Sketch to Face Transformation for Criminal Investigation

Guide - Mr. Vaibhav Ambhire
Group No - 42
Group Members - Aryan Nayak, Rachaell Nihalaani, Sparsh Nagpal

Introduction

- Criminal Investigation has recently advanced it's technical capabilities and methodologies. Criminal identification using face sketches is still an implementation that requires good human visual and memory skills.

- Recognising actual human faces from those sketches become really difficult.

- The implementation of computer vision into this will reduce the human effort to relate a black and white drawing to actual faces by means of image translation.

- Our model will utilize the Generative Adversarial Networks (GANs)

Problem Statement

To create a model that takes an artistic face-sketch of a person as the input image file, enhances certain features of it and transforms it into a realistic (or as close as possible) photograph of the face of that person. Train the dataset over a sizeable dataset for better results.

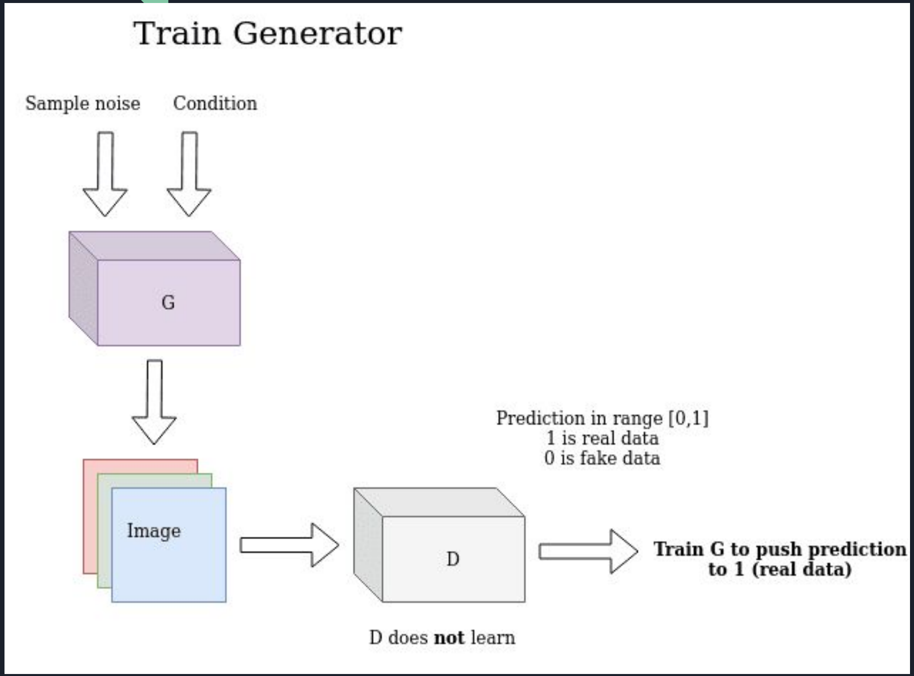
General Methodology

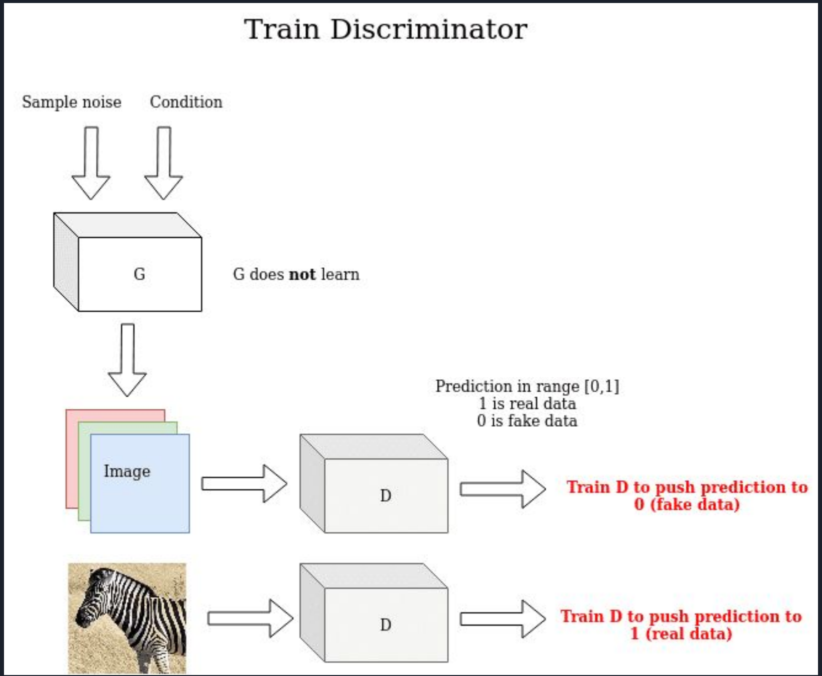
- What is GAN ?
  - The GAN model architecture involves two sub-models: a *generator model* for generating new examples and a *discriminator model* for classifying whether generated examples are real, from the domain, or fake, generated by the generator model.
    - **Generator**. Model that is used to generate new plausible examples from the problem domain.
    - **Discriminator**. Model that is used to classify examples as real (*from the domain*) or fake (*generated*).

General Methodology

- What is GAN ?
  - Generative Adversarial Networks
  - are an approach to generative modeling using deep learning methods, such as convolutional neural networks.
  - Generative modeling is an unsupervised learning task in machine learning that involves automatically discovering and learning the regularities or patterns in input data in such a way that the model can be used to generate or output new examples that plausibly could have been drawn from the original dataset.

Vanilla GANs architecture





Existing Solutions & Drawbacks

- **AutoEncoders**
- **Pix2Pix using conditional GANs**
- **DCGANs**
- **CycleGANs**

Existing Solutions & Drawbacks

- Common approaches used now in conditional generative adversarial networks (cGAN) incorporate hard conditions like pixel-wise correspondence alongside the translation process, which makes the output strictly align with the input edges. This can be highly problematic in sketch-to-image generation when the input is a free-hand sketch.
- On the other hand, the translation should respect the sparse input content, but might need some deviation in shape to generate a realistic image.

Proposed System - Contextual GANs

- To tackle the challenges in the existing solutions, we are using a novel contextual generative adversarial network for sketch-to-image generation. We pose the image generation problem as an image completion problem, with sketch providing a weak contextual constraint.
- In conventional image completion, the corrupted part of an input image is completed using surrounding image content as context.
- A generative adversarial network is trained to learn the joint distribution and capture the inherent correspondence between a sketch and its corresponding image using the defined joint image.

Evaluation

There are two types of evaluations: Qualitative and Quantitative [Link] [Paper]

Qualitative - Qualitative measures are those measures that are not numerical and often involve human subjective evaluation or evaluation via comparison.

1. Nearest Neighbors.
2. Rapid Scene Categorization.
3. Rating and Preference Judgment.
4. Evaluating Mode Drop and Mode Collapse.
5. Investigating and Visualizing the Internals of Networks.

Evaluation

Quantitative- Quantitative GAN generator evaluation refers to the calculation of specific numerical scores used to summarize the quality of generated images.

1. Average Log-likelihood
2. Coverage Metric
3. Inception Score (IS)
4. Modified Inception Score (m-IS)
5. Mode Score
6. AM Score
7. Frechet Inception Distance (FID)
8. Maximum Mean Discrepancy (MMD)
9. The Wasserstein Critic
10. Birthday Paradox Test
11. Classifier Two-sample Tests (C2ST)
12. Classification Performance

1. Boundary Distortion
2. Number of Statistically-Different Bins (NDB)
3. Image Retrieval Performance
4. Generative Adversarial Metric (GAM)
5. Tournament Win Rate and Skill Rating
6. Normalized Relative Discriminative Score (NRDS)
7. Adversarial Accuracy and Adversarial Divergence
8. Geometry Score
9. Reconstruction Error
10. Image Quality Measures (SSIM, PSNR and Sharpness Difference)
11. Low-level Image Statistics
12. Precision, Recall and F1 Score

Evaluation

- Perhaps the most used qualitative GAN generator model is an extension of the manual inspection
- Two widely used quantitative measures are
  - Inception Score (classification into classes)
  - Frechet Inception Distance (Vision specific features)

These measures capture the quality and diversity of generated images, both alone (former) and compared to real images (latter) and are widely

Database

CUHK Face Sketch FERET Database (CUFSF) in Mar., 2011. [Link]

CUHK Face Sketch database (CUFS) is for research on face sketch synthesis and face sketch recognition. It includes 188 faces from the Chinese University of Hong Kong (CUHK) student database, 123 faces from the AR database [1], and 295 faces from the XM2VTS database [2]. There are 606 faces in total. For each face, there is a sketch drawn by an artist based on a photo taken in a frontal pose, under normal lighting condition, and with a neutral expression.

Database

More Face Images. Could be augmented to black and white sketch like Images for higher data.

- Large-scale CelebFaces Attributes
- Labeled Faces In The Wild
- CUHK Face Sketch FERET Database
- XM2VTS data set
- SCface - Surveillance Cameras Face Database
- Face Recognition Databases
- adiencedb - homepage

Design

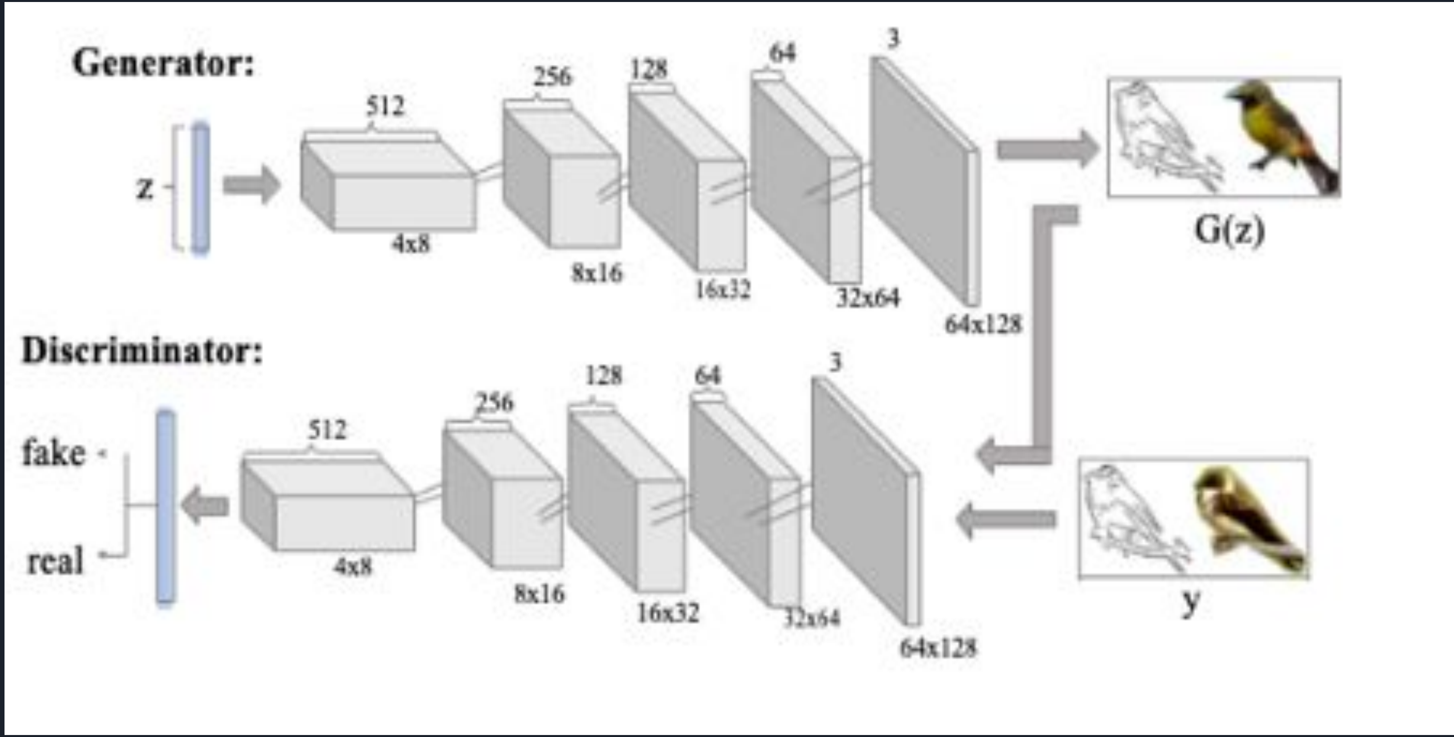


Figure 4.1A : Generator Discriminator Architecture for Contextual GANs

Design

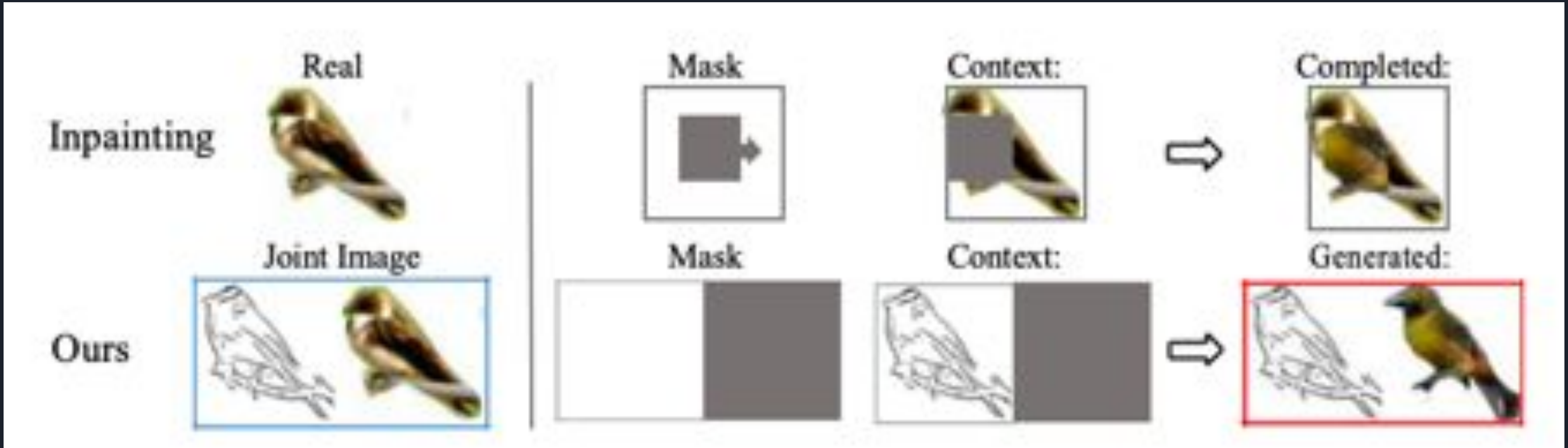


Figure 4.1B : Contextual GANs impainting methodology

Hardware and Software Requirements

- Python 3.7
- HTML/CSS/JavaScript/Flask - for Web Design
- Web Browser: Safari, Google Chrome or others
- Operating System: Windows 10 , MacOS

Functional Requirement

- Input must be an image of a sketch of a person's face.
- System must perform transformation of the input sketch to an output face image.
- Output must be a corresponding face image of the input sketch image.

Non-Functional Requirement

- Input must be in the form of one of these image file formats - .jpg, .jpeg or .png
- Transformation of the sketch image to the face image should be done within 30 seconds.
- The site should load in 2 seconds and be compatible with all web browsers such as Google Chrome, Firefox, Microsoft Edge etc.
- Output face image must be downloadable.

Plan for Implementation

- Dataset augmentation from the limited CUHK dataset.
- The GAN model created with a defined Generator model and a Discriminator model.
- Create a loss function to train the model.
- Load the dataset and train the model.
- Save the trained model as h5 files to have a real time implication.
- The model is implemented and tested.
- We measure the image quality using evaluation metrics.

Results

- We were able to generate real-life-like images through image translation method using contextual GANs. To measure the performance, we used metrics like SSIM score and L2-normalization score. SSIM score was tested as an approximate of 0.77 and L2-normalization score as 93.

Future Work

- The CUHK Face Sketch Dataset includes images of students of the Chinese University of Hong Kong and so we can expect most of them to be Chinese. As a result, our model is well trained to recognize and predict Chinese facial features, and provides accurate images from the given sketch.
- Other than this, training the model on a much larger dataset should also improve the model accuracy.
- Hardware requirements like GPU will also be needed with larger dataset for training

References

[1] Ali Borji, Pros and Cons of **GAN Evaluation** Measures, p 5

[2] Yongyi Lu, Shangzhe Wu, Yu-Wing Tai, Chi-Keung Tang., Image Generation from Sketch Constraint Using Contextual GAN. In ECCV (2018)

[3] JYeh, R.A., Chen, C., Lim, T.Y., Schwing, A.G., Hasegawa-Johnson, M., Do, M.N.: Semantic image inpainting with deep generative models. In: CVPR (2017)

[4] Pathak, D., Kr¨ahenb¨uhl, P., Donahue, J., Darrell, T., Efros, A.: Context encoders: Feature learning by inpainting. In: CVPR (2016)