Visual Recognition with Fashion Dataset

Group members:
Haiyan Tian
Fengyi Zheng
Ziwei Zhao

Outline

- Introduction
- Data Preparation
- ResNet152
- Back Propagation
- Deep Convolutional Gan
- Summary

Introduction

Image source: https://blog.valohai.com/clothes-detection-for-fashion-recommendation

Data Preparation

Original Dataset:
- Obtained from Kaggle
- Over 30 GB
- 44442 images
- 44447 image ids and descriptions

id	gender	masterCategory	subCategory	articleType	baseColour	season	year	usage	productDisplayName
15970	Men	Apparel	Topwear	Shirts	Navy Blue	Fall	2011	Casual	Turtle Check Men Navy Blue Shirt
39386	Men	Apparel	Bottomwear	Jeans	Blue	Summer	2012	Casual	Peter England Men Party Blue Jeans
59263	Women	Accessories	Watches	Watches	Silver	Winter	2016	Casual	Titan Women Silver Watch
21379	Men	Apparel	Bottomwear	Track Pants	Black	Fall	2011	Casual	Manchester United Men Solid Black Track Pants
53759	Men	Apparel	Topwear	Tshirts	Grey	Summer	2012	Casual	Puma Men Grey T-shirt
1855	Men	Apparel	Topwear	Tshirts	Grey	Summer	2011	Casual	Inkfruit Mens Chain Reaction T-shirt

Background Info

- Clothing Detection Tools have been adapted into fashion businesses

- Related Work:
  Google's Image Swirl Project

- For this Project:
  ResNet 152+
  Transfer Learning

Image source:
https://www.semanticscholar.org/paper/Clothes-detection-and-classification-using-neural-Cychnerski-Brzeski/b245580bd7bbdd8f96fd079fcfa23e05e731a0df

Data Preparation

- Original Dataset

- Resizing Images and Id mismatch error

- Classifying

Class 0: Men/Accessories   Class 1: Women/Accessories.

Class 2: Men/Appeal.   Class 3: Women/Appeal.

Class 4: Men/Other.   Class 5: Women/Other.

Resizing and id mismatch

- 5 mismatched image ids

- Original size range 150*200 ～ 3744*5616

- Resized: model training: **224*224** ; GAN: **64*64**

Objectives

1, Use Convolutional Neural Network to extract feature and classify images

2, Apply Guided Backpropagation to understand the latent factors behind model and interpret the prediction

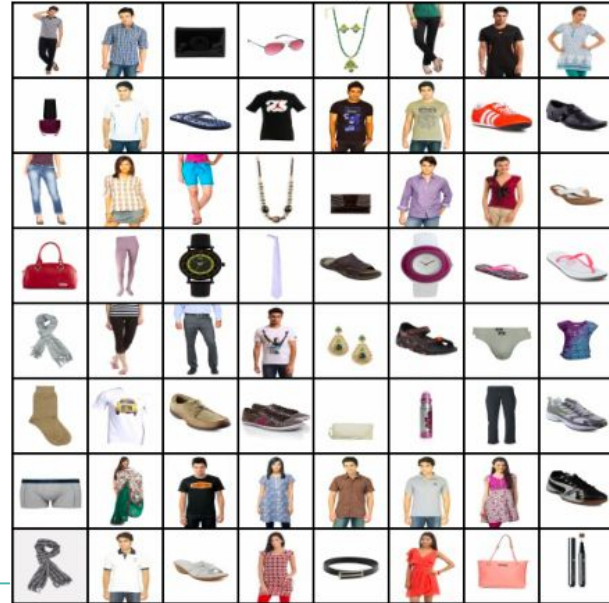3, Leverage DCGAN to generate new clothes images

Proportion

Training : Validation : Testing
  75   :   10   :   15

31400 :   3240   :   5600

Training Images

ResNet 152

- Epoch: 25
- Batch size: 256
- Learning rate = 0.0001
- Train: 31, 395 Test: 5,600 Validation: 324
- What is Resnet152 and Why we use it
  - ResNet152 is build upon the original ResNet. It has 152 total layers and also has the least top 1 and top 5 error out of all the proposed ResNet model in pytorch.
  - It is able to increase depth while eliminated unnecessary complexity
- Transfer learning
  - A technique where you train a model on a large dataset (ImageNet) and then fine tune it to your target dataset (our dataset)
  - Why?
    - i. Our dataset is relatively small. To help the model train better and increase performance, we decided to employ transfer learning with ResNet152 that was previously trained on ImageNet
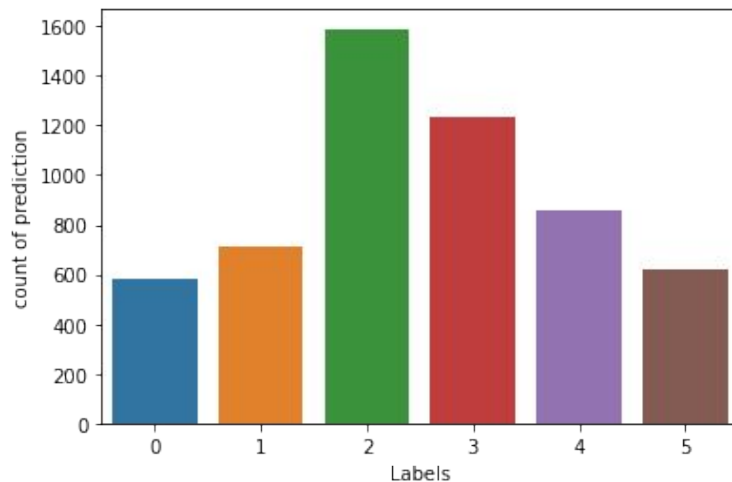- Optimization: Adams

ResNet 152 and Results

- Training Accuracy: 93.26%

- Validation Accuracy: 91.26 %

- Test Accuracy: 91.75%

Classifying

Train distribution of classes
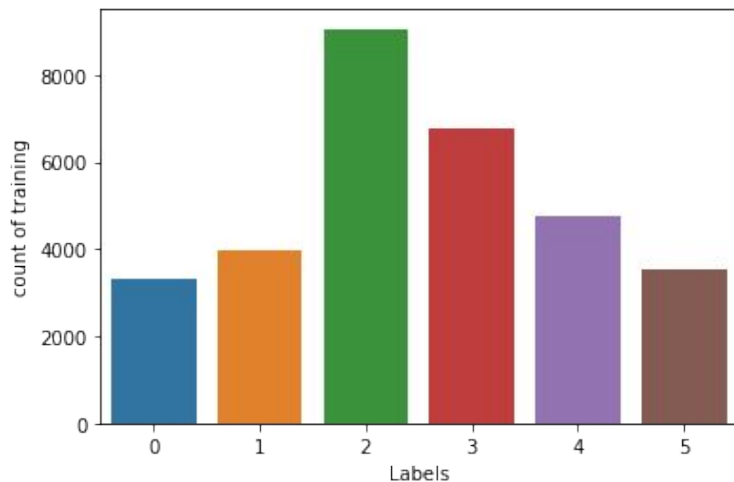
Predicted distribution of classes

The distribution of trained label is very similar to the distribution of predicted label which shows our model does not have a biases towards any class.

Results

Guided Backpropagation

- We want to interpret model:
  - By shed light on what our model see
- What is Backpropagation:
  - It is an algorithm that allows us to see which pixels in the inputs has the most influence on the prediction
- Interesting thing to notice
  - The guided backpropagation shows appropriate outline of the image
  - Detect face more than cloth

predicted label: 0
Actual label:0, Men/Accessories

predicted label: 1
Actual label:1, Women/Accessories

predicted label: 2
Actual label:2, Men/Apparel

predicted label: 3
Actual label:3, Women/Apparel

predicted label: 4
Actual label:4, Men/Other

predicted label: 5
Actual label:5, Women/Other

Average Guided Backpropagation

predicted label: 0, Men/Accessories   predicted label: 1, Women/Accessories   predicted label: 2, Men/Apparel

Average Guided Backpropagation

predicted label: 3, Women/Apparel

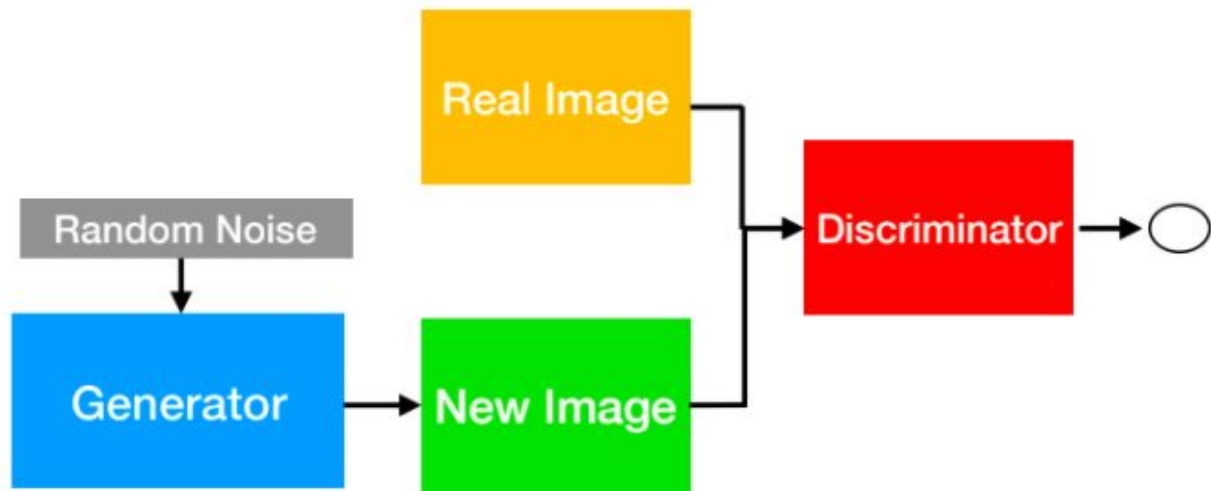predicted label: 4, Men/Other

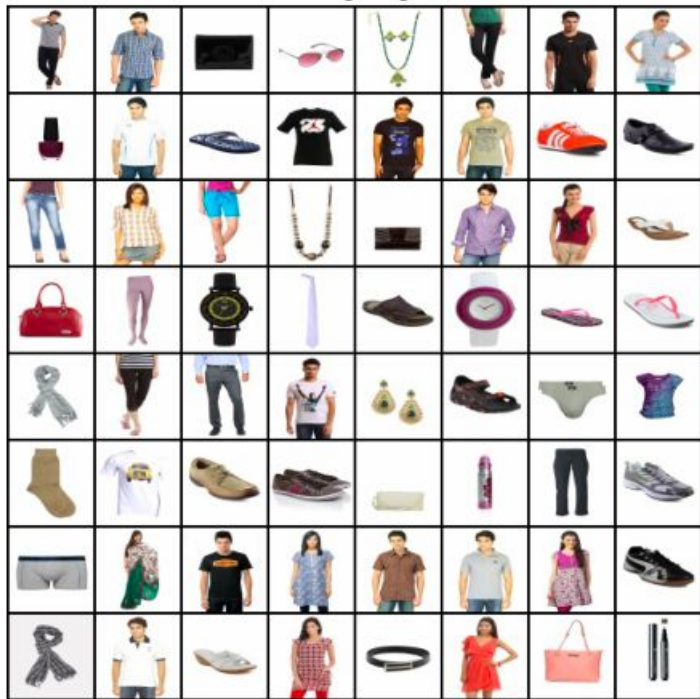predicted label: 5, Women/Other

DCGAN

- Generator: Learns to generate better images to fool the discriminator.
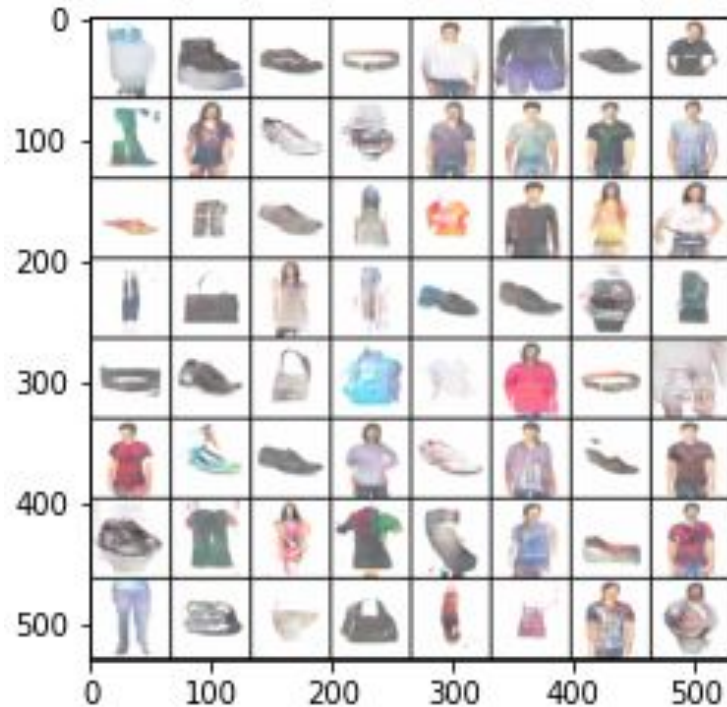- Discriminator: Learns to become better as distinguishing real from fake images

Image source:
https://github.com/rasbt/stat479-deep-learning-ss19/blob/master/L17_gans/L17_gan_slides.pdf
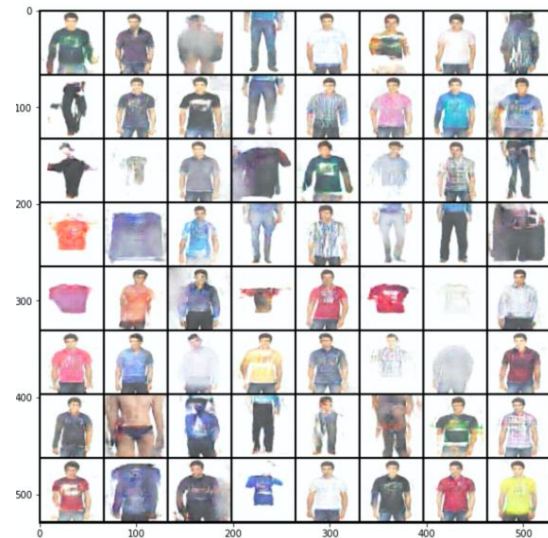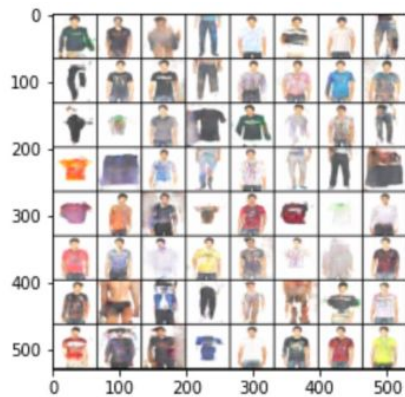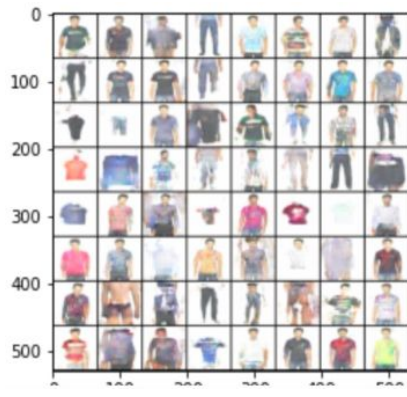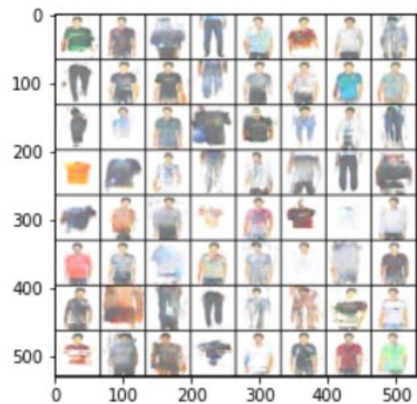
DCGAN

Original Images (real)

Generated Images (fake)

DCGAN with Class 1 (Women/Accesorries)
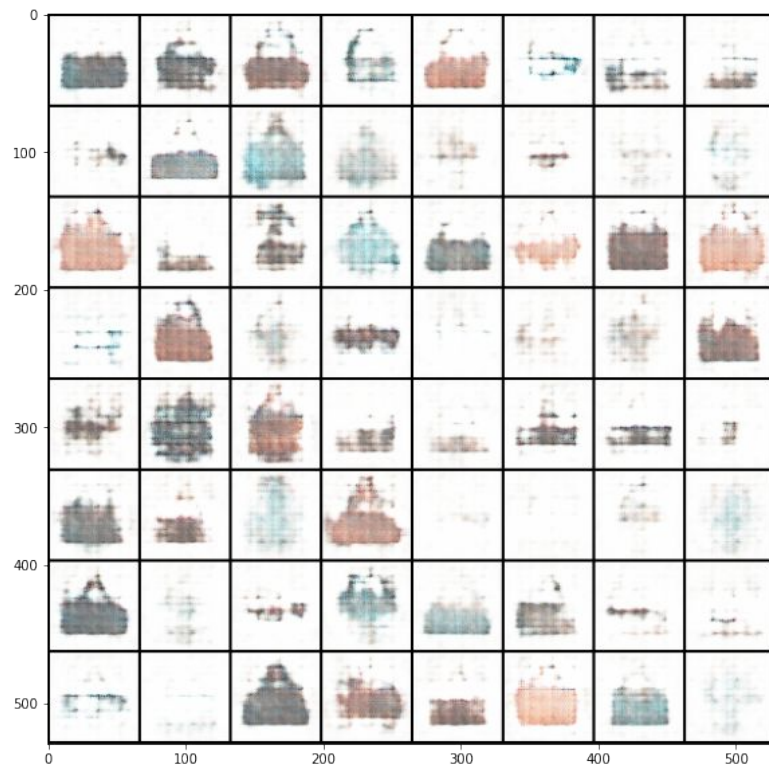
DCGAN with Class 2 (Men/Appeal)
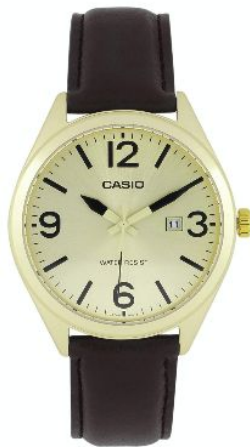
Summary

Pros:

- ResNet152 was chosen after trying out with Inception and Vgg16, with transfer learning technique, relatively higher accuracy.

Future Directions:

- Back Propagation
- Dataset not so targeted due to transfer learning
- Hyperparameters for Gan

Cons:

- Information lost during resizing.

1080*1440   64*64

Thank You

Image source:
https://www.chinabrands.com/dropshipping/article-top-6-best-wholesale-clothing-websites-114.html