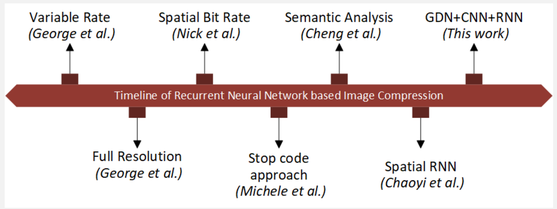
Image Compression with Recurrent Neural Network and Generalized Divisive Normalization

Khawar Islam[1], Dang Lien Minh[1], Sujin Lee[2] and Hyeonjoon Moon[1,2]
Computer Vision and Pattern Recognition Lab[1], Department of Artificial Intelligence[2]
Sejong University, South Korea

History and Timeline:

- The development of image compression networks based on RNN is relatively small compare to CNN and auto-encoders.
- These techniques proposed a full image resolution network using residual scaling, recurrent neural network (RNN), and entropy coding based on deep learning. This network simultaneously generated three models during training.

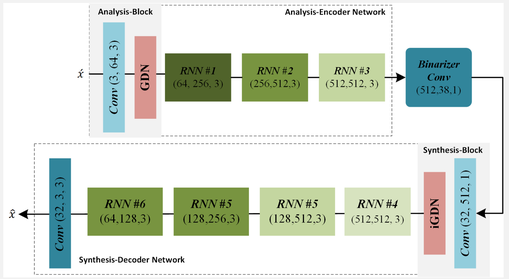


Introduction:

- In this paper, two effective novel blocks are developed: analysis and. synthesis block that employs the convolution layer and GDN in the variable-rate encoder and decoder side.
- Our network utilizes a pixel RNN approach for quantization. Furthermore, to improve the whole network, we encode a residual image using LSTM cells to reduce unnecessary information.
- Experimental results demonstrated that the proposed variable-rate framework with novel blocks outperforms existing methods and standard image codecs, such as George's and JPEG in terms of image similarity.

Architecture:

- Iterative architecture of image compression framework based on recurrent neural network. Each input patch was first passed to the analysis-encoder block to enrich image representation.
- Similarly, the synthesis-decoder block reconstructs a de-coded image with the help of recurrent neural network cells.

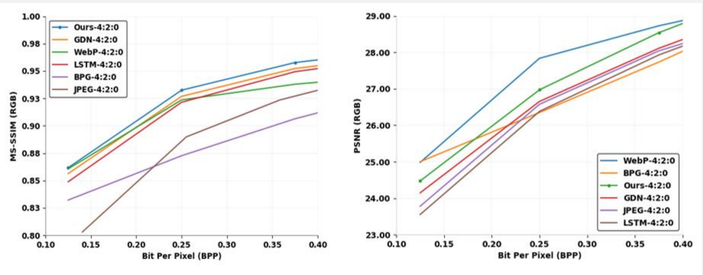


Entropy Coding:

- The approach of pixel RNN as a binary RNN for image compression with a single convolution layer. The quantization approach describes in and applies quantization noise during training.
- The binarizer part comprises a single linear convolutional layer with an activation function. The entire binarizer module generates binary codes within an interval of (1,-1) with the help of the sign function.

Results:

- We visualize the performance of our reconstructed images in each iteration with our variable-rate model. The comparison results of our proposed network and other image codecs are presented in below figure.
- The results are compared with the original image, George's [1] (4:2:0). As illustrated in below fig, the proposed network with blocks gives a high similarity compare to the previous methods [1].
- The visual performance of JPEG and JPEG2000 has worst because the edges contain ringing artifacts. The results of HEVC based BPG are clearer, and smoother compare with JPEG. Compare with George's [1] and JPEG in terms of SSIM. The PSNR score is a little lower than BPG and WebP (in sampling factor 4:2:0 format). The proposed variable-rate method achieves better MS-SSIM, especially at the 10 epochs.



References:

1. G. Toderici, D. Vincent, N. Johnston, S. Jin Hwang, D. Min-nen, J. Shor, and M. Covell. Full resolution image compres-sion with recurrent neural networks. In Proceedings of theIEEE Conference on Computer Vision and Pattern Recogni-tion, pages 5306–5314, 2017.

2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) At Nashville, Tennessee, United States