(F_1)$$

$$[0075] \quad F_3 = g_2(F_2)$$

$$[0076] \quad F_4 = g_3(F_3)$$

[0077] 其中 g_{pre} 表示预处理卷积操作, g_1 、 g_2 和 g_3 表示第一、二和三个下采样卷积模块。

[0078] 3) 特征融合 f_2 :基于注意力机制的特征过程可表示如下:

$$[0079] \quad F_{\text{out}} = g_{\text{att}}(F), F \in F_1, F_2, F_3, F_4$$

[0080] 其中, F_{out} 表示特征融合的输出。特征融合具体过程为:将特征提取阶段产生的特征图 F_1 、 F_2 、 F_3 和 F_4 作为输入,通过注意力模块和多核感知模块对特征图进行改善后输出与输入同大小的特征图。具体过程为,对于特征图 F_1 ,首先经过两组注意力模块对特征图进行改善,再经过多核感知模块提取出更丰富的细节信息,最后经过两组注意力模块得到输出 F_{1_out} ;对于 F_{2_out} 、 F_{3_out} 和 F_{4_out} ,为了减少参数量并加速训练,只通过注意力模块对其进行改善,特别地,相应的注意力模块的个数为4、2和1。

[0081] 4) 生成彩色图像 f_3 :改善后的特征图 F_{1_out} 、 F_{2_out} 、 F_{3_out} 和 F_{4_out} 有着不同的分辨率,为了融合其中的信息,依次从低分辨率的特征图逐步上采样到高分辨率,具体步骤如下。首先,特征图 F_{4_out} 通过上采样卷积模块生成与 F_{3_out} 同尺寸的特征图并与 F_{3_out} 连结起来生成特征图 F_{3_final} ;此后连接起来的总特征图经过上采样卷积模块生成与 F_{2_out} 同尺寸的特征图并与 F_{2_out} 连结起来生成特征图 F_{2_final} ;最终 F_{2_final} 通过上采样卷积模块生成与 F_{1_out} 同尺寸的特征图并与 F_{1_out} 连结起来生成特征图 F_{1_final} , F_{1_final} 通过输出模块生成最终的着色图像 I_{out} 。

[0082] 虽然,上文中已经用一般性说明及具体实施方案对本发明作了详尽的描述,但在本发明基础上,可以对之作一些修改或改进,这对本领域技术人员而言是显而易见的。因此,在不偏离本发明精神的基础上所做的这些修改或改进,均属于本发明要求保护的范围。

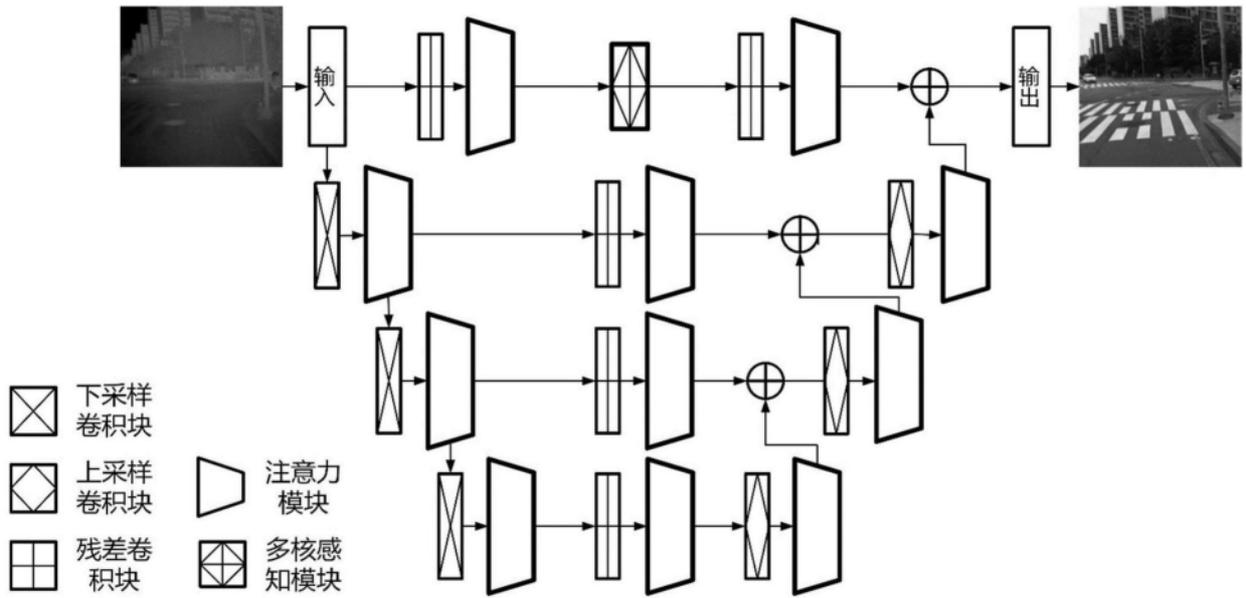


图1

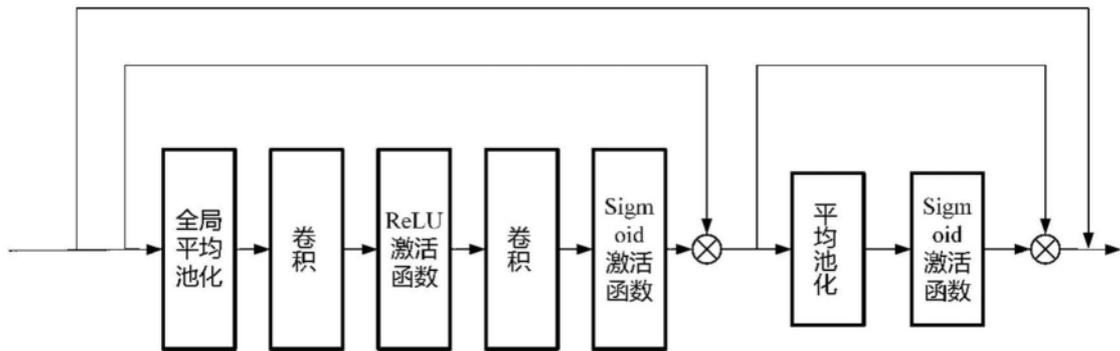


图2

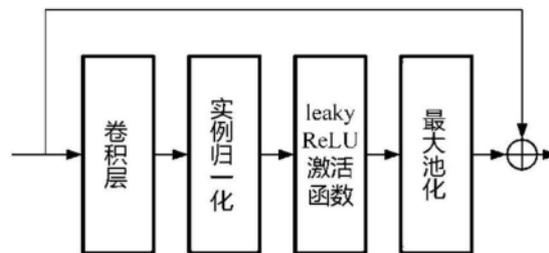


图3

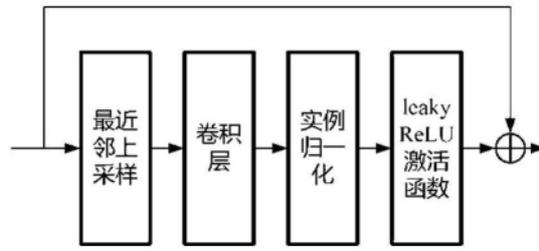


图4

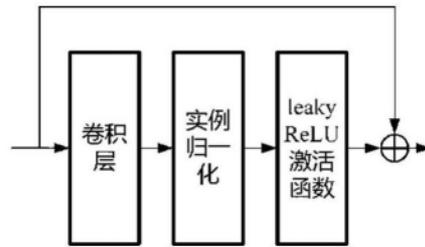


图5

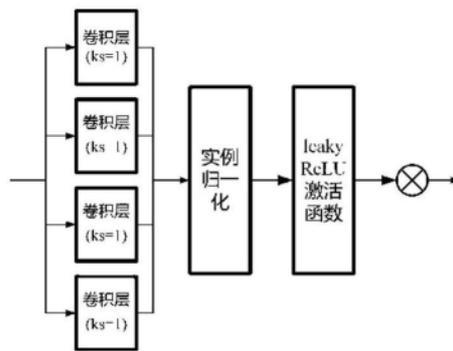


图6