

基于深度学习译码器的图像压缩

汪进

摘要

本文基于深度学习和深度学习网络，对随机码和结构化编码（比如极化码）进行一次性译码。可以观察到，结构化编码更容易学习，神经网络能够正确译码在训练期间未见过的码字。此外，通过信道编码与信源编码之间的对偶关系，基于神经网络译码器实现了图片的压缩。

关键词：深度学习；极化码；信道编码；信源编码

1 引言

近年来，采用神经网络设计信道译码算法成为热门的研究方向，由于神经网络具有并行结构，在提升译码吞吐率方面有较大潜力。极化码是由 Arikan 教授^[1]提出的一种目前唯一一个被证明可以达到香农极限的编码方式。其核心是通过信道极化，使各个子信道呈现出不同的可靠性。当码长接近于无穷时，部分信道的容量接近于 1 的完美信道，而另一部分信道容量接近于 0 的纯噪声信道，选择在接近于 1 的信道上直接传输信息以逼近信道容量。在解码端可以使用简单的逐次干扰抵消解码的方法，以较低的复杂度获得与最大似然解码相近的性能。

极化码也可以采用神经网络译码，本文复现的文章是最早提出基于全连接神经网络的极化码译码方案^[2]。极化码具备一定的编码规则结构，因此神经网络可以学习解码算法。不过受到维度灾难的影响，神经网络解码目前仅限于短码，即 $N \leq 64$ 。

2 相关工作

近年来，采用神经网络（Neural Network, NN）设计信道译码算法成为热门的研究方向。由于极化码具有并行结构，在提升译码吞吐率方面有较大潜力。本文复现的文章是最早提出基于全连接神经网络的极化码译码方案^[2]。文献^[3]提出 min-sum 神经网络译码算法，文献^[4]研究了加权 BP 译码算法。文献^[5]应用遗传算法对极化码的构造与译码进行了优化。高健和牛凯等在^[6]设计了单比特神经网络译码算法。

对于极化码而言，信源编码是信道编码的反向应用，即可采用信道译码算法进行信源压缩，采用信道编码方法进行信源恢复。该思想由 Hussami、Korada 和 Urbanke 提出^[7]，Arikan 也研究了无失真信源编码的应用问题，讨论了极化码的信源编码与信道编码的对偶问题，并给出了一些初步分析^[8]。Korada 等指出信道极化码对于有损信源编码是渐进最优方法，能够达到率失真函数 $R(D)$ 的下界^[9]。

3 本文方法

3.1 本文方法概述

在这项工作中，想使用一个神经网络来译码噪声码字。在发射端，将 k 个信息位编码成长度为 N 的极化码码字，经过调制后，在有噪声信道上传输。在接收端，接收到有噪声版本的码字，神经网络解码器的任务是恢复 k 位的信息位。该神经网络的输入层有 N 个神经元，用来接收训练和测试样本，输

出层有 k 个神经元，分别输出 k 位信息比特。与迭代解码相比，神经网络只需通过每一层，一次就能得到估计值。由于该原理支持低延迟实现，我们将其称为一次性解码。

神经网络模型结构如下图所示。神经网络分别采用 128、64 和 32 神经元的三层隐藏层。每个隐藏层都使用一个 ReLU 激活函数，因为它非线性，同时非常接近线性，这有助于模型优化。输出层表示信息位，Sigmoid 函数使得输出神经元在 0 和 1 之间，可以解释为传输信息位为“1”的概率。如果概率接近标签的位，则损失应只略微增加，而较大的误差应导致非常大的损失。可以使用均方误差 (MSE) 和二元交叉熵 (BCE) 损失函数来衡量。

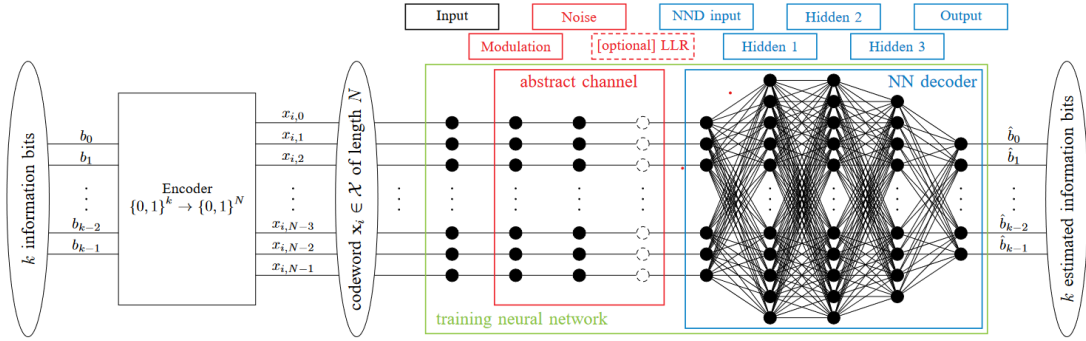


图 1: 方法示意图

损失函数的梯度在整个训练集 X 上使用 Adam 计算，这是一种随机梯度下降优化的方法。由于我们的体系结构中的噪声层在每次使用时都会生成一个新的噪声实现，所以神经网络解码器永远不会看到相同的输入两次。由于这个原因，尽管训练集只有 2^k 码字的有限大小，但我们可以通过简单地增加 Mep 的 epoch 数来训练一个本质上无限的训练集。

对于机器学习领域来说，获取标记训练数据通常是一项非常困难和昂贵的任务。但是使用神经网络进行信道编码是具有优势的，因为我们要处理的是人造信号，因此我们可以生成尽可能多的训练样本。此外，由于产生了有噪声码字，传输的信息位显然是已知的，因此可以直接获得期望的 NN 输出，也称为标签。为了简单起见，采用了二相移键控 (BPSK) 调制和加性高斯白噪声 (AWGN) 通道。其他信道可以直接采用，这种灵活性可能是基于神经网络的解码的特殊优势。对于极化码译码，Mep = 2^{18} epoch 已经达到了接近 MAP 的性能，而对于随机码，则需要更大的 NN 或更多的训练 epoch。实验结果如下图所示：

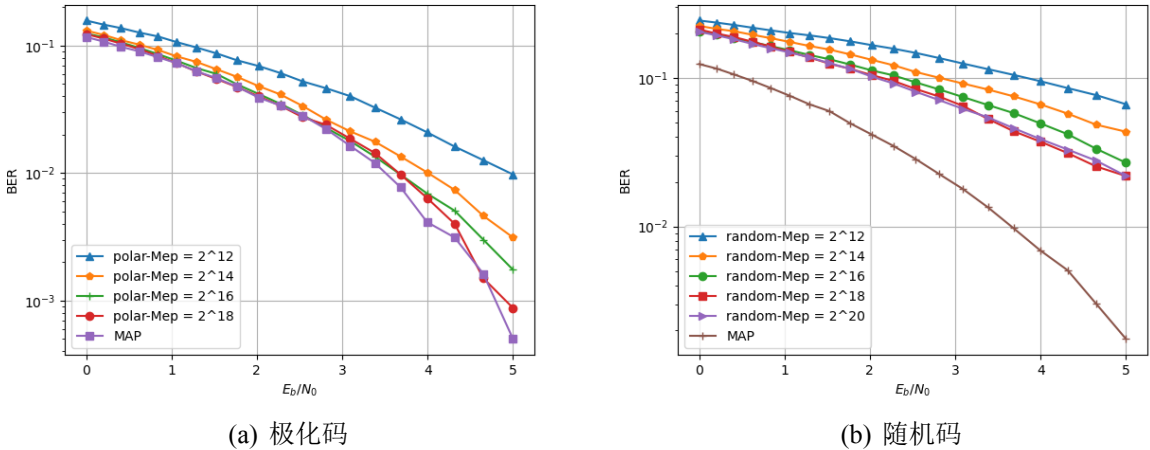


图 2: epoch 增加误码率下降

在固定的 NN 设计和训练周期数下，极坐标编码的 NNDs 总是比随机编码的性能更好。这第一次表明，结构化码 (如极坐标码) 比随机码更容易学习。对于小块长度，我们能够实现 MAP 性能解码随机码和极化码。但是，随着码字中信息位的增加，学习受到指数复杂度的限制。

在复现论文中的极化码神经网络译码器的基础上，利用信道编码与信源编码之间的对偶性，可以使用译码器来实现信源的压缩过程。

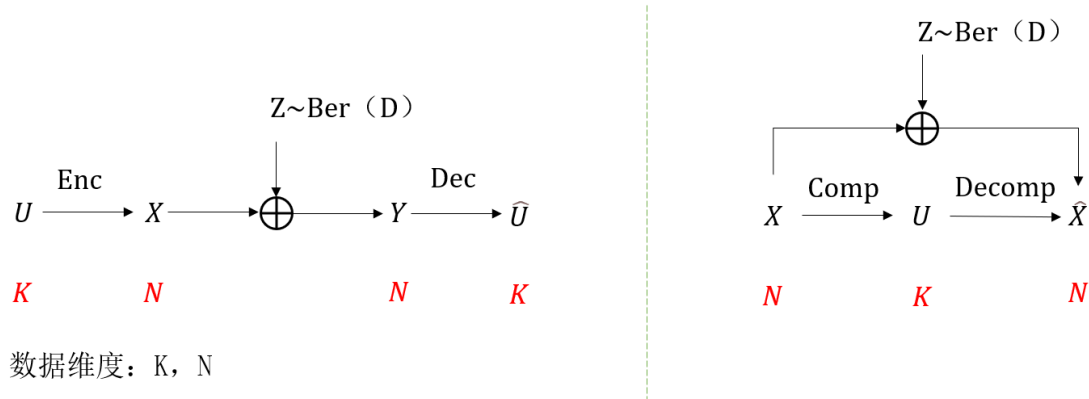


图 3: 信道编码与信源编码

左图对应的是信道编码译码的过程，右图是信源压缩和解压缩的过程，他们之间存在一个对偶的关系^[8]。信道编码是增加冗余，信道译码是减少冗余，恢复信息位的过程。而信源编码是减少冗余的过程，信源译码是增加冗余的过程。那么可以发现，压缩的操作和信道译码端的操作是一样的，都是数据从 N 维到 k 维减少冗余的过程。因此，我们可以使用神经网络译码器来进行数据的压缩操作。比如码长 N 为 16， k 为 8 的神经网络译码器，就可以看作是一个压缩率为 0.5 的一个压缩器。

选择的要压缩的图像是 Lena 的灰度图像，一共有 256×256 个像素点，每个像素的灰度值在 0-255 之间，所以可以使用 8 位二进制表示。把图像相邻的 2 个像素值拼接在一起，组成一个 16 位的 01 比特序列，作为模型的输入，然后让神经网络译码器帮我们完成压缩的操作，从 16 维数据压缩到 8 维，压缩率是 0.5。

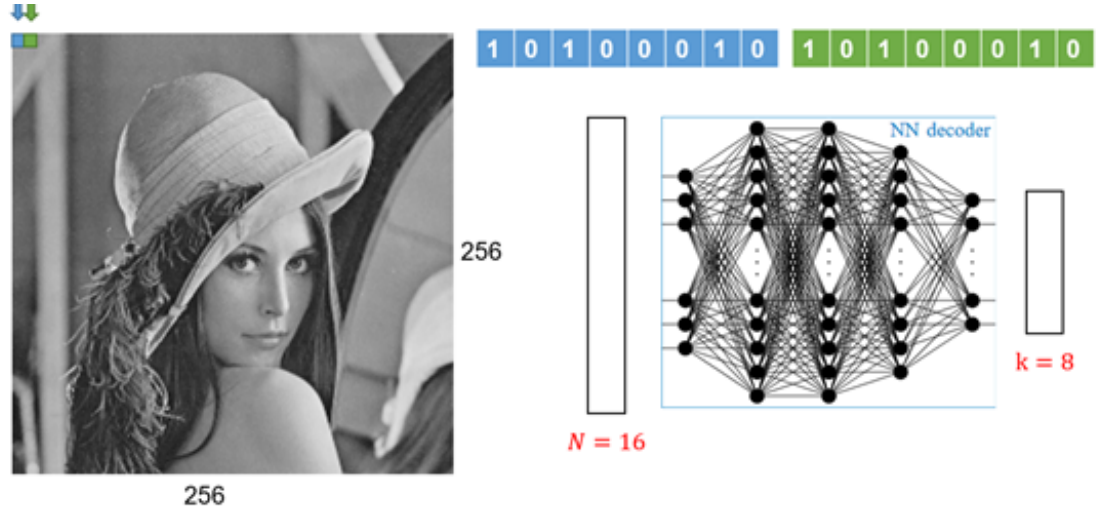


图 4: Lena 灰度图

解压缩对应着信道的编码操作，添加冗余把 k 维数据恢复到 N 维。图片解压缩时，把神经网络输出的信息位和冻结比特位拼起来，乘上极化码的生成矩阵，就恢复到了 N 维的数据，也就完成了像的解压缩操作。

4 复现细节

4.1 与已有开源代码对比

在开源的神经网络模型上，对信道进行了修改，在原文中使用的是 AWGN 加性高斯白噪声信道，修改为 BSC (0.11) 信道，失真率为 0.11。由于信道编码与信源编码之间存在对偶性，因此可以使用神经网络译码器实现图片的压缩操作。利用训练好的神经网络模型以及极化码的编码过程，对图片分别进行压缩和解压缩操作。

4.2 实验环境搭建

Python3.6, Anaconda22.9.0, tensorflow 框架

4.3 创新点

在原文神经网络译码器上，通过信道编码与信源编码之间的对偶关系，利用训练好的神经网络模型以及极化码的编码过程，对图片分别进行压缩和解压缩操作。选择压缩的图像是 Lena 的灰度图像，将像素点灰度值用 8 位二进制表示，把图像相邻的 2 个像素值拼接在一起，组成一个 16 位的 01 比特序列，作为模型的输入，使用训练好的神经网络模型完成压缩的操作。

5 实验结果分析



图 5: lena 原图与解压缩后复原图

压缩失真比较严重，原因是将像素点对应的二进制比特全部进行了压缩，然而每个位置对应的比特的重要性是不一样的。最高位的比特如果解压缩错误，会造成灰度值 128 的误差，因此，可以考虑

压缩低位的比特，高位比特保持不变，以减少误差。

改进方法：

可以将像素点灰度值二进制展开的后四位低位比特进行压缩。将相邻的 4 个像素点灰度值后四位低位比特进行拼接，作为神经网络译码器的输入，进行压缩。

6 总结与展望

在原文的基础上，通过信道编码与信源编码之间的对偶关系，利用训练好的神经网络模型以及极化码的编码过程，对图片分别进行压缩和解压缩操作。

参考文献

- [1] ARIKAN E. Channel Polarization: A Method for Constructing Capacity-Achieving Codes for Symmetric Binary-Input Memoryless Channels[J]. *IEEE Transactions on Information Theory*, 2009, 55(7): 3051-3073. DOI: 10.1109/TIT.2009.2021379.
- [2] GRUBER T, CAMMERER S, HOYDIS J, et al. On deep learning-based channel decoding[J]., 2017: 1-6. DOI: 10.1109/CISS.2017.7926071.
- [3] LUGOSCH L, GROSS W J. Neural offset min-sum decoding[C]// *2017 IEEE International Symposium on Information Theory (ISIT)*. 2017: 1361-1365. DOI: 10.1109/ISIT.2017.8006751.
- [4] XU W, WU Z, UENG Y L, et al. Improved polar decoder based on deep learning[C]// *2017 IEEE International Workshop on Signal Processing Systems (SiPS)*. 2017: 1-6. DOI: 10.1109/SiPS.2017.8109997.
- [5] ELKELESH A, EBADA M, CAMMERER S, et al. Genetic Algorithm-based Polar Code Construction for the AWGN Channel[C]// *SCC 2019; 12th International ITG Conference on Systems, Communications and Coding*. 2019: 1-6. DOI: 10.30420/454862007.
- [6] GAO J, NIU K, DONG C. Learning to Decode Polar Codes With One-Bit Quantizer[J]. *IEEE Access*, 2020, 8: 27210-27217. DOI: 10.1109/ACCESS.2020.2971526.
- [7] HUSSAMI N, KORADA S B, URBANKE R L. Polar Codes for Channel and Source Coding[J]. *ArXiv*, 2009, abs/0901.2370.
- [8] ARIKAN E. Source polarization[J]., 2010: 899-903. DOI: 10.1109/ISIT.2010.5513567.
- [9] KORADA S B, URBANKE R L. Polar Codes are Optimal for Lossy Source Coding[J]. *IEEE Transactions on Information Theory*, 2010, 56(4): 1751-1768. DOI: 10.1109/TIT.2010.2040961.