基于深度学习的数据压缩实验说明

一、实验要求

**1.1 编程实现**

结合参考文献和网上相关资源，在PyTorch环境下实现一种基于深度学习的数据压缩模型，课程组会在教学云平台中提供参考文献[1]的参考代码。

使用所给的Open Images数据集，以MSE为优化指标，在训练集上训练自己实现的模型；改变速率和失真的权衡参数𝜆，训练得到不同bpp下对应的模型；在Kodak数据集上对模型性能进行测试，计算重建图像集的平均PSNR（对重建图像集计算平均MSE后使用其计算平均PSNR）和相应的bpp（参考代码中已给出相应计算方法），得到以bpp为横坐标、PSNR为纵坐标的率失真曲线。

探索改变模型架构和相关超参数，如使用Transformer作为核心架构、改变batch\_size等，尽力提升bpp-PSNR的压缩性能曲线。

**1.2 实验报告**

每一组应给出bpp在0.0到1.5的范围内至少4个点（两点间间隔不应太小，尽量不小于0.2，）相应的PSNR数值，需要给出运行结果的截图，截图中需包含训练和测试过程中的相关指标，包括但不限于总loss、MSE loss、bpp、PSNR，否则判定实验结果无效。参考代码中给出了输出相应指标的代码，结果类似下图（模型测试时输出PSNR的代码请注意补充）。



每一组应在实验报告中给出相应bpp下Kodak测试集的重建图片，只需给出测试集中顺序为本组组号的重建图片即可（例如第1组给出序号为1的图片即可）。

按照所给实验报告模板给出相应结果分析和问题分析，并需要回答相应思考问题，具体问题之后会以单独的word给出。

二、数据集

**2.1 训练集**

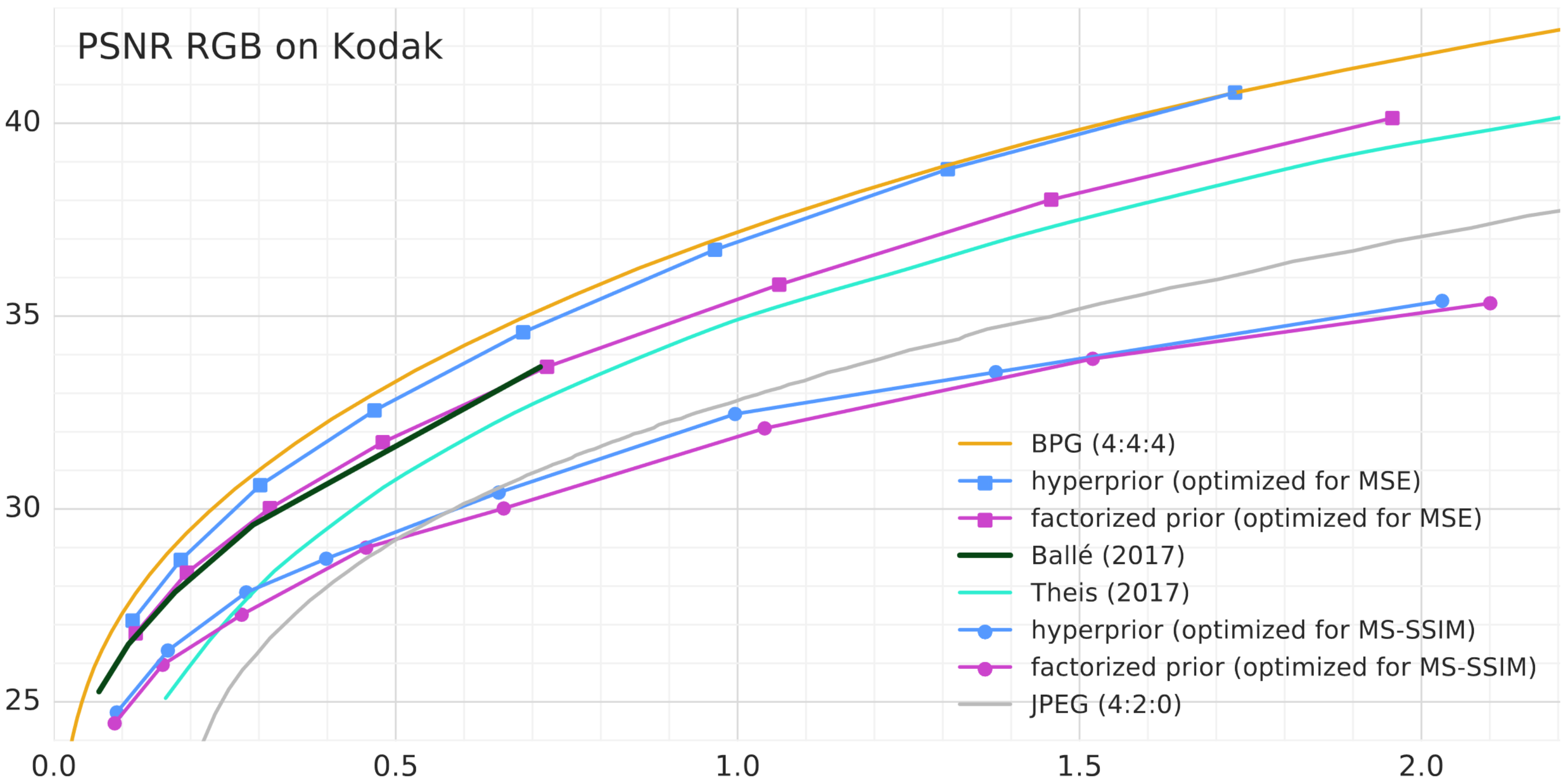
模型的训练使用Open Images数据集。Open Images 是包含约 900 万张图像的数据集，带有图像级标签、对象边界框、对象分割掩模、视觉关系和局部叙述。课程组从其中提取共40441张jpg格式的图像（约12G，课程组会直接提供）用于模型训练，要求训练的图像尺寸全部随机裁剪为（课程组提供的参考代码中已包含相关操作）。

**2.2 测试集**

模型的测试使用Kodak数据集。Kodak数据集是由伊士曼科达公司发布的图像数据集，在图像表示领域中被广泛用于测试算法的性能。它包含24张分辨率为768×512或512×768的无损真彩色图像，包含RGB三个通道，每个像素点使用24比特表示，全部采用PNG格式（课程组会直接提供）。

三、评分要求

**3.1 benchmark**



曲线2

曲线1

上图中**蓝色**箭头所指向的是hyperprior方案以MSE为优化指标的性能曲线，命名为**曲线1**，**紫色**箭头所指向的是factorized prior方案以MSE为优化指标的性能曲线，命名为**曲线2**。曲线中具体数值如下。

曲线1： bpp-psnr

0.13129340277777776——27.581536752297392,

0.20889282226562503——29.196703405493214,

0.3198581271701389——30.972162072759534,

0.47835625542534727——32.83818257445048,

0.6686876085069443——34.52626403063645,

0.9388258192274304——36.74334835426406,

1.2591722276475694——38.58384824012354,

1.6602240668402775——40.556865931529494

曲线2：bpp-psnr

0.1226230197482639——26.908817394289304,

0.18852742513020834——28.217925699960002,

0.2878078884548611——29.616915231568353,

0.44037882486979174——31.27708728897609,

0.6481357150607638——32.956122820153084,

0.9669765896267358——35.380922291056244,

1.351165771484375——37.39693190227357,

1.8332655164930556——39.621314292092684

**3.2 评判标准**

性能表现：将小组的bpp-PSNR曲线与benchmark曲线进行对比，达到曲线1即可得到实验性能部分80%的分数；若未达到曲线1但达到曲线2可得到实验性能部分70%的分数；若未达到曲线2会得到60%及以下的分数。

性能排名：对所有小组的bpp-PSNR曲线性能进行排名，前3名分别得到性能部分剩余分数的40%、35%和30%分数，4-6名得到20%的分数，7-10名得到15%的分数，剩余小组得到10%的分数。（例，第1名可以得到实验性能部分总分的120%，即获得20%的额外加分）

问题回答：实验报告中对思考问题的回答，综合考察其合理性与创新性，给出相应分数。

四、参考文献

下列参考文献中，第1篇文献可用于参考实现实验性能的基本要求即benchmark，其余文献可参考实现更佳的实验性能。每篇参考文献后给出了相应的访问链接，同时也会在教学云平台上上传下列参考文献。

[1] Variational image compression with a scale hyperprior. https://arxiv.org/pdf/1802.01436.pdf

[2] Joint Autoregressive and Hierarchical Priors for Learned Image Compression. https://arxiv.org/pdf/1809.02736.pdf

[3] Learned Image Compression with Discretized Gaussian Mixture Likelihoods and Attention Modules. https://arxiv.org/pdf/2001.01568.pdf

[4] Channel-wise Autoregressive Entropy Models for Learned Image Compression. https://arxiv.org/pdf/2007.08739.pdf

[5] QARV: Quantization-Aware ResNet VAE for Lossy Image Compression. https://arxiv.org/pdf/2302.08899.pdf