多源融合的暗图增强：ILLVFusion

**①项目介绍**

本项目旨在解决在极端低光环境下图像质量问题，通过红外与低光可见光图像的融合提升图像细节纹理和对比度。我们提出了名为ILLVFusion的图像融合与增强方法，创新地应对了红外图像与低光可见光图像融合时的挑战。首先，我们使用DIPNet进行低光图像预处理，包括色彩恢复和亮度增强，从而改善图像的视觉效果；其次，我们引入CBFNet，其基于卷积神经网络（CNN），准确提取并融合红外和低光可见光图像的关键信息，实现了对不同模态图像特征的有效融合；最后，我们采用LCPNet进行图像的后期增强，利用拉普拉斯算子提升图像的纹理细节和对比度，进一步提高图像质量。

**②负责内容**

作为项目负责人，我主要负责CBFNet融合网络的设计和实施，采用多层编解码器结构以优化特征提取和图像重建的过程。同时设计了LCPNet进行图像后期增强。

**③数据集**

我们使用开源的MSRS数据集，并特别挑选了其中的夜间图像。数据集内容为适用于红外和可见光融合的多光谱道路场景，包含1444对高质量对齐的的红外与可见光图像；涵盖多种场景和条件，包括城市夜景、森林、建筑物等，提供了丰富的测试样本。

**④创新点**

项目的核心创新点是通过将红外图像与可见光图像融合来实现暗图像增强。红外图像具有在低光环境下清晰捕捉物体轮廓的能力，而可见光图像则包含丰富的颜色和纹理细节。我们的创新之处在于利用红外图像的轮廓信息和可见光图像的细节信息，通过融合两种图像，解决低光环境下图像细节和对比度不足的问题。融合后的图像不仅保留了红外图像的清晰度，还结合了可见光图像的色彩和纹理，从而显著提升了图像的整体质量。

**⑤评价指标**

我们的方法在MSRS数据集上的各项指标都取得了优异的结果，证明了其在低光环境下的有效性和优越性：

- EN: 熵, 衡量图像信息量，熵值越高，表示图像中的细节和信息越丰富。

- SD: 标准差, 衡量图像灰度值分布的离散程度，标准差越高，图像对比度越大，看起来越清晰和有层次感。

- SF: 空间频率, 用于评估图像细节的丰富程度，频率越高，图像的细节越多。

- MI: 互信息, 衡量两幅图像之间的共享信息量，用于评估融合图像保留原始图像信息的能力。

- SCD: 差异相关性之和, 衡量融合图像与原始图像之间的相关性，数值越大，表示融合效果越好。

- VIF: 视觉信息保真度, 衡量图像的视觉质量，特别是视觉感知效果，数值越高，视觉效果越好。

- AG: 平均梯度, 计算图像中灰度变化的梯度，用于反映图像的清晰度，梯度值越大，图像的边缘和细节越清晰。

- CC: 相关系数, 衡量两幅图像之间的线性相关性，数值越高，表示融合图像与原始图像的相似度越高。

**⑥需要改进的地方**

1）计算效率：模型结构较为复杂，导致计算时间长，特别是在处理高分辨率图像时。可以通过模型压缩或简化网络结构来提高计算效率。

2）数据依赖性：模型的性能可能依赖于训练数据的质量和多样性。如果训练数据覆盖不足，会导致模型在实际应用中表现不稳定。引入更丰富、更具代表性的训练数据有助于提高模型的鲁棒性。

3）适用性：当前模型在静态图像处理上效果不错，但在实时处理或处理视频等动态场景时无法满足需求。可以针对这些特定场景进行优化，例如提升模型的处理速度和适应动态变化的能力。

**⑦网络设计**

1）**DIPNet（暗图像预增强网络）**由亮度增强模块（Enhance Module）和自校准模块（Self-Calibrate Module）组成。

亮度增强模块负责对输入的低光图像进行初步的亮度提升。该模块通过多个3×3卷积层提取并增强图像的亮度特征。卷积层捕捉图像中的亮度信息，利用 ReLU 激活函数则用于对这些特征进行非线性处理，从而增强亮度不足区域，使得图像更接近理想的曝光水平。为了在增强过程中不丢失原始图像的信息，亮度增强模块采用了残差连接的设计，即将增强后的特征与原始输入图像相加。这样既保留了原始图像的细节，又提升了图像的整体亮度。

自校准模块对初步增强的图像进行进一步的处理。自校准模块的任务是优化图像的亮度，使其更加均匀和平衡，避免出现过度增强导致的过曝问题。这个模块同样使用一系列3×3卷积层对亮度特征进行细化调整，并通过 Sigmoid 激活函数将亮度特征归一化到 [0, 1] 的范围内，确保亮度的平滑调整。此外，自校准模块也采用了残差连接，将校正后的特征与之前的增强结果融合，从而保持图像的自然细节，避免过度调整带来的不自然感。

**损失函数**：1.5\*均方误差损失+1\*平滑度损失。选择均方误差损失（MSELoss）和平滑度损失（SmoothLoss），并将它们加权组合形成最终的损失函数。

均方误差损失（MSELoss） 用于衡量增强后的图像与原始输入图像之间的像素级差异。它计算了每个像素在增强后与原始图像之间的平方差异，并对这些差异求平均。通过最小化这个损失，模型被引导生成与输入图像在结构和亮度上更为接近的增强结果，保持图像的整体一致性。

平滑度损失（SmoothLoss） 的目的是保持增强后的图像在空间上的平滑性，防止在增强过程中产生不自然的伪影或过度锐化。该损失通过计算图像中相邻像素之间的亮度差异来实现，并且这种差异是通过高斯加权来平滑的。这种设计使得增强后的图像在视觉上更加自然，减少了因为亮度提升而可能产生的噪声和不连贯的过渡。

通过这样的损失函数设计，DIPNet 能够在提升低光图像亮度的同时，保持图像的结构完整性和自然感，使得增强后的图像具备良好的视觉效果。

**问题预测**：

**Q1: “为什么DIPNet需要分为亮度增强模块和自校准模块？这两个模块分别解决了什么问题？”**

亮度增强模块主要解决低光图像的亮度不足问题，通过提取和增强亮度特征，使图像更加接近理想的曝光水平。自校准模块则进一步优化亮度分布，确保图像不会因过度增强而导致过曝，同时保持亮度调整的平滑性和均匀性。

**Q2: “残差连接在DIPNet中具体起什么作用？为什么选择使用它？”**

残差连接主要用于防止信息丢失，通过将增强后的特征与原始输入图像相加，确保在提升亮度的同时保留原始图像的细节。它还帮助缓解梯度消失问题，使得网络更容易训练。

**Q3: “3×3卷积层为什么适合用于图像亮度的提取和增强？它有什么优势？”**

3×3卷积层具有良好的局部特征提取能力，能够捕捉图像中的亮度信息。同时，其小尺寸使得网络的参数量适中，计算效率高，适合逐步提取和增强图像的亮度特征。

**Q4: “为什么在亮度增强模块中使用ReLU激活函数，而在自校准模块中使用Sigmoid激活函数？它们分别解决了什么问题？”**

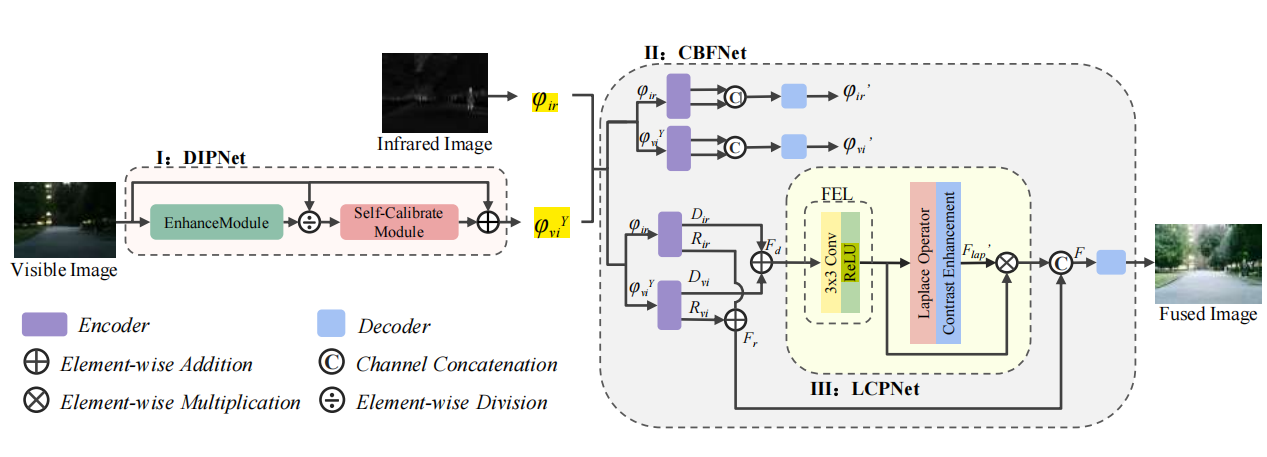
ReLU激活函数能够有效地增强亮度不足的区域，因为它对负值输出为零，对正值保持原值，适合处理亮度不足的情况。Sigmoid激活函数则将输出值限制在[0, 1]范围内，确保亮度调整后的图像不会过曝，同时使得亮度分布更加平滑。

**Q5: “为什么选择了均方误差损失和平滑度损失？这两种损失在DIPNet中分别起到什么作用？”**

均方误差损失用于衡量增强图像与原始图像之间的像素级差异，确保增强后的图像在结构和亮度上保持一致性。平滑度损失则确保图像在亮度增强过程中不会产生不自然的伪影或过度锐化，使得图像在视觉上更为自然和平滑。

**Q6: “为什么在损失函数中均方误差损失的权重设置为1.5，而平滑度损失的权重是1？这种加权组合对模型的效果有什么影响？”**

设置较高的权重给均方误差损失是为了确保图像增强的精度，即亮度的提升要足够明显和准确。而平滑度损失的权重较低，是为了在保持图像自然性的同时，不影响主要的亮度增强效果。这个加权组合保证了模型在增强亮度的同时不会引入过多的伪影和噪声。



2）**CBFNet（基于CNN的融合网络）**负责将红外图像和低光可见光图像的特征进行融合。CBFNet由两个主要阶段组成：图像重建阶段（Stage I）和图像融合阶段（Stage II）。它采用编码器-解码器架构，能够有效地提取并融合不同模态的特征，从而生成高质量的融合图像。

在图像重建阶段（Stage I），CBFNet 通过双分支编码器对红外图像和经过 DIPNet 增强后的可见光图像进行特征提取。编码器部分包含 Restormer block、相关特征编码器（RFE）和差异特征编码器（DFE）。Restormer block 引入了注意力机制，能够有效地提取输入图像的全局特征。RFE 负责提取两种模态之间的共同特征，如背景和大环境特征，而 DFE 则专注于捕捉红外和可见光图像之间的差异特征，如纹理和热辐射信息。在这些特征被提取后，融合层将它们进行融合，生成重建的原始图像。解码器随后将这些特征还原为红外和可见光图像，为第二阶段的融合操作提供基准。

图像融合阶段（Stage II） 是 CBFNet 的关键部分，它负责将经过 DIPNet 增强的可见光图像与红外图像融合在一起。这个阶段的编码器结构与 Stage I 相同，继续使用 Restormer block、RFE 和 DFE 提取输入图像的特征。在融合层中，RFE 提取的相关特征被融合生成 Fr，而 DFE 提取的差异特征则被融合生成 Fd。这些融合后的特征在通道上进行拼接，形成最终的融合特征 F。最后，解码器将这些融合特征还原为最终的融合图像。这张融合图像不仅保留了丰富的细节，还具有良好的对比度和视觉效果，能够在低光条件下提供更加清晰和自然的视觉体验。

**损失函数：**



STAGE 1:

STAGE 2:

编码器损失：在特征提取过程中，最大化相关特征（即在两个模态下都存在的相似特征）之间的相似性，同时最小化差异特征（即在两个模态下表现不同的特征）之间的相似性；

Lir 和 Lvi：分别表示红外图像和可见光图像的损失，通过结构相似性指数（SSIM）来衡量。

**问题预测：**

**Q1: “为什么选择使用编码器-解码器架构来实现图像融合？它有什么优势？”**

编码器-解码器结构是一种广泛应用于深度学习领域的神经网络架构，主要用于处理输入和输出之间存在复杂映射关系的任务。编码器（Encoder）负责接收输入数据，将其转换为紧凑的特征表示（隐含表示），其目标是提取输入数据的关键信息，并过滤掉不相关的细节，从而生成能够表示输入数据核心内容的编码；解码器（Decoder）接收编码器生成的特征表示，将其解码为目标输出格式。

编码器-解码器架构能够有效地压缩和提取图像的关键特征，减少信息冗余。同时，它能够将提取的特征还原成高质量的图像，确保在融合过程中不会丢失重要的信息。对于多模态图像融合，编码器-解码器架构可以在不同模态间提取相关和差异特征，从而生成更丰富的融合图像。

**Q2: “Restormer block 在编码器中起到了什么作用？为什么选择它？”**

Restormer block 引入了注意力机制，能够有效地提取图像中的全局特征，而不会增加过多的计算量。它在编码器中有助于捕捉输入图像中的长距离依赖关系，这对于融合不同模态的特征特别重要。选择 Restormer block 可以提高特征提取的精度，确保最终的融合图像质量。

**Q3: “RFE 和 DFE 的作用分别是什么？为什么要区分这两种特征？”**

相关特征编码器（RFE）提取的是红外图像和可见光图像之间的共同特征，如背景和环境信息，这些信息在两种模态中都是一致的。差异特征编码器（DFE）则捕捉两种模态间的独特特征，如可见光图像中的纹理和红外图像中的热辐射信息。区分这两种特征可以使融合过程更具针对性，确保最终的融合图像既保留了共同信息，又展示了各自的特征。

**Q4: “详细介绍是如何提取相关和差异特征的？”**

首先，RFE模块负责提取图像中的基础特征，通常是红外和可见光图像中共同存在的部分，如整体轮廓和背景信息。在这个模块中，首先使用7×7的卷积核对图像进行初步处理，这样的卷积操作可以捕捉到更多的上下文信息，提取出图像中的低频特征。这些特征在后续的批量归一化（BatchNorm）和ReLU激活后被进一步强化，保证了图像的尺度一致性和非线性表达。接着，通过多个残差块（Residual Blocks），进一步提取和强化这些基础特征。残差块中的残差连接确保了信息的传递和保留，避免了梯度消失问题，使得深层网络能够稳定地学习到图像的全局信息。这些提取出的基础特征在红外和可见光图像中都是一致的，因此被视为相关特征。

接下来，DFE模块则专门用于提取图像中的差异特征。与基础特征不同，差异特征通常是红外和可见光图像中不一致的部分，如红外图像的热辐射特征与可见光图像中的纹理细节。在这个模块中，使用了多个反向残差块（Inverted Residual Blocks），这些块通过1×1点卷积和3×3深度卷积来提取图像的高频信息。这种结构设计有助于捕捉图像中的细节特征，特别是那些在两个模态之间存在显著差异的部分。随后，DetailNode模块通过theta\_phi、theta\_rho和theta\_eta三个特征变换操作进一步处理这些差异特征，分别实现了特征的增强、乘性融合和加性融合。这些操作使得差异特征在最终的融合过程中得到了充分的放大和展示。

最终，在Restormer\_Decoder中，编码器提取的基础特征和细节特征被融合在一起。通过通道拼接的方式，将来自不同模态的相关特征和差异特征结合起来，并通过1×1卷积进行通道压缩，再经过若干层卷积处理后，生成最终的融合图像。这种融合方法确保了图像中既保留了红外和可见光图像的全局信息，又能够清晰地展示它们之间的差异，使得最终输出的图像既具备全局一致性，又在细节上呈现出丰富的模态特征。

**Q5: “融合层是如何将相关特征和差异特征进行融合的？为什么要在通道上进行拼接？”**

在融合层中，相关特征和差异特征分别进行融合，形成 Fr 和 Fd。然后，这些特征在通道上进行拼接，形成最终的融合特征 F。这种设计可以在特征级别上结合不同模态的信息，使得融合后的图像既能体现两种模态的优势，又能够保持结构上的一致性。通道拼接有助于保留更多的特征信息，为解码器的还原提供更丰富的输入。

**Q6: “为什么需要两个阶段来完成图像融合？Stage I 与 Stage II 之间有什么关系？”**

Stage I 的主要任务是为图像融合阶段（Stage II）提供基准，模型的编码器和解码器被训练来提取红外和可见光图像的特征，并通过这些特征重建出原始红外和可见光图像。在Stage II，模型利用Stage I中学到的特征表示，将红外和可见光图像的特征融合在一起，生成最终的融合图像。

Stage I通过图像重建的任务帮助模型的编码器学习如何提取和表示图像的特征，为Stage II的图像融合奠定基础。在Stage II，模型利用Stage I学到的特征表示进行跨模态融合，从而提升最终融合图像的质量。两者相辅相成，使得模型既能提取出高质量的图像特征，又能生成高质量的融合结果。

3）**LCPNet（拉普拉斯后增强网络）**是图像融合网络中的一个关键模块，旨在增强图像的纹理和对比度。它主要由三个部分组成：特征提取层、拉普拉斯算子和对比度增强模块。

首先，特征提取层通过一个3×3的卷积层和LeakyReLU激活函数来提取浅层特征，这些浅层特征包含了图像的基本信息，为后续处理打下基础。

接着，拉普拉斯算子进一步提取图像的深层特征，并通过模拟拉普拉斯操作增强图像的纹理和边缘。这一步骤也是通过一个3×3的卷积层来实现的，能够有效地突出图像中的细节部分。

最后，在对比度增强模块中，LCPNet通过将拉普拉斯增强后的特征与一个对比度增强因子k相乘，进一步提升图像的对比度。这里k被设置为5，以确保图像的清晰度与视觉效果之间的平衡。最终，LCPNet使得融合后的图像在细节和对比度上都有显著提升，特别是在低光条件下，能够提供更为清晰和自然的视觉体验。

**问题预测：**

**Q1: “为什么要使用3×3卷积层来提取浅层特征？这样的设计有什么优势？”**

3×3卷积层在提取局部特征方面表现出色，同时相对较小的卷积核能够保持计算的高效性。结合LeakyReLU激活函数，它能有效捕捉图像中的基本信息，为后续的深层特征提取奠定基础。

**Q2: “拉普拉斯算子在LCPNet中具体起到什么作用？为什么选择用它来增强图像纹理？”**

拉普拉斯算子是图像梯度的二阶导数，用于捕捉图像中灰度变化最剧烈的区域，即边缘和轮廓。在LCPNet中，拉普拉斯算子通过模拟边缘检测，增强了图像的高频信息，使得图像的纹理更加清晰。这对于在低光条件下提升图像的细节表现尤为重要。

**Q3: “为什么在对比度增强模块中，选择将拉普拉斯增强后的特征与对比度因子k相乘？”**

通过将拉普拉斯增强后的特征与对比度因子k相乘，可以进一步强化图像的对比度，使得图像在视觉上更为突出。k的选择（设为5）是为了在提升图像清晰度的同时，避免过度锐化带来的伪影和视觉不适。

**⑧项目中遇到的困难**

1. 在融合层中，如何选择适当的融合方式以确保最终输出图像的质量？

我们考虑了直接相加、通道拼接等多种融合方法。经过对比实验，我们最终选择了通道拼接的方式，这种方式能够保留更多的特征信息，并在解码器阶段通过1×1卷积进行通道压缩和细化处理，使得最终融合图像既具备丰富的细节，又能保持较高的视觉一致性。

2）在完成初步融合后，我们发现图像虽然融合了两种模态的信息，但整体看上去有些平淡，特别是在一些关键细节和边缘部分，显得不够锐利，对比度也不足，影响了图像的实际应用效果。

因此，我们设计了LCPNet来处理这个问题。使用拉普拉斯算子增强图像的边缘和纹理细节，同时结合对比度增强模块进行进一步调整。起初在设置对比度因子时，曾经出现过图像过度锐化的情况，导致一些细节部分失真。我通过多次实验，最终找到了一个合适的对比度增强因子，使得图像在清晰度和视觉舒适度之间达到了平衡。