10-Final Presentation

COSMOP-Project-1

BY: Xiaoqing Liu, Subin Varghese, Sunita Villarreal

Mentor: Athina Bikaki

Objective

- Generate various quasi-realistic, using Generative Adversarial Networks, immigration images from a smaller dataset of real ones
- Criteria:
  - People carrying bags
  - People with children
  - People crowded around at the border
  - People with border patrol

Workflow

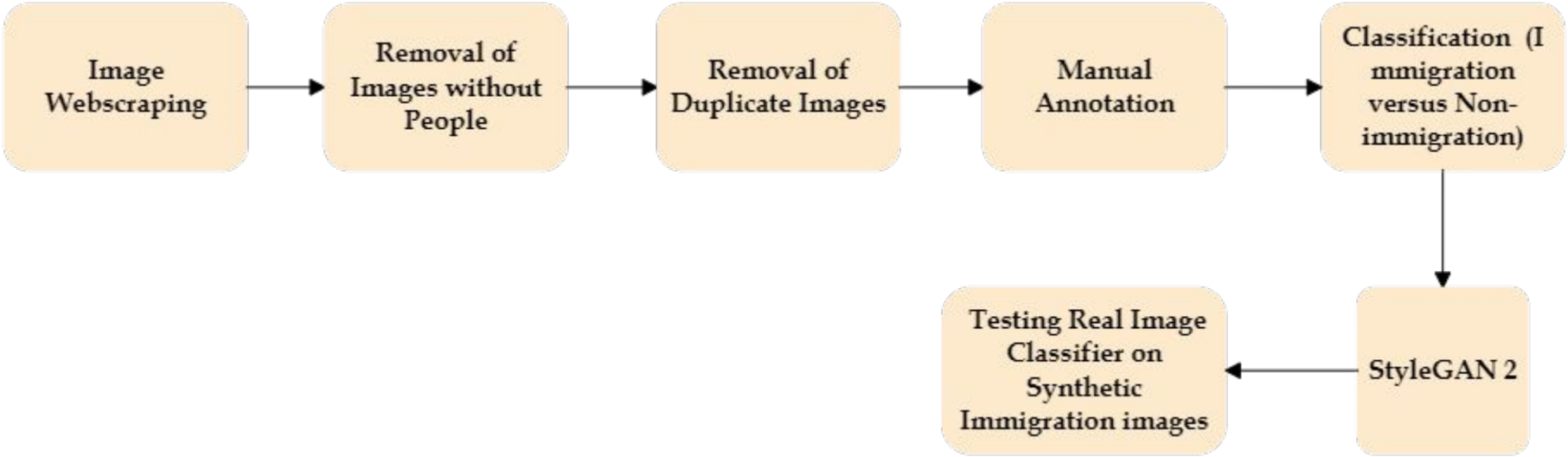
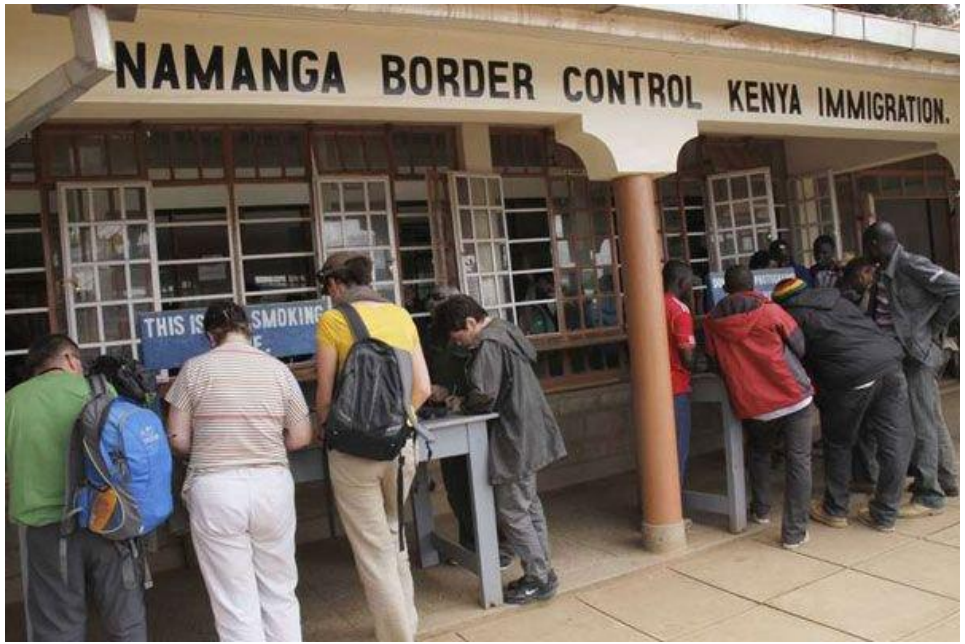


Image Webscraping
Removal of Images without People
Removal of Duplicate Images
Manual Annotation
Classification (Immigration versus Non-immigration)
StyleGAN 2
Testing Real Image Classifier on Synthetic Immigration images

People carrying bags

People with children

People crowded around at the border



People with border patrol

Image Webscraping

- Webscraping using
  - Selenium webdriver
    - Keywords:
      - “[country] + immigration border” [1]
  - Twitter API
    - Keywords and hashtags: “ukraine refugees”, “ukraine”, etc., based on word clouds [2]

[1] https://en.wikipedia.org/wiki/List_of_countries_and_territories_by_land_borders

[2] Athina

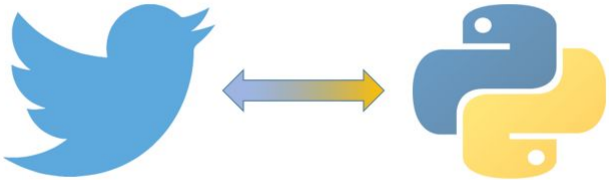
Image webscraping using Selenium

1. Use Selenium to open up a web browser, e.g., Chrome
2. Search for query of interest
3. Move into the images section
    a. Get image links
4. Save images into folder of choice

https://medium.com/@wwwanandsuresh/web-scraping-images-from-google-9084545808a2

Image webscraping using the Twitter API

1. Register own 'app' in Twitter so that we can have our own credentials (consumer key and secret)
2. Access the Twitter API using the Tweepy package
3. Search for tweets with the specified keywords/hashtags
4. Obtain the full path for the images and download them into folder of choice



https://miguelmalvarez.com/2015/03/03/download-the-pictures-from-a-twitter-feed-using-python/

Problems encountered

- More irrelevant images than relevant ones
    - Dependence on hashtags
- Added only about 500 more images from Twitter

Pre-processing our dataset

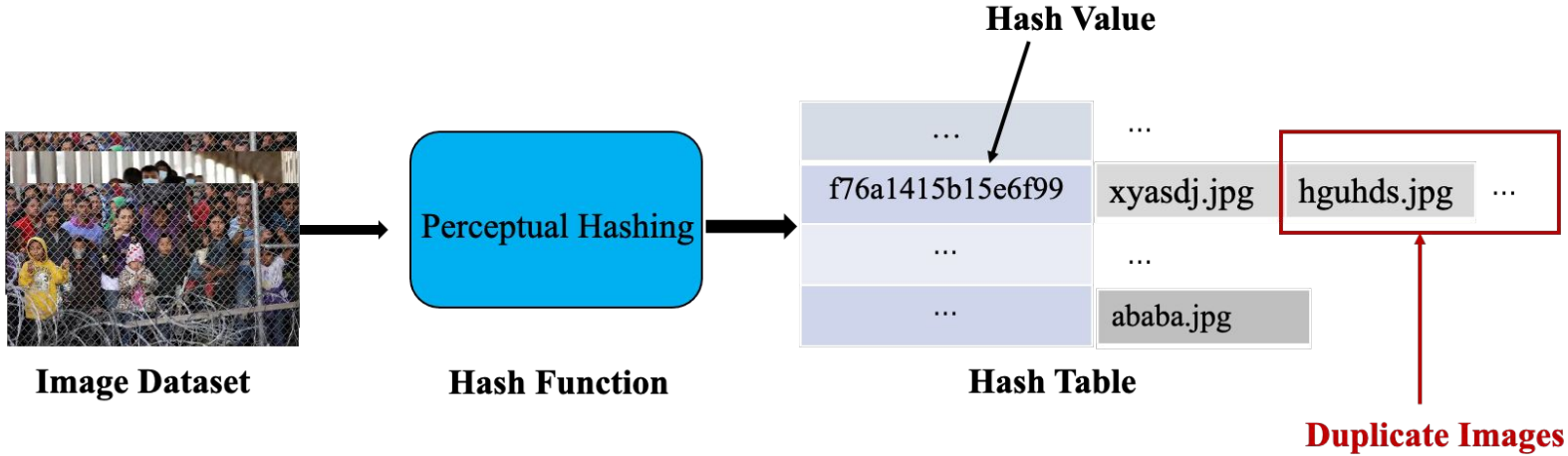
- Pre-processing:
  - Remove duplicates
  - Remove images without people
  - Get only relevant images
  - > 2200 results

Annotation

- Using a classification
  - Test accuracy of 80%
- Decided on manual annotation

Real Images

Remove Duplicate Images From Dataset



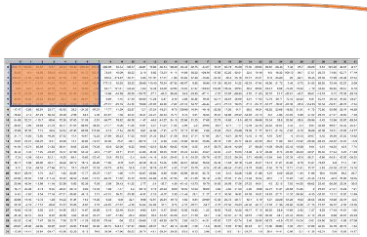
https://backend-live-coc.cfr.org/cdn/ff/BOg_dKOSmNuEIFzJF5pDXyubS8SGjNPFRcSWanThunU/1578015356/public/publications/global-memos/RTS2F7RC_1.jpg

Remove Duplicate Images From Dataset

Original Image

Gray Image

Resize Image
(32 x 32)



DCT Result
(32 x 32)

	0	1	2	3	4	5	6	7
0	2695.75	183.03	25.67	-9.87	-28.53	190.49	-209.49	-129.36
1	-35.63	-9.4	48.99	-183.2	-164.72	103.19	94.72	162.46
2	-180.23	-17.94	-187.33	48.69	-97.08	1.02	128.6	-8.86
3	170.07	171.71	110.75	-15.24	175.67	-5.7	-4.93	-46.96
4	0.83	-23.5	36.72	32.19	74.75	-1.5	46.93	-0.62
5	-34.72	-57.99	-18.34	68.62	43.66	-91.67	-12.05	23.01
6	-36.69	41.82	30.53	-50.22	35.3	7.33	-86.91	7.91
7	102.35	17.89	-57.32	13.7	-73.18	-12.67	9.38	19.77

Low Frequency
(8 x 8)

Calculate
Average Value

Threshold

1	1	0	0	0	1	0	0
0	0	1	0	0	1	1	1
0	0	0	1	0	0	1	0
1	1	1	0	1	0	0	0
0	0	0	0	1	0	0	0
0	0	0	1	0	0	0	0
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0

Binary Matrix

c42712e808100080

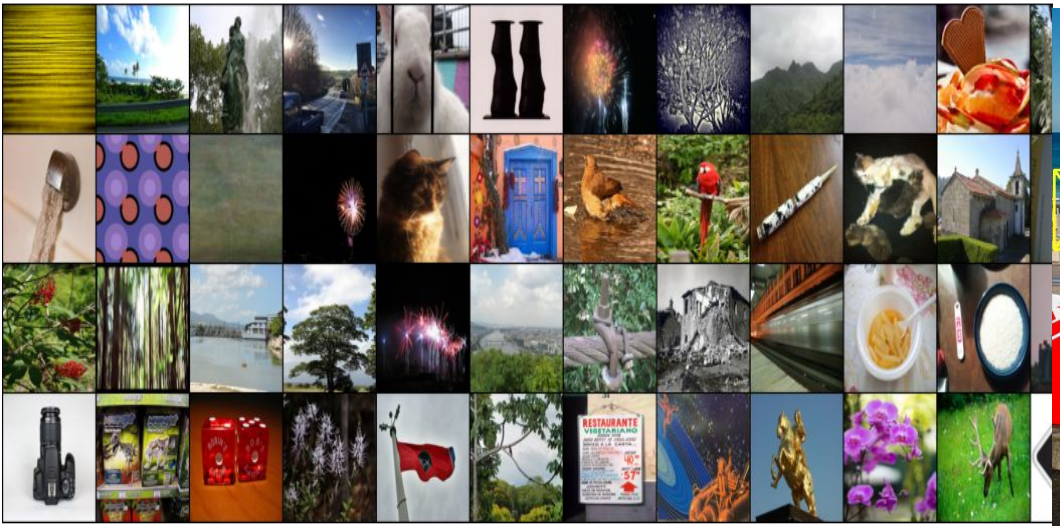
Hash Value

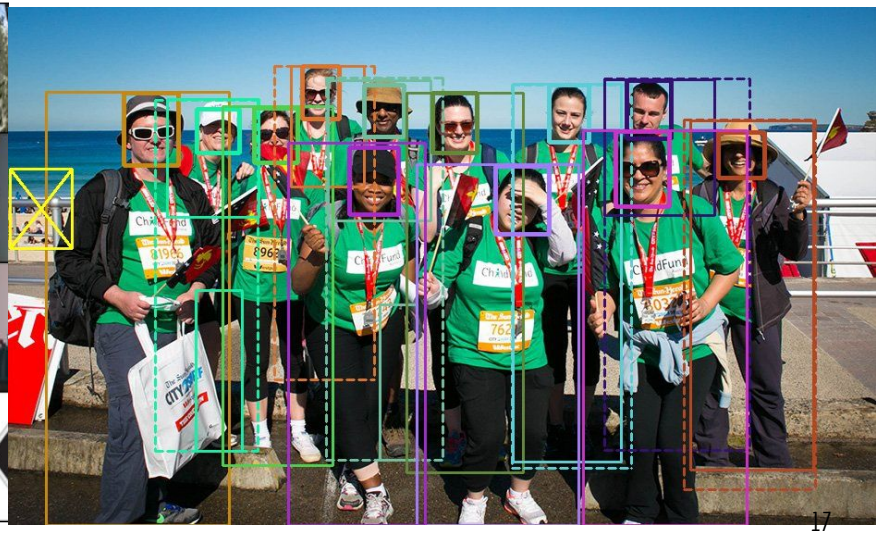
https://backend-live-coc.cfr.org/cdn/ff/BOg_dKOSmNuEIFzJF5pDXyubS8SGjNPFRcSWanThunU/1578015356/public/publications/global-memos/RTS2F7RC_1.jpg

Filter for Humans in a photo

Combine PASS (Asano et al. 2021) and CrowdHuman (Shao et al. 2018)dataset.

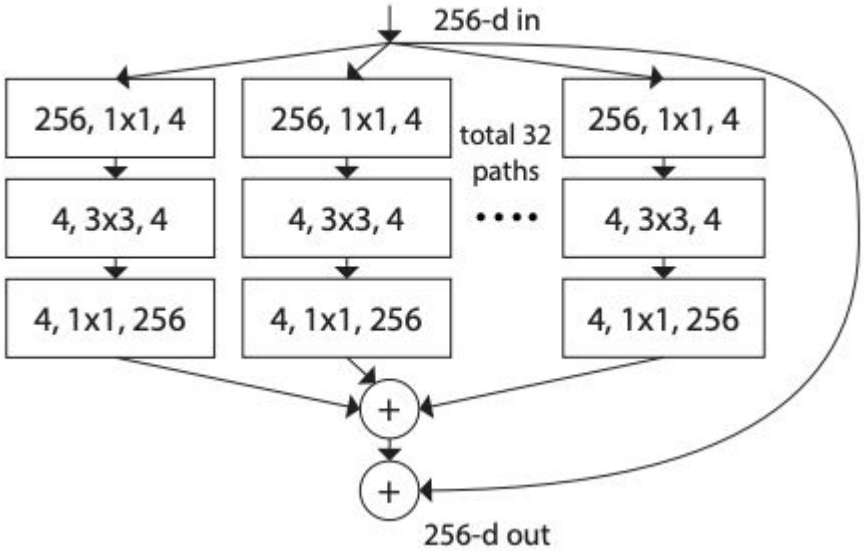
Train a CNN ResNeXt-101_32X8D (Xie et al. 2017) with transfer learning from ImageNet and 3 additional linear layers. 90% accuracy





ResNext 101 32x8D

- 101 layers deep
- 32 paths in each block
- 8 is the depth of each each block



(Xie et al. 2017)

StyleGAN 2

Why we chose StyleGAN 2?

This was trained on the Flickr-Faces-HQ Dataset(Karras et al 2019).

The dataset consists of 70,000 images

(Karras et al 2020)

StyleGAN 2 Generator architecture intro

Random noise z is mapped to A using mapping network f.

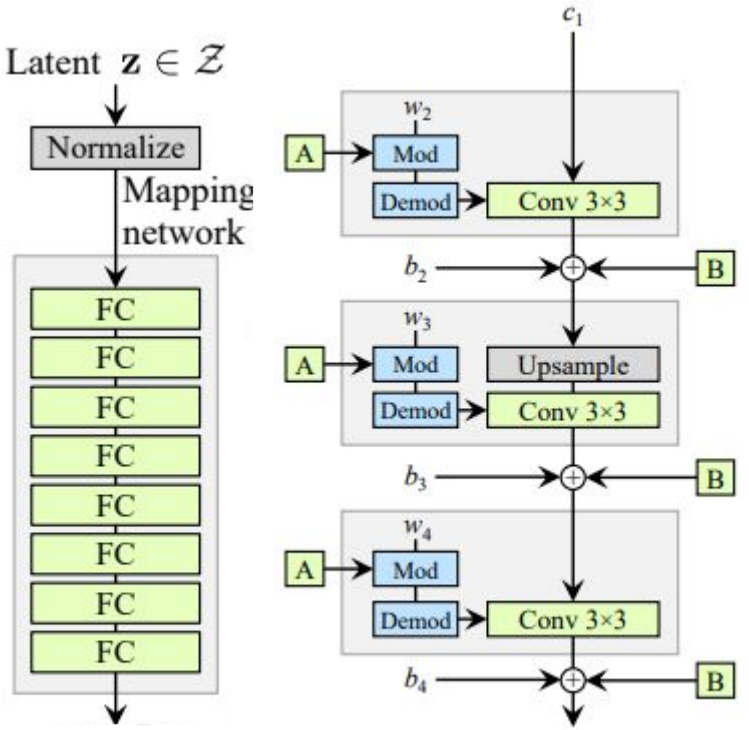
Modulate and Demodulate

W and b are learned weights and biases

B is random noise

$$w'_{ijk} = s_i \cdot w_{ijk},$$

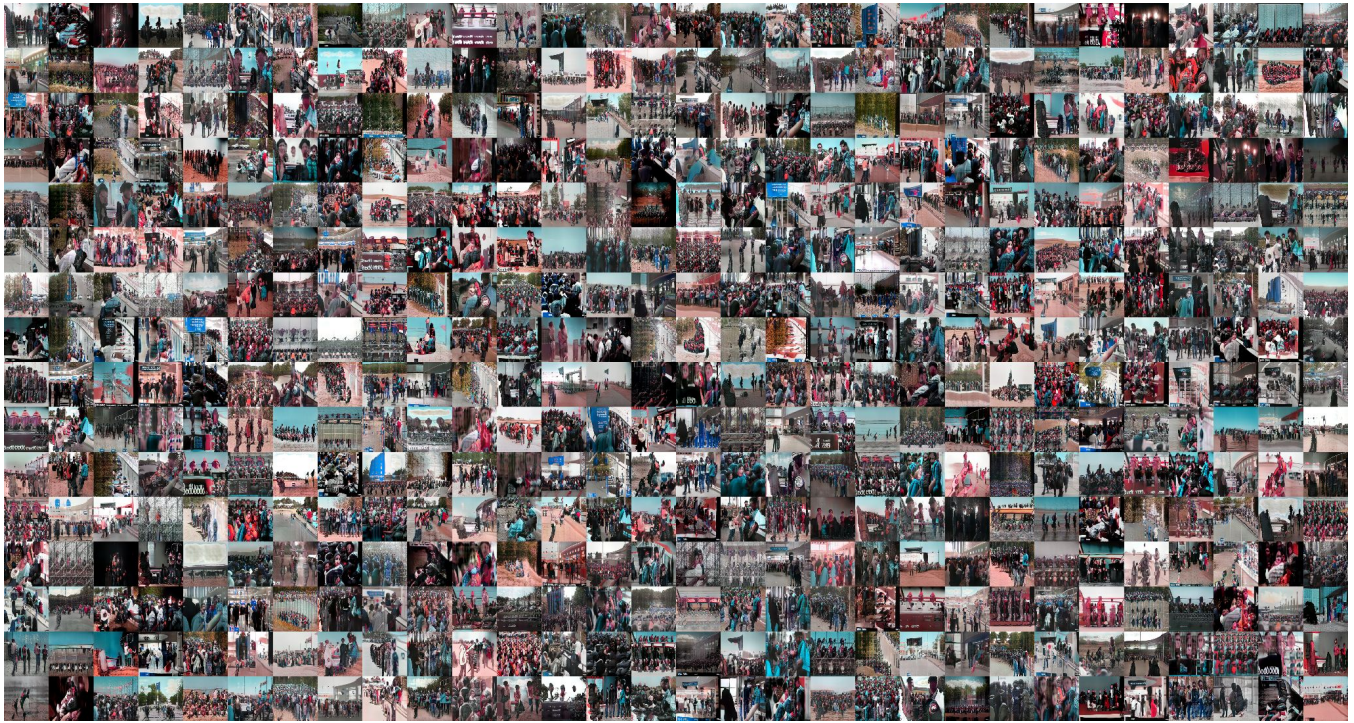
$$w''_{ijk} = w'_{ijk} \Big/ \sqrt{\sum_{i,k} {w'_{ijk}}^2 + \epsilon},$$



(Karras et al 2019,2020)

GAN results at 10,000k images

GAN results after training on 10,000k images



GAN results after training on 15,000k images

GAN results at 15,000k images

Frechet Inception Distance

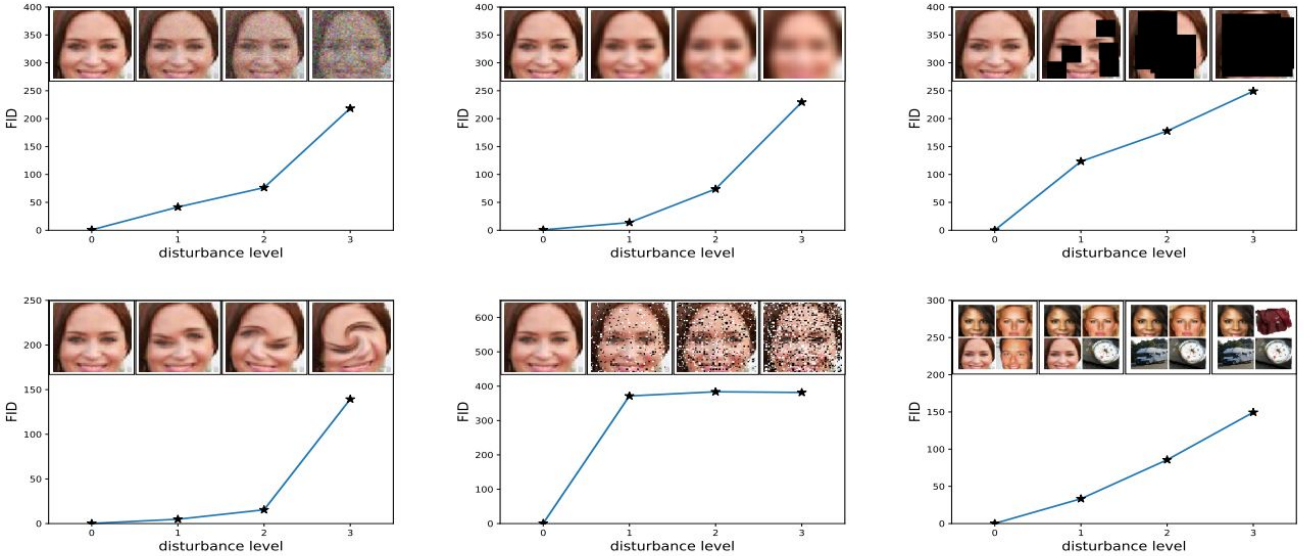
- Extract features from a set of real and synthetic images using the Inception V3 (Szegedy et al 2016) network trained on ImageNet (Deng et al 2009)
- Assume Gaussian distribution for real and synthetic image features.
- Measure distance between these distributions. Frechet or Wasserstein 2 distance, using the mean and covariance matrices of both distributions.

(Huesel et al 2017)

Frechet Inception Distance

In short it is a measure of image “quality”.

A lower score is better



(Huesel et al 2017)

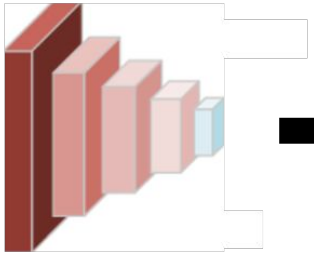
Our FID scores

	FID
0 images	207.25
10,000k images	46.06
15,000k images	127.72

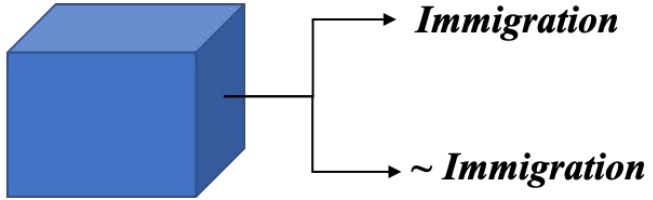
Evaluation Method

- **Feature Extraction + One-Class-Classifier**

Real Immigration Images (Annotation Manually)



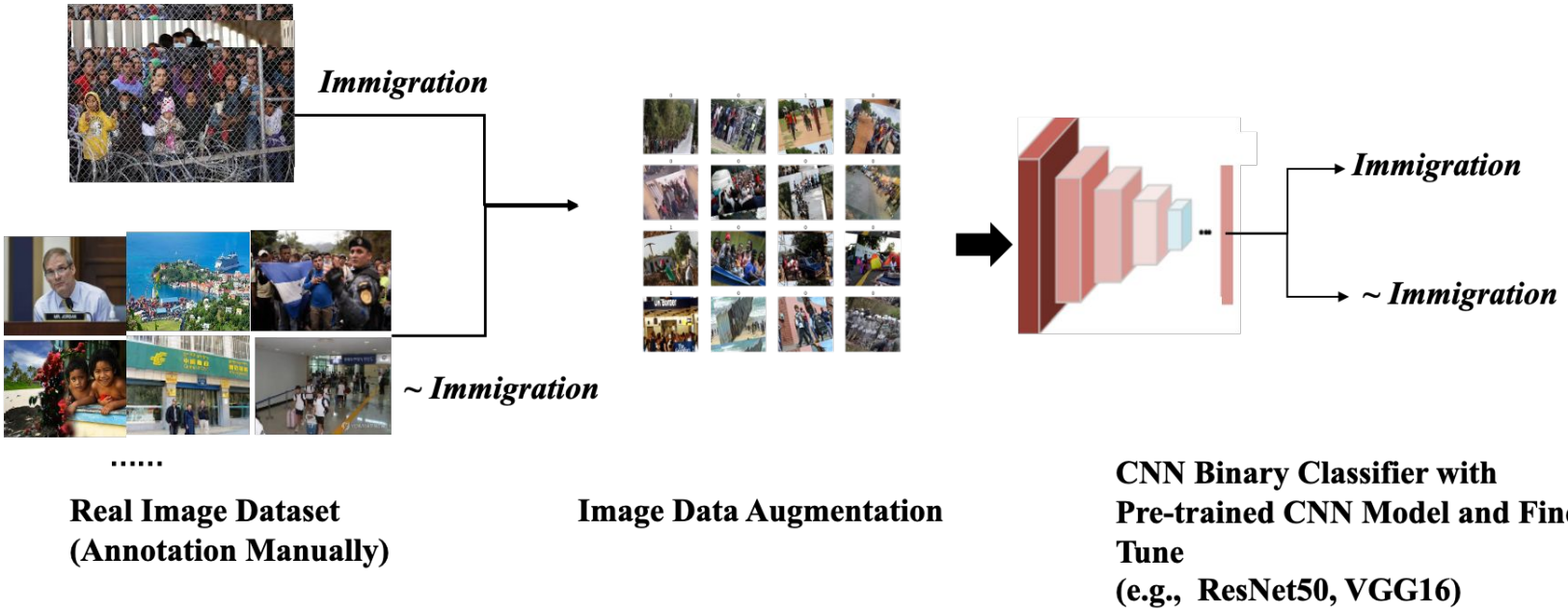
Feature Extraction with Pre-trained CNN Model (e.g., ResNet50, VGG16)



One-Class-Classifier (e.g., One-Class-SVM, Isolation Forest)

Evaluation Method

- **CNN Binary Classifier with Transfer Learning**



Real Image Dataset (Annotation Manually)

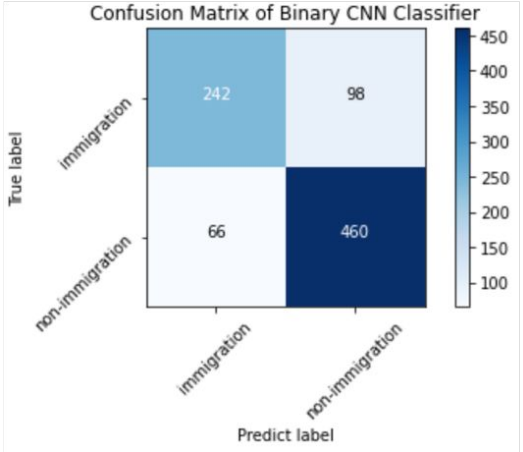
Image Data Augmentation

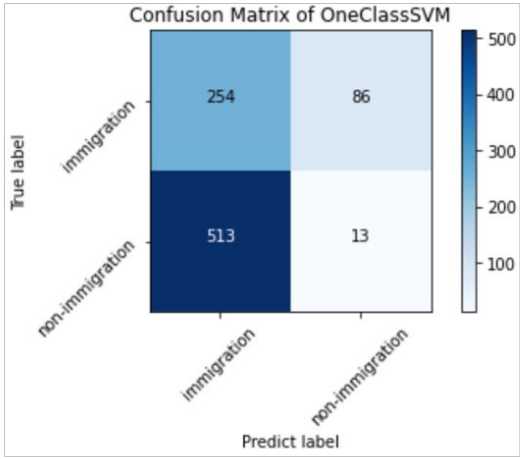
CNN Binary Classifier with Pre-trained CNN Model and Fine Tune (e.g., ResNet50, VGG16)

Evaluation Result

Performance Measure on Real Image Dataset

Method	Precision	Recall	Accuracy
Binary CNN Classifier	**0.79**	0.71	**0.81**
One-Class-SVM	0.33	**0.75**	0.31





Examples of Mis-Classified Images

Evaluation on Synthetic Images

$$\frac{\text{Classified as Immigration Images Number}}{\text{Total Number of Synthetic Images}} = \frac{1605}{2276} = 0.70$$

Conclusion

6. Tasks

T.1. Find real images related to immigration using web scraping methods

T.2. Generate synthetic images from the real ones using GAN

T.3. Extract features, using pre-trained CNN, from real and synthesized images and feed a classification algorithm to

T.4. Evaluate the outcome, get a classification score for the real ones, and a classification score for the augmented dataset, and compare

COMPLETE ✓

Based on our initial results, we conclude that GAN's can potentially be used to augment our immigration dataset, but we believe must first increase our dataset size further in order to generate photo realistic synthetic images.

THANK YOU