

4 月工作汇报

Ku Jui

April 2024

Content

1 实验进度

2

1 实验进度

已知问题

信息丢失导致的图像细节质量恢复不高 [1]。即当前的深度学习方法增强后的图像仍然存在纹理模糊、光照不准确、色差等问题。几乎所有现有的深度模型都是基于 U-Net 结构，其中包含多个特征缩放操作 [2]。然而，特征缩放不可避免地会失去某些信息视觉原语 (Visual primitives) [3]。特征缩放导致的信息丢失使得增强后的图像丢失了重要的细节，包含了不需要的纹理和颜色。

我们认为目前模型不理想的原因如下：

- 1) 第一，未能在模型的初始阶段获取更多的图像细节和纹理信息，使得后续的高层次特征处理部分始终停留在图像语义和抽象信息上；
- 2) 第二，在特征缩放的过程，高级特征虽然深度增加了，但是图像尺寸裁剪使得细节和纹理信息不断丢失，后续在上采样的过程中则放大了这样一种缺失。
- 3) 第三，使用深度可分离卷积和 ReLU 激活的组合会导致低维流形 (Low-dimensional manifolds) 上感兴趣信息 (Interest information) 的丢失 [4]。

解决第一个问题：我们尝试以下方法通过增加 Stem 层的深度感知能力，Stem 两个卷积层组成，其中每个层有 64 个 3×3 的过滤器，stride 为 1，padding 为 1。

解决第二个问题：我们尝试引入一种称为自注意蒸馏 (Self-attention distillation) 的注意模块 [5] 以解决上述提到的特征缩放使得图像细节丢失的问题¹，使得低级特征可以学习到高级特征的语义信息 (Semantic information)，并通过正则化损失 (Regularization loss) 来约束这一过程，以有效地增强网络的表示学习能力。在自注意蒸馏中，不同层次主干网络中基于激活的 (Activation-based) 特征图会被提取并构建为自上而下的 (Top-down) 注意力蒸馏，以增强表示学习过程。具体来说，我们在 U-Net 的编码器网络中添加 Mimic 操作²

除此之外，我们尝试引入 FPN (Feature Pyramid Network) [6] 来处理不同尺度的目标³。具体而言，我们在 U-Net 网络中的每一个跳跃连接 (Skip connection) 加入 FPN 模块，采用层层递进的策略，尽可能的减少特征图下采样带来的特征损失。

解决第三个问题：我们需要使用一个称为线性瓶颈 (Linear bottlenecks) 和反向残差 (Inverted residuals) 的基本模块来构建编码器网络，基本模块如 Fig. 1 所示，使用激活函数为 ReLU6⁴

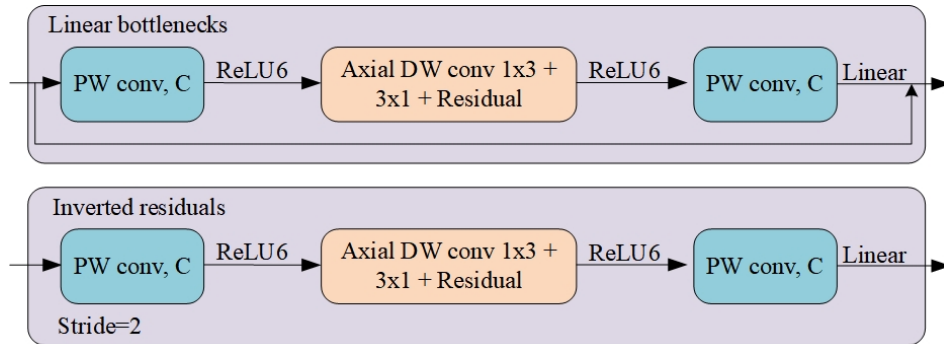


Figure 1: Linear bottlenecks 和 Inverted residuals

¹ 自注意力蒸馏 (Self-attention distillation): 这是一种注意力机制模块，旨在使低层次特征能够学习高层次特征的语义信息。通过这种方式，可以增强网络的表示学习能力。具体来说，它通过提取不同层次的特征图，并将它们构建为一个自上而下的注意力蒸馏过程，从而实现这一目标。

² Mimic 操作 (Mimic operation): 在知识蒸馏 (Knowledge Distillation) 过程中，让一个模型（通常是较小的学生模型）模仿另一个模型（通常是较大的教师模型）的行为或输出。在自注意力蒸馏的上下文中，Mimic 操作指的是让低层次特征模仿高层次特征的行为，以学习更丰富的语义信息。

³ FPN (Feature Pyramid Network) 主要包含两个关键步骤：自底向上的特征提取和自顶向下的特征融合。

1) 自底向上的特征提取阶段通过一个基础网络从输入图像中提取出不同尺度的特征图。这些特征图具有不同的感受野和语义信息。

2) 自顶向下的特征融合阶段将高层次特征与低层次特征进行融合，以获得更加丰富和具有多尺度信息的特征表示。具体来说，FPN 使用上采样操作将较高级别的特征图进行插值得到与相应低级别特征图尺寸相匹配的特征图，然后通过逐元素相加的方式将它们进行融合。

$$f(x) = \min(\max(0, x), 6)$$

References

- [1] Zhao Zhang, Huan Zheng, Richang Hong, Jicong Fan, Yi Yang, and Shuicheng Yan. Frc-net: A simple yet effective architecture for low-light image enhancement. *IEEE Transactions on Consumer Electronics*, 2023.
- [2] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III* 18, pages 234–241. Springer, 2015.
- [3] Yulun Zhang, Kunpeng Li, Kai Li, Gan Sun, Yu Kong, and Yun Fu. Accurate and fast image denoising via attention guided scaling. *IEEE Transactions on Image Processing*, 30:6255–6265, 2021.
- [4] Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, and Liang-Chieh Chen. Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4510–4520, 2018.
- [5] Yanhui Guo, Xue Ke, Jie Ma, and Jun Zhang. A pipeline neural network for low-light image enhancement. *IEEE Access*, 7:13737–13744, 2019.
- [6] Tsung-Yi Lin, Piotr Dollár, Ross Girshick, Kaiming He, Bharath Hariharan, and Serge Belongie. Feature pyramid networks for object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2117–2125, 2017.