

Summarize

2074389405

June 2023

1 动机

JPG 压缩算法在网络上已经得到了广泛的应用，即使出现了 JPG2000 等更高效的压缩方法，JPG 算法依旧是主流的压缩方法之一。JPG 存在提升空间，我们通过深度学习方法对图像进行“前处理”，在不破坏 JPG 压缩流程的前提下，提升压缩的熵率性能，从而使我们的工作具有应用价值。

如图一所示，对原始图片进行 JPEG 压缩，我们使用网络进行预处理，从而使图片变得“适应”JPG 压缩，同时达到更低的码率和更低的失真。

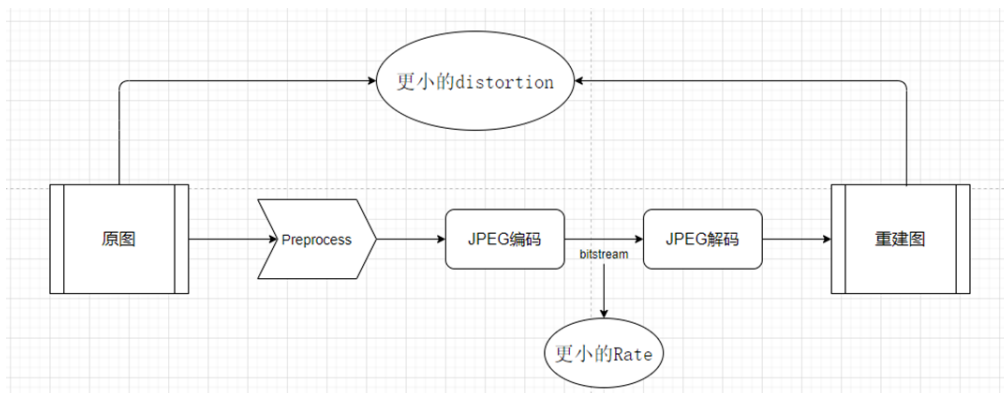


图 1: 预期流程

2 实现方式

在图 1 中，JPG 压缩、JPG 解压缩这一过程产生了码率和重建的图像，然而这一过程是不可微的，因此我们设计了两个辅助网络解决这个问题。网络一预测码率，网络二预测重建图像，他们都是可微的，从而解决不可微的问题。

训练时，设置 $Loss = Distortion + \lambda * Rate$ 。

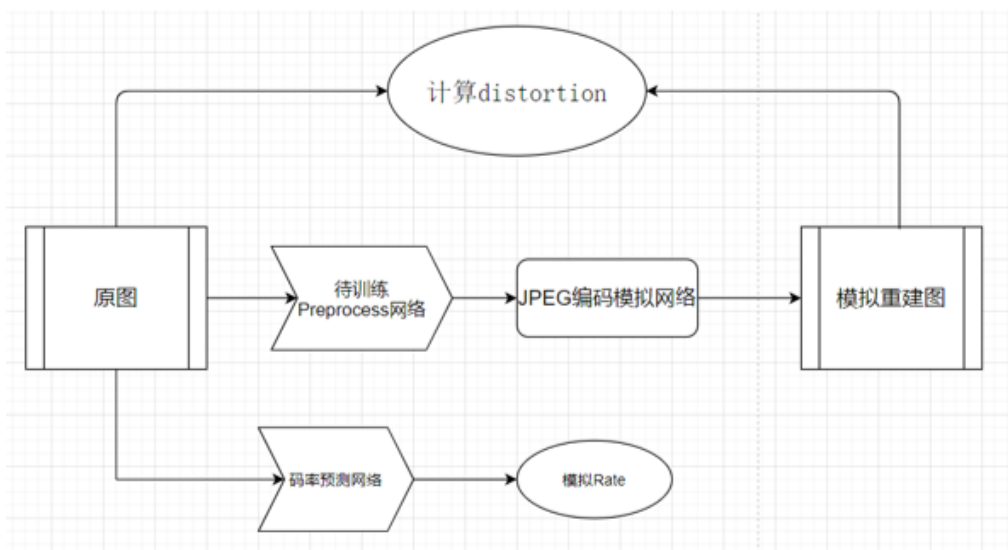


图 2: 实现方法

3 数据集

对 imao 数据集，对其中的每张图片进行 Quality=5 的 JPEG 压缩，记录产生的重建图和码率，

4 预处理网络

采用 vdsr 网络，该网络含有残差结构，能够学习到对图像的精细化处理。实验发现，使用该网络进行图像还原，相对误差约为 $1e-5$ 。

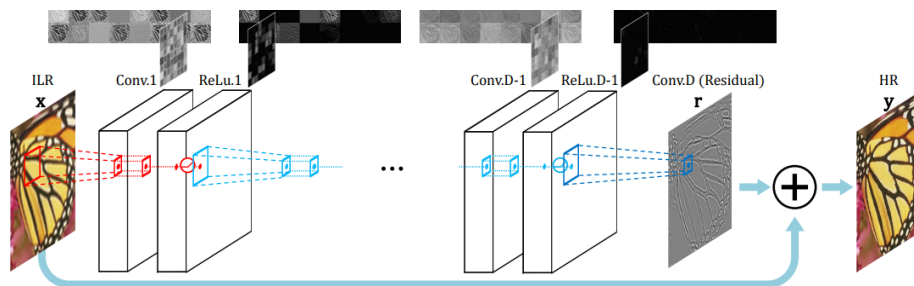


图 3: vdsr 网络结构

5 重建图预测网络

6 码率预测网络

6.1 网络结构与训练

受到 NIMA 图像评分的启发，使用了 vgg 网络提取特征，将提取的多维特征归一化，然后视为码率的概率分布，将码率均值作为最终预测值。

使用了 E 类 VGG 网络，在 quality=5 的数据集上将最终输出特征数调整为 500，并将输出 feature 点乘码率的取值集合：**【0.000,0.001, ...,0.499】**得到预测结果。在 quality=25 的数据集上调整为 1500，并将输出 feature 点乘一个向量：**【0.000,0.001, ...,1.499】**得到预测结果。

训练时需要注意：不能直接使用该输出值进行梯度回传，需要将码率标签转换为类别标签，计算 feature 与类别标签之间的交叉熵，将交叉熵作为损失函数，否则不收敛；输入图片的像素值范围需要缩放至 **【0,255】**，否则会导致网络不收敛；optimizer 需要使用 SGD 优化器，学习率设置为 1e-4 量级，使用 Adam 优化器会导致网络不收敛。

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224×224 RGB image)					
conv3-64	conv3-64 LRN	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 conv1-256	conv3-256 conv3-256 conv3-256	conv3-256 conv3-256 conv3-256 conv3-256
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

图 4: vgg 网络结构

6.2 与其他方法比较

7 结果

7.1 $loss = Rate$

仅使用预测码率进行梯度回传，这里使用平均值作为预测值，和训练时不同。观察到随着预测码率下降，真实码率也下降，图象丢失细节内容。

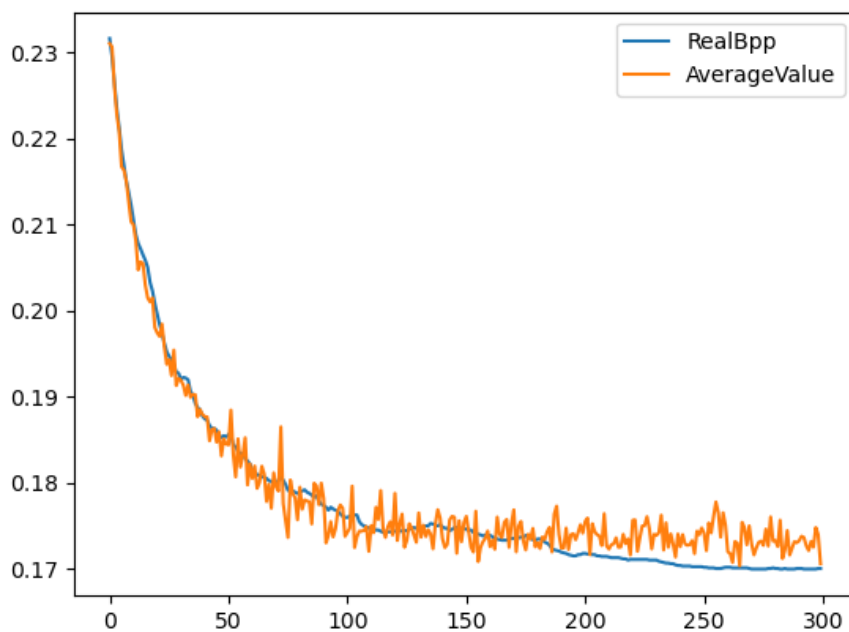


图 5: 码率下降过程

7.2 $Loss = Distortion + \lambda * Rate$

仅使用码率预测网络几乎不能提升熵率性能，在严格限制失真的条件下，性能有微弱提升

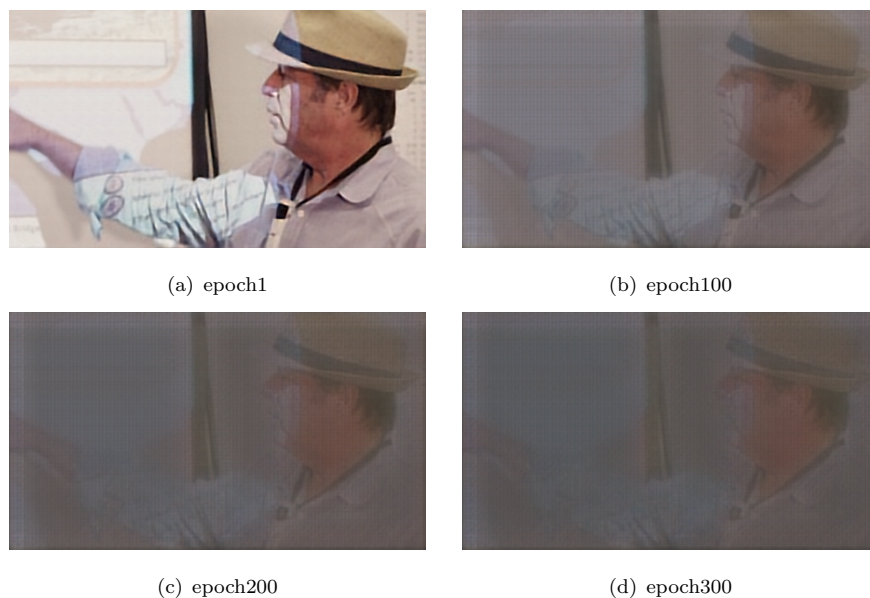


图 6: 码率下降过程

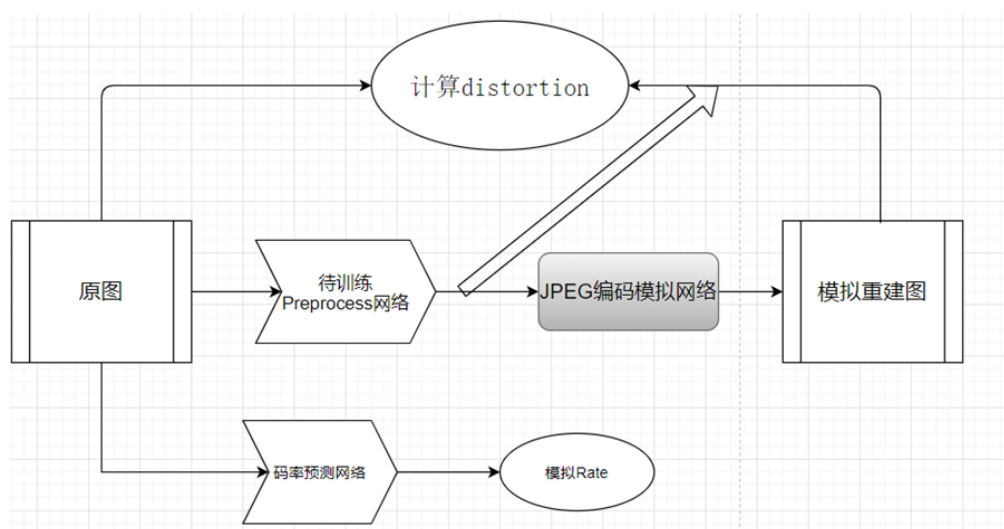


图 7: 仅使用码率预测网络

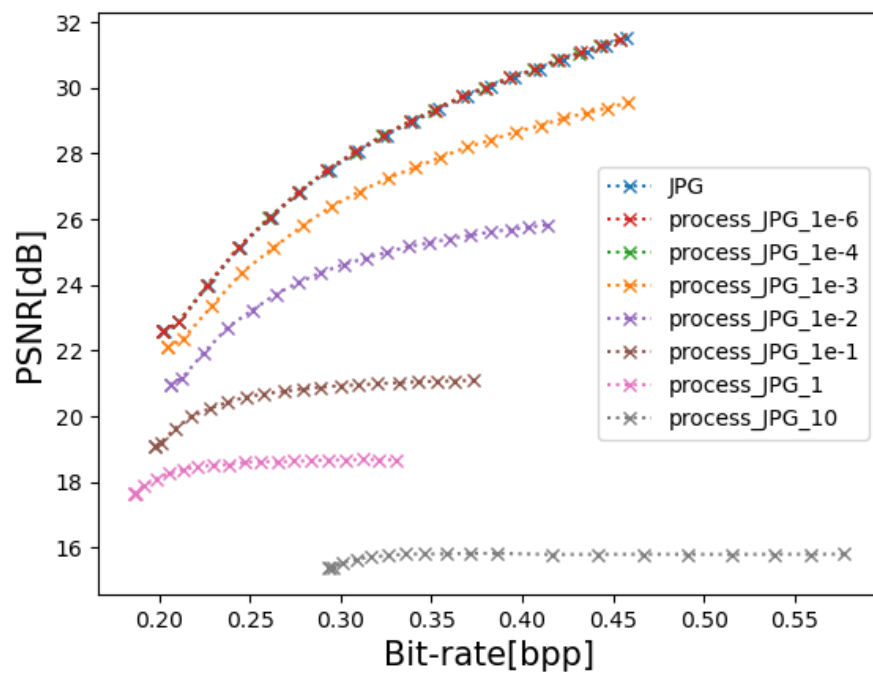


图 8: 码率曲线

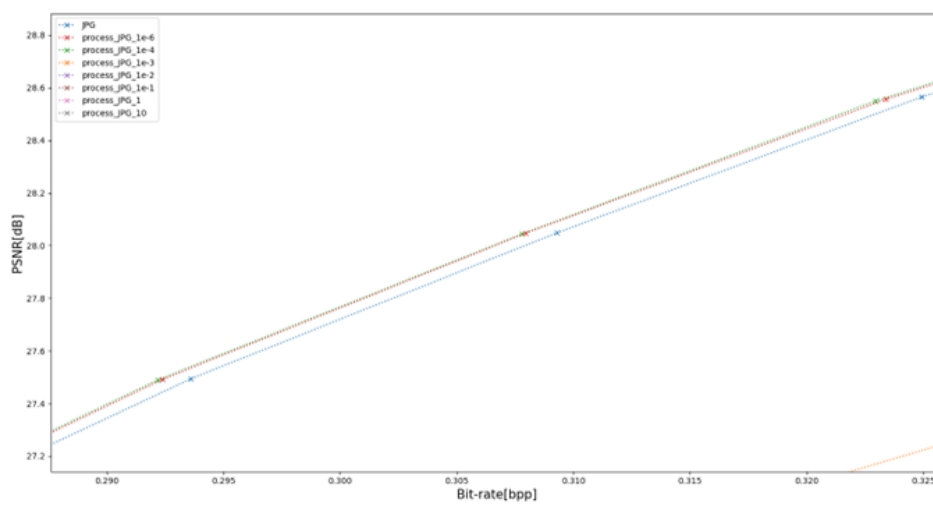


图 9: 码率曲线