Image Classification and Product Recommendation System

codes and docs are in https://github.com/561-Proj

Zihe Yang (zy151), Jianing Sun (js4770) and James Gao (hg356)

Introduction

The coronavirus pandemic is rapidly changing our behavior toward e-commerce, and the shifts are likely to stick post-pandemic.

According to the data released by Howard University[1], over 56 % of surveyed U.S. adults stated that they shopped more online, 67 % report going to the shop less, and even those who had never used an e-commerce service felt motivated to do so.

This growing e-commerce industry presents us with a large dataset waiting to be researched on. We plan to start with fashion product images with several label attributes (category, color, size and etc)

Our Business Case:

(1) to train a an image classifier.

(2) build a recommendation model outputs similar images.

Dataset

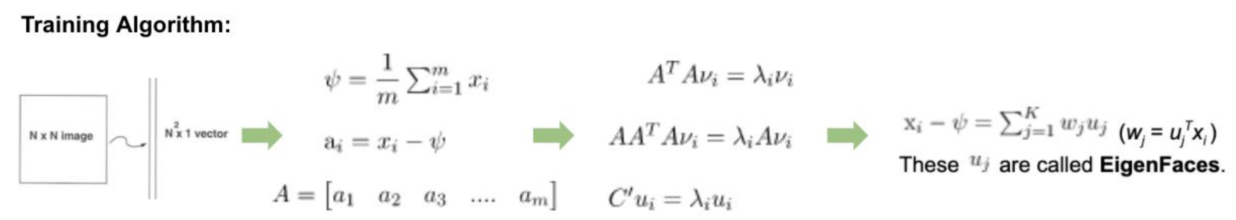
The dataset contains mainly contains professionally shot images in the fashion industry and attributes of those images. There are 4 sub datasets in it.

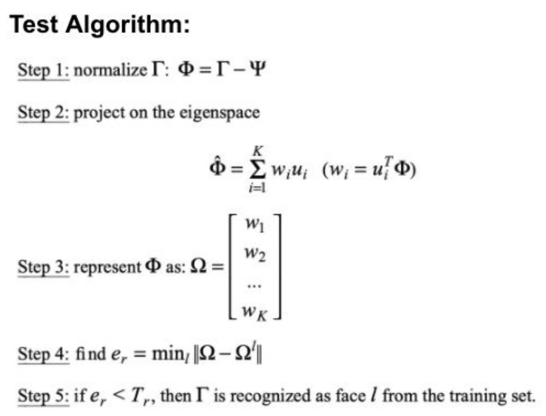
(1) The first dataset contains over 44,000 images in jpg format with unique IDs used for identification. Each image is about footwear, shirts or accessories like jewelry and fragrance.

(2) The second dataset contains the attributes of these images in json format.

Methodology

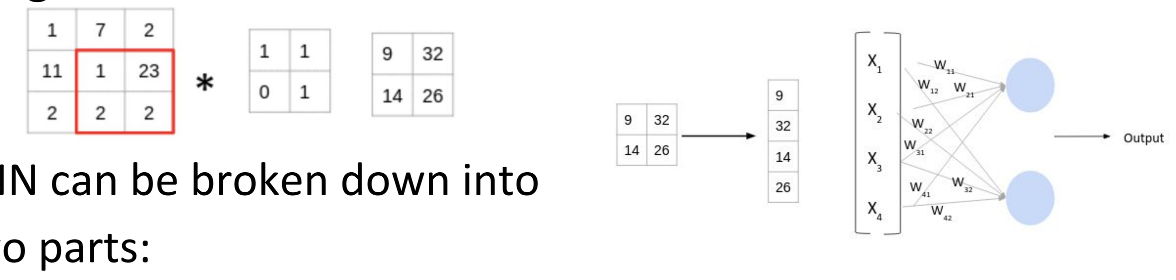
Image Classification - Eigenfaces





Eigenfaces method creates a low-dimensional basis of images for a set of features using PCA, and project a training sample on small feature space.

Image Classification - Convolutional Neural Network



CNN can be broken down into two parts:

- **The convolution layers**: Extracts features from the input
- **The fully connected (dense) layers**: Uses data from convolution layer to generate output.

Methodology

Product Recommendation:

(1) Image Processing: Apply ResNet-50 for image embedding.

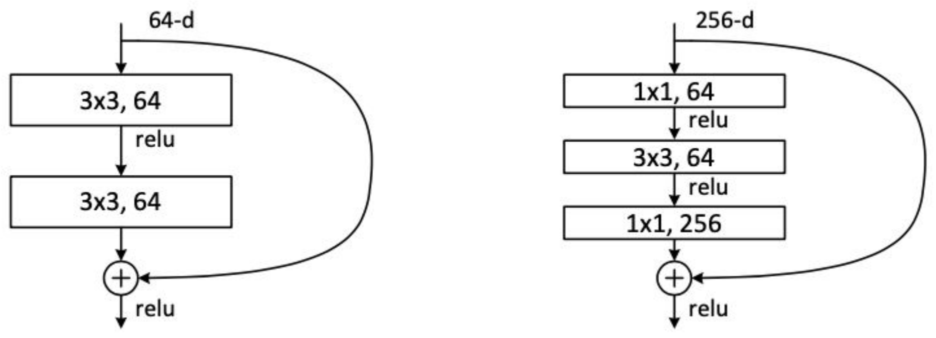
ResNet (Residual Network) solves the problem of no convergence and vanishing gradient in deep convolutional neural network.

Image Processing, in essence, is Boundary Value Problem (BVP), because we are minimizing loss of "interpolation" in every grid boundary. And it's proved that multigrid method for BVP converges at a *log* rate by minimizing residual function.

Therefore, In contrast to *min Loss(f)*, where $f = h_1(a_1 h_2 + b_1)$ and $h_1 = h_2(a_2 h_3 + b_2)$ … in normal stacked deep-layer neural network problem, ResNet is *min Loss(residual)*. Every layer is : $H(x) = F(x) + x$. Adding this shortcut connection makes target function as minimizing residual, and will not cause increase of parameters.

(2) Compute similarity of image embeddings with cosine metrics.

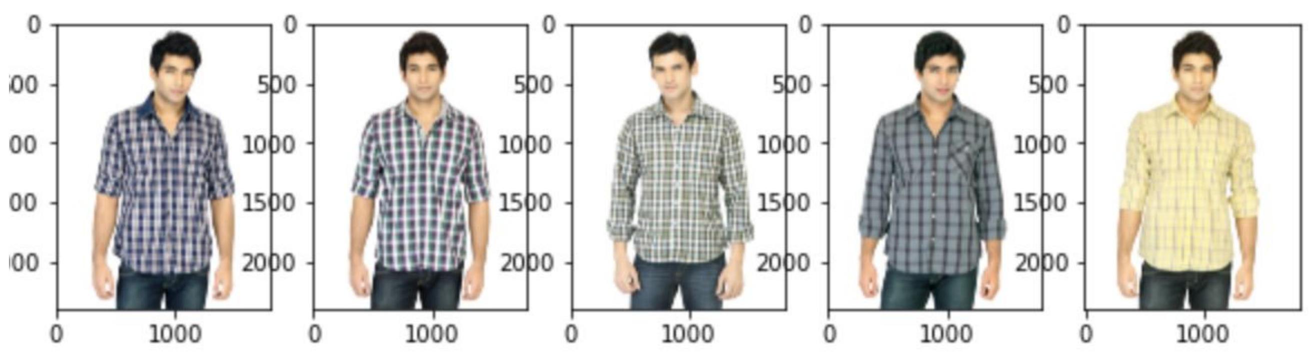
Cosine is between [0,1] and commonly used in comparing similarities. Wasserstein distance results from a partial differential equation (PDE) formulation of Monge's optimal transport problem, and was proved to be an efficient numerical solution method for solving Monge's problem and image comparison.



Results

Product Recommendation:

1. **Retrieval:** retrieve the items that is in the same category as target id in result of image classification. This step shrinks candidates pool from 5000+ down to 500-800 items for saving computation resources.
2. **Ranking:** Apply ResNet over cropped images, and obtain embeddings of dimension (2048,) array for every product image; Compute similarities over retrieved dataset, and sort out top 10 items for recommendation.



When test_id = 15970, these ranking scores are 1, 0.949, 0.946, 0.946, 0.944, which means the features, plaid shirt, are significantly learnt. When test_id = 58183, the scores are 1, 0.928, 0.909, 0.906, 0.901; the features, buckle and bag handle, are relatively vague.

Results

Preprocessing

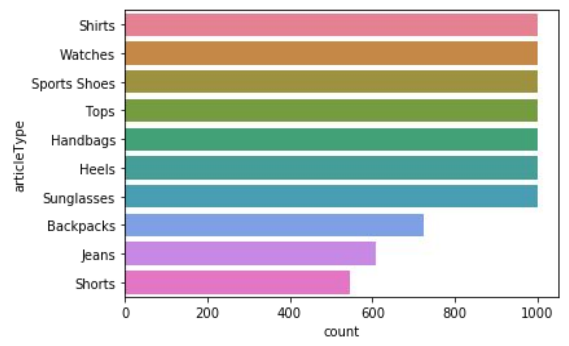
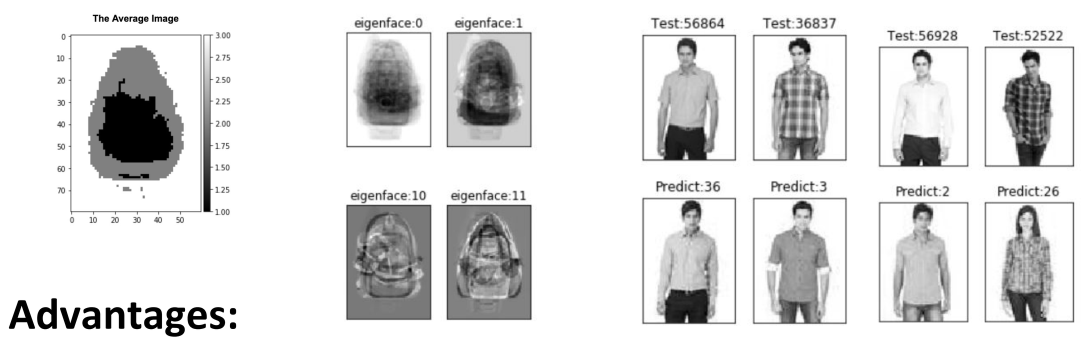


Image Classification - Eigenfaces



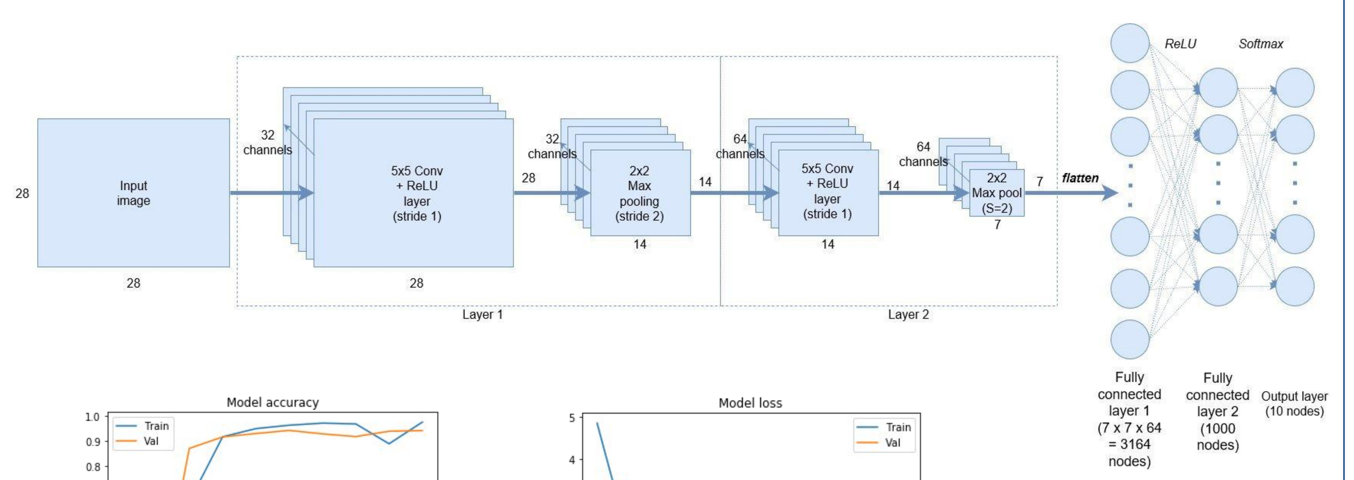
Advantages:

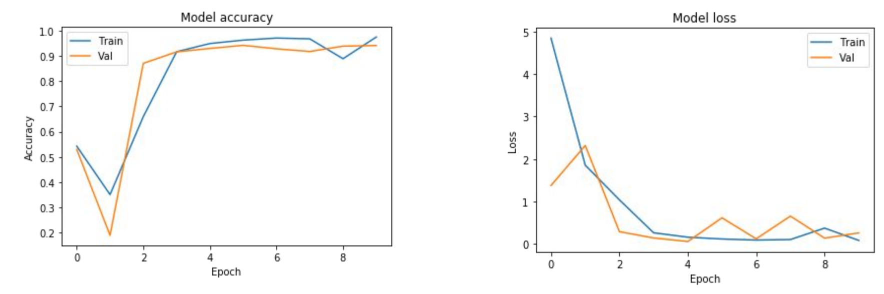
- Easy to implement and computationally less expensive.
- No knowledge of the image required.

Limitations :

- Proper centered subject is required for training/testing.
- The algorithm is sensitive to lightning, shadows.
- Front view of the image is required to work properly.

Image Classification - Convolutional Neural Network





Future Work

In product recommendation,

1. Some photos are with models while some are not, and this may cause bias in similarity scores, and affect the recommendation. More work in instance segmentation should be done.
2. Cosine metric was used in measuring similarities between images, and more researches could be focus on Wasserstein distance for comparison.

To obtain a more accurate set of predictions, there can be some improvements made to the algorithm.

Challenges involved in image processing and classification:

(1) Recognition

(2) Viewpoint variation

(3) Illumination

(4) Background clutter

References

1. Kaiming He, et, Deep Residual Learning for Image Recognition
2. W. L. Briggs, S. F. McCormick, et al. A Multigrid Tutorial. Siam, 2000.
3. R. Szeliski. Locally adapted hierarchical basis preconditioning. In SIGGRAPH, 2006.
4. Monge's Optimal Transport Distance for Image Classification
5. Sheng Zhang and Matthew Turk (2008) Eigenfaces. Scholarpedia, 3(9):4244.
6. Vincent Dumoulin, et al, "A guide to convolution arithmetic for deep learning", 2015, arXiv:1603.07285
7. A. W. Harley, "An Interactive Node-Link Visualization of Convolutional Neural Networks," in ISVC, pages 867-877, 2015

Georgetown University