Generative Neural Network Based Image Compression

E. Meltem Tolunay, Ahmad Ghalayini
Email: {meltem.tolunay, ahmad2}@stanford.edu

Stanford | ENGINEERING
Computer Science

CS 229: Machine Learning

INTRODUCTION

MOTIVATION

Standard lossy image compression techniques such as *JPEG*:

- are **not** data-specific → do **not** make use of the semantic relations among the images in a specific dataset
- thus, cannot achieve the best possible compression rates

OUR SUGGESTED APPROACH

1. Train a GAN on a dataset, capturing the semantic relations in that dataset
2. **Gan Reversal:** recover the latent space representation of an image from the GAN generator **[1]**
   - *Experiment with different loss functions (L1, L2, SSIM)*
3. This latent vector will be the compressed representation
4. To decode image, pass latent vector through GAN generator

OUR AIM

A ***better extreme-compression scheme*** **with two main objectives:**

1. scheme must achieve higher compression rates than other standard lossy image compression techniques
2. reconstructed images must still be of high **perceptual** quality and true to their originals → use **suitable metrics** to measure that

DATA

- Our model is trained on the well-known **CelebA benchmark dataset [2]**
  - Consists of > **200K** celebrity images
- Our test set consists of 10 images from CelebA
  - Our approach is not general yet, so we must manually adjust parameters → test set small, **but promising results**
- We crop and center the images to **128x128** for training and testing
- **(Left) An image that the GAN outputs for some latent vector**
  (Right) An uncompressed image that is not a GAN output

GAN OUTPUT

NEW IMAGE

MODELS

BASELINE MODELS

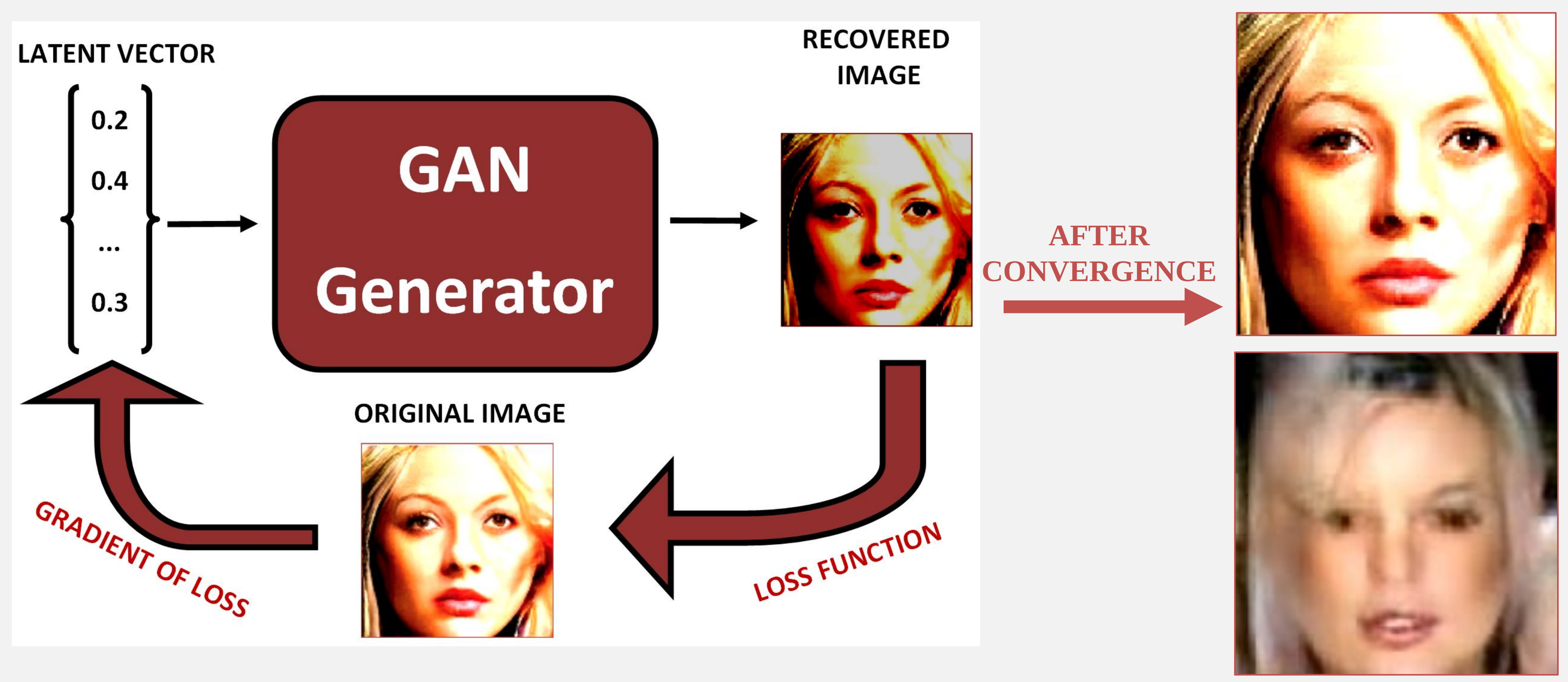
- **KNN:** *K=4***,** cannot achieve extreme compression while maintaining good quality
- **JPEG (optimized):** a popular lossy image compression technique
  - We use two values for *quality* parameter: *1% and 10%*

KNN (K=4)

JPEG: 10%, 1%

OUR MODEL: GAN REVERSAL

- First train a GAN on the dataset → captures the semantic relations in the dataset
- Using Gradient Descent *(with some modification)*, find a latent vector which when passed through the trained GAN Generator, is closest to the original image in terms of some loss function (we try L1, L2, and **SSIM**)



GAN REVERSAL

METRICS

- **Compression Magnitude Measure**: Bits per Pixel (*BPP*)
- **Traditional Similarity Measures:** Mean Square Error (*MSE*), Peak Signal to Noise Ratio (*PSNR*)
- **Perceptual Quality Measure:** Structural Similarity Index (*SSIM*) **[3]**
  - SSIM is a perception-based model that considers image degradation as *perceived change in structural information*

$$\mathrm{SSIM}(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

RESULTS

- Below is a summary of the performance of the baselines and our approach "GAN Reversal", based on the proposed metrics

Method	BPP	PSNR	MSE	SSIM
KNN (K=4)	1.486	23.012	342.46	0.7145
JPEG (10%)	0.4848	26.161	161.87	0.7711
JPEG (1%)	0.2176	20.465	589.15	0.5280
GAN Reversal (Our Approach)	**0.2152**	21.192	540.06	**0.7073**

CONCLUSION AND FUTURE WORK

- The main contributions of our project are:
  - Introducing GAN Reversal as a novel tool for image compression
  - Using *SSIM* as a custom loss function that yields much better performance when recovering the latent vector
- Our results indicate that using GAN reversal, we can perform extreme compression while maintaining acceptable perceptual quality compared to other approaches like *JPEG*
- **NEXT STEPS**
  - Try other loss functions when recovering latent vector
  - Try to make latent vector distribution better follow the uniform distribution to improve GAN output
  - Increase test set size

REFERENCES

[1] Zachary C. Lipton and Subarna Tripathi. Precise recovery of latent vectors from generative adversarial networks. CoRR, abs/1702.04782, 2017. URL http://arxiv.org/abs/1702.04782.

[2] Ziwei Liu, Ping Luo, Xiaogang Wang, and Xiaoou Tang. Deep learning face attributes in the wild. In Proceedings of International Conference on Computer Vision (ICCV), December 2015.

[3] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. IEEE transactions on image processing, 2004.