GENERATIVE COMPRESSION

Vineeth S

M.Tech Artificial Intelligence, SR No. 16543

ABSTRACT

In this work, we explore the use of generative models for image or video compression. We know that generative models such as GANs and VAE can reproduce an image from its latent vector. We ask the question whether we can go the other direction — from an image to a latent vector. We show that we can compress images with a high compression rate.

***Index Terms*—** image compression, generative networks

1. INTRODUCTION

In the modern world, a huge amount of data is being sent across the internet every day. Efficient data compression and communication protocols are of great research interest today. We focus on the compression of images and video (sequence of image frames) using deep generative models and show that they achieve a better performance in compression ratio and perceptual quality. We explore the use of GAN[1] and VAE-GAN[2] for this task. Research, though limited, has shown that these types of methods are quite effective and efficient[3][4].

2. TECHNICAL DETAILS

In our first approach we show that given an image and a pre-trained GAN on the same domain, we can find the latent vector that produces the image. In our second approach we propose a VAE-GAN architecture that achieves high rate of compression when trained of a sufficiently large dataset (we use internet scrapped images and Flicker30K Dataset).

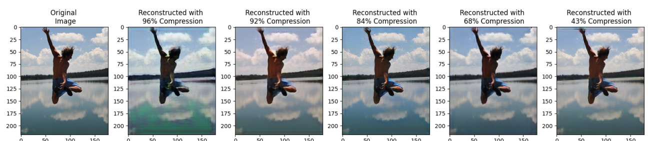
2.1. GAN Compression

In this approach ,we use a pre-trained PGAN[5] from Facebook trained on CelebA-HQ dataset. We freeze the network parameters throughout this approach. We provide an image and use gradient descent to obtain the latent vector. We further employ lossy compression [6] and use Elias coding to reduce the vector size further. Figure 2.1 compares the image quality of reconstruction for a sample image.

Original(500KB) JPEG(5.2KB) GAN(2.8KB) Lossy (1KB)

2.2. VAE-GAN Compression

We propose a novel VAE-GAN architecture for learned image compression. We train the VAE to reconstruct an image and train the discriminator to classify the image as original or reconstructed simultaneously. We provide the (image, encoded image) pair as inputs to the discriminator to classify. The number of channels in the last convolutional layer of the encoder characterizes the compression rate. Figure 2.2 compares the image quality of reconstruction for a sample image.



3. RESULTS

We evaluate the models on the Structural Similarity Index (SSIM) and Peak Signal to Noise Ratio (PSNR) between the original image and reconstructed image. Though VAE-GAN seems to have lower compression ratio (CR), they practically perform better across different domains (unlike GANs).

Table 1: Approach 1: GAN Compression

Compression method	SSIM	PSNR	CR
GAN	0.79	26.48	176 ×
Lossy compression 8bits	0.77	25.06	412.5 ×
Lossy compression 6bits	0.67	25.06	495 ×
Lossy compression 4bits	0.46	25.06	559 ×
JPEG Quality 1%	0.51	19.96	99 ×

Table 2: Approach 2: VAE-GAN Compression

Number of Channel	SSIM	PSNR	CR
28 Channels	0.83	24.79	1.74 ×
16 Channels	0.80	24.37	3.06 ×
8 Channels	0.81	24.06	6.12 ×
4 Channels	0.79	23.57	12.24 ×
2 Channels	0.71	21.57	24.49 ×

4. CONTRIBUTIONS

We implement GAN Compression and VAE-GAN Compression from scratch. Note that more details and images are not shown here due to lack of space and are available at the GitHub repositories here and here.

5. RESOURCES

Dataset: Web Scrapped images

Dataset: Flicker30K http://bryanplummer.com/Flickr30kEntities/

PGAN(Approach 1) : Facebook Research https://github.com/facebookresearch/pytorch_GAN_zoo

Toolkits: PyTorch, Torchvision, Scikit-image, PIL Image

References

[1] Ian J. Goodfellow et al. *Generative Adversarial Networks*. 2014. arXiv: 1406.2661 [stat.ML].

[2] Anders Boesen Lindbo Larsen et al. *Autoencoding beyond pixels using a learned similarity metric*. 2016. arXiv: 1512.09300 [cs.LG].

[3] Shibani Santurkar, David Budden, and Nir Shavit. *Generative Compression*. 2017. arXiv: 1703.01467 [cs.CV].

[4] Eirikur Agustsson et al. *Generative Adversarial Networks for Extreme Learned Image Compression*. 2019. arXiv: 1804.02958 [cs.CV].

[5] Tero Karras et al. *Progressive Growing of GANs for Improved Quality, Stability, and Variation*. 2018. arXiv: 1710.10196 [cs.NE].

[6] Dan Alistarh et al. *QSGD: Communication-Efficient SGD via Gradient Quantization and Encoding*. 2017. arXiv: 1610.02132 [cs.LG].