A Study of GANs

(Generative Adversarial Networks)

by Stephen E. Worley

Outline

- Overview
- Discriminative vs Generative
- Architecture
- Adversarial
- Issues
- Use Cases

Overview

- Introduced in a paper by Ian Goodfellow and other researchers at University Montreal in 2014.
- Facebook AI research director Yann LeCun called it, “the most interesting idea in the last 10 years in machine learning.”
- GANs can be used to mimic any distribution of data:
    - Given a data set, it can learn to produce new data points in line with the dataset.
    - Images, music, speech.
    - Almost like creators in their own right

Discriminative Algorithms

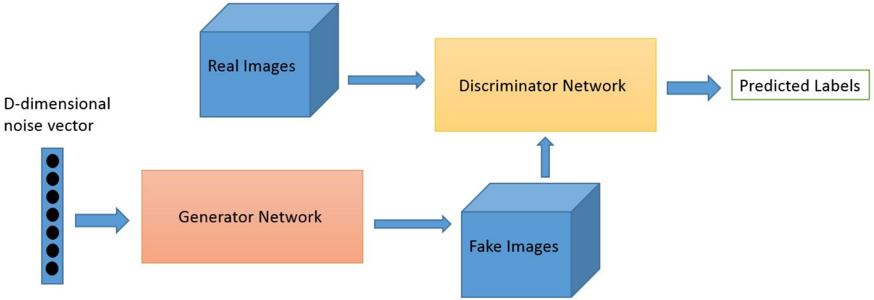
- Classify input data
- Given the features/attributes of a data instance, predict the class label or category of that instance
    - Ex: Spam Email Detector
- Mathematically, $p(y|x)$ where y is the label and x are the features
- Mapping features -> labels
- Learn the boundaries between classes

Generative Algorithms

- Opposite of Discriminative algorithms
- Given the class label or category, predict features/attributes of a data instance
    - Ex: Assuming the email is spam, how likely are these features?
- Mathematically, $p(x|y)$, the probability of features given a class
- Model the distribution of individual classes

They can also be used to classify data as well!

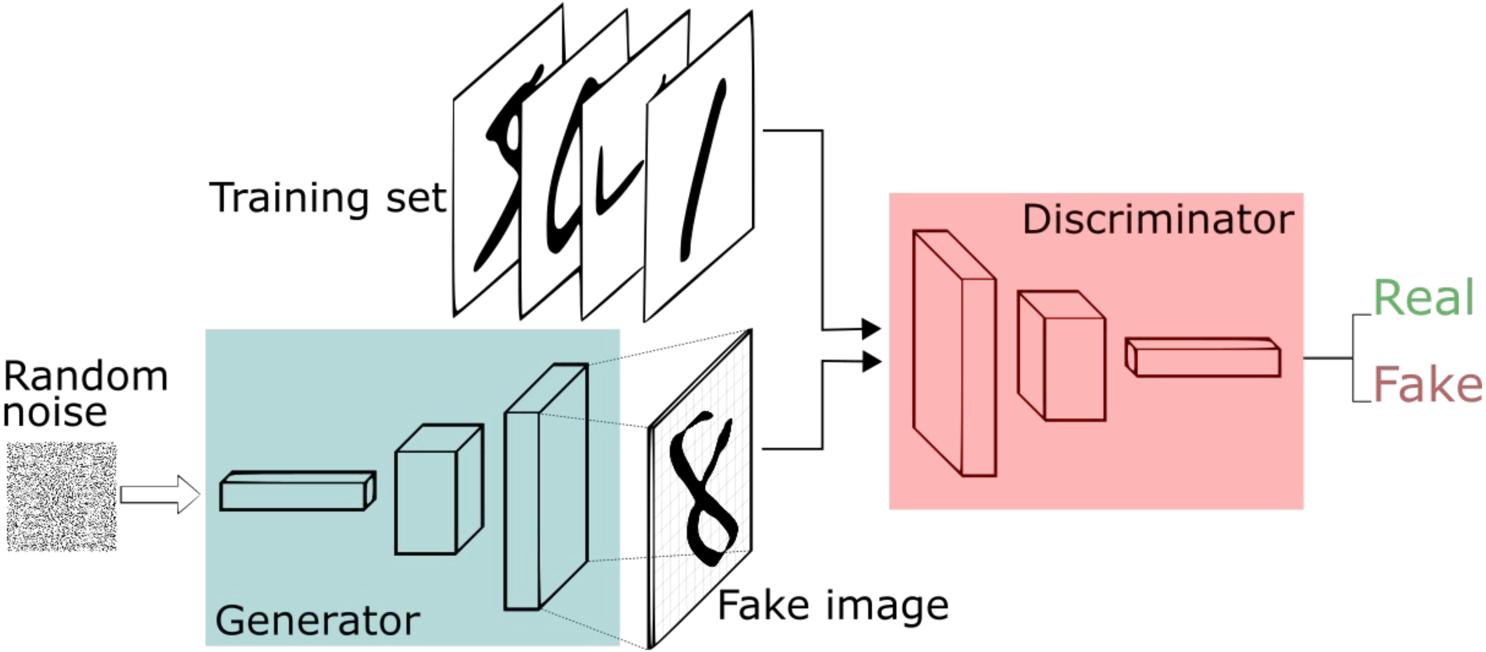
Architecture of GANs



- Two neural networks:
  - Generator: generates new data instances
  - Discriminator: evaluates data for authenticity
- Steps:
  - Generator takes in random numbers (noise) and returns an image
  - Generated image is fed into discriminator alongside a stream of images taken from the actual dataset
  - The discriminator takes in both real and fake images and returns probabilities [0,1] 1=authentic, 0=fake.
- Produces double feedback loop:
  - Discriminator in feedback loop with the ground truth of images (known class labels = authenticity)
  - Generator is in a feedback loop with the discriminator (whether it tricked it or not)

Adversarial

- Both nets trying to optimize a different and opposing objective function (loss function) in a zero-sum game.
    - The win of one is the loss of the other.
- As the generator changes its behavior, so does the discriminator, and vice versa.
- GANs can be thought of as the combination of a counterfeiter and a cop in a game of cat and mouse
    - Counterfeiter is learning to pass false notes
    - Cop is learning to detect them
    - Both are dynamic:
        - Central bank is flagging the bills that slipped through to train the cop
        - Each side comes to learn the other's methods in a constant escalation


Training set
Discriminator
Real
Fake
Random noise
Generator
Fake image

Issues

- Each side of the GAN can overpower the other during training.
  - If the discriminator is too good, it will return values so close to 0 or 1 that the generator will struggle
  - If the generator is too good, it will persistently exploit weaknesses in the discriminator that lead to false negatives.
  - Fix these by adjusting respective learning rates
- Training time is LONG.
  - GPU = hours, CPU = days
- Generative model can trick Discriminator but not humans
- Generative model tends to look very similar to original data set
  - Proposed that this can be fixed with very large datasets

Use Cases

- Predicting next frame in a video
- Increasing resolution of an image
- Text to image generation
- Interactive image generation
- Image to image translation
    - Facebook is using this research area to map faces to other pictures

Increase Image Resolution

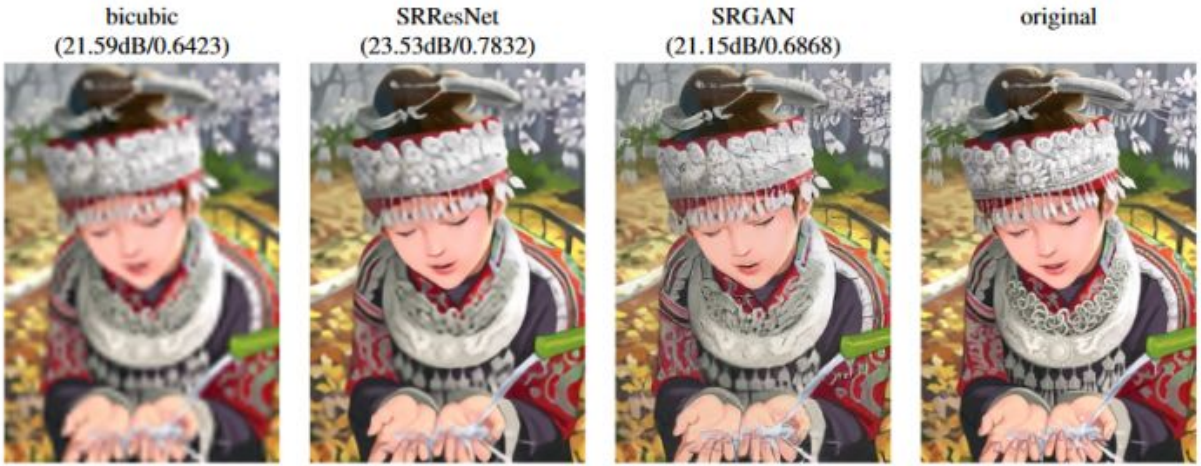


Figure 2: From left to right: bicubic interpolation, deep residual network optimized for MSE, deep residual generative adversarial network optimized for a loss more sensitive to human perception, original HR image. Corresponding PSNR and SSIM are shown in brackets. [4× upscaling]

Questions?

References

https://www.analyticsvidhya.com/blog/2017/06/introductory-generative-adversarial-networks-gans/

https://deeplearning4j.org/generative-adversarial-network

https://arxiv.org/pdf/1609.04802.pdf

https://arxiv.org/pdf/1511.06380.pdf

https://github.com/junyanz/iGAN

https://affinelayer.com/pixsrv/