Learned Image Compression

EE 276, Information Theory | Project Proposal

Group Members

Yanpei Tian, yanpeit@stanford.edu
Helena Huang, yhuang77@stanford.edu
Boxiao Pan, bxpan@stanford.edu

Problem Motivation and Definition

Image compression intends to reduce the size of an image so as to save the cost during transmission as well as storage. There are mainly two types of image compression, namely lossless compression and lossy compression. Traditional codecs include PNG (lossless) and JPEG (lossy). Modern approaches apply deep learning on this task, i.e. learned image compression. The general pipeline includes an autoencoder architecture, where an encoder first compresses the input image into a compact representation, which is subsequently fed to a decoder to reconstruct the input image. The learning objective is thus to minimize the difference between the reconstructed image and the original image. In this project, we aim at exploring different ways to improve this framework. One idea we currently have is to insert an additional entropy encoder-decoder dual in between the original encoder-decoder bottleneck. The intuition is to leverage the redundant information of the compressed image so as to further reduce the size of the compressed image from the autoencoder.

Literature Review

Full Resolution Image Compression with Recurrent Neural Networks [6]: This project is built on top of Variable Rate Image Compression With Recurrent Neural Networks [2], which shows that it is possible to train a single RNN and achieve better than current image compression schemes at a fixed output size. The authors incorporate the idea of “entropy encoder” into this project to make it achieving competitive compression rates on images of arbitrary size.

Lossy image compression with compressive autoencoders [3]: In this paper, the author introduced an effective way of dealing with non-differentiability in training autoencoders for lossy image compression. They proposed three alternatives to the un-differentiable quantization step: 1. Replace gradient of quantization by the gradient of identity function, effectively passing gradients without modification from the decoder to the encoder; 2. Additive uniform noise; 3. Stochastic rounding operations;

Entropy Encoding in Wavelet Image Compression [4]: The author reviews several entropy encoding schemes such as Huffman coding and arithmetic coding, providing detailed mathematical bases for the encoding schemes and their applications.

Dataset

We propose to use the dataset provided by “WORKSHOP AND CHALLENGE ON LEARNED IMAGE COMPRESSION” (http://www.compression.cc), a workshop aims to gather publications which will advance the field of image and video compression using state of the art machine learning and computer vision techniques.

Specifically, we focus on the low-rate compression track and use its provided Dataset-P and Dataset-M. Dataset-P consists of 1.9G training and 129M validation images, while Dataset-M includes 3.8G training and 226M validation images.

Implementation

The training procedure is shown in Fig. 1. Specifically, the loss consists of two components: the reconstruction loss between the original and reconstructed image, as well as the regularization loss which penalizes the size of the compressed image from being too large.

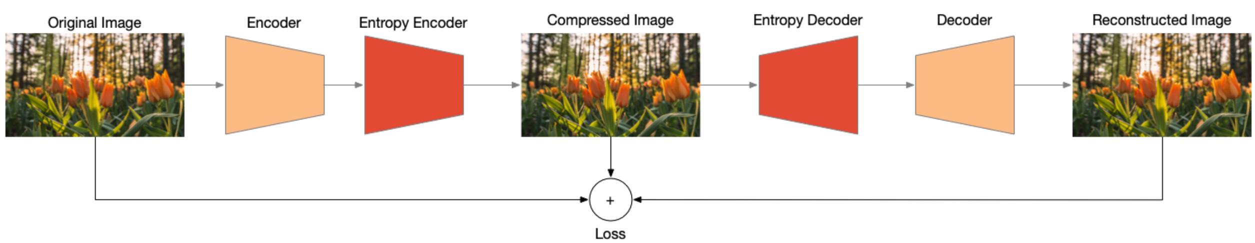

Figure 1. Illustration of the training procedure.

Evaluation

We will evaluate the performance of our project based on these three metrics: compression ratio, loss function, and human perceptual rating.

Compression Ratio: Represented as bits per pixel (bpp), the number of bits needed in the compressed image to represent 1 pixel in the original image;

Loss Function: As far as our literature review goes, there are no such metric capable of correlating with human raters for all types of distortions. Current state-of-the-art models utilize L_p norms as the optimization metric.

Human Perceptual Rating: This is an easy one, the ultimate goal for all image compression schemes is to create a faithful but smaller representation of the original image. So the reconstruction of the compressed image needs to be similar to the original image according to human perception.

References

[1] "Workshop and Challenge on Learned Image Compression." *Workshop and Challenge on Learned Image Compression*, www.compression.cc/

[2] G. Toderici, S. M. O'Malley, S. J. Hwang, D. Vincent, D. Minnen, S. Baluja, M. Covell, and R. Sukthankar. Variable rate image compression with recurrent neural networks. ICLR 2016, 2016.

[3] Lucas Theis, Wenzhe Shi, Andrew Cunningham, and Ferenc Huszár. Lossy image compression with compressive autoencoders. arXiv preprint arXiv:1703.00395, 2017.

[4] Song, M.S.: Entropy encoding in wavelet image compression. In: Representations, Wavelets and Frames A Celebration of the Mathematical Work of Lawrence Baggett, pp. 293–311 (2007)

[5] Mwiti, Derrick. "A 2019 Guide to Deep Learning-Based Image Compression." *Medium*, Heartbeat, 25 Oct. 2019, heartbeat.fritz.ai/a-2019-guide-to-deep-learning-based-image-compression-2f5253b4d811.

[6] G. Toderici, D. Vincent, N. Johnston, S. J. Hwang, D. Minnen, J. Shor, and M. Covell. Full Resolution Image Compression with Recurrent Neural Networks. arXiv preprint arXiv:1608.05148, 2016.