Machine Learning II:
Identifying Clothing Items using Convolution Neural Network

Jason Achonu, Kyi Win, Manoah Farmer

Dec 4, 2018

Overview

Introduction

Dataset

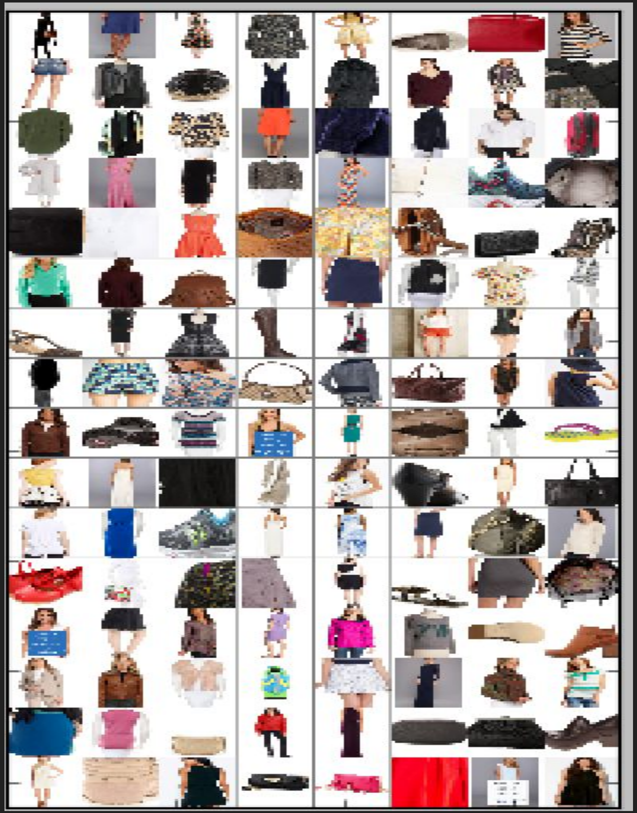
Data Preprocessing

Model

Experimentation

Results

Discussion/Conclusion



Dataset

The dataset was compiled by University of North Carolina at Chapel Hill. It consists of 404,683 images of clothing from different online stores.

Original 11 categories: bags, belts, dresses, eyewear, footwear, hats, leggings, outerwear, pants, skirts, and tops.

Due to a lack of computational and storage resources, the dataset was reduced from eleven categories to the six categories which are: tops, skirts, bags, outerwear, dresses, and footwear resulting in a total of 70,500 images which were randomly selected from the six categories.

Data Preprocessing

- Code split data into training and test sets
- Input image transformation
  - Resize
  - Centercrop
  - Normalization

Image Classification

Set of Predefined categories: Bags, Dresses, Footwear, Outwear, Skirts and Tops

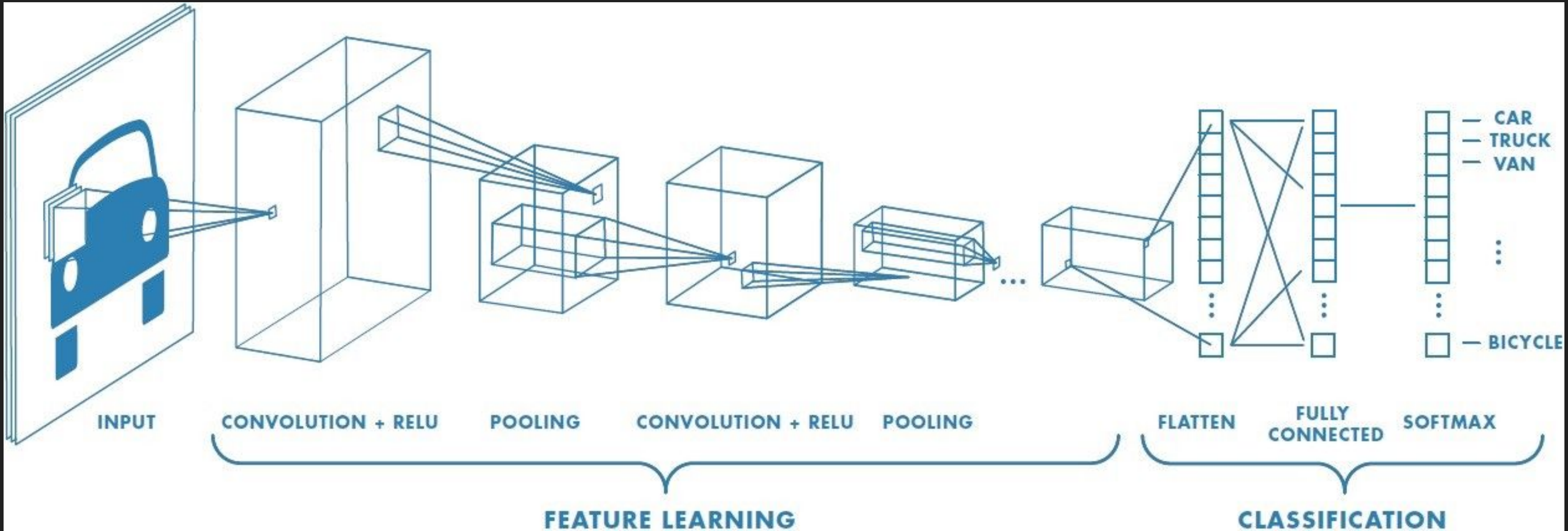
SKIRTS

Image Classification : Clothing Items



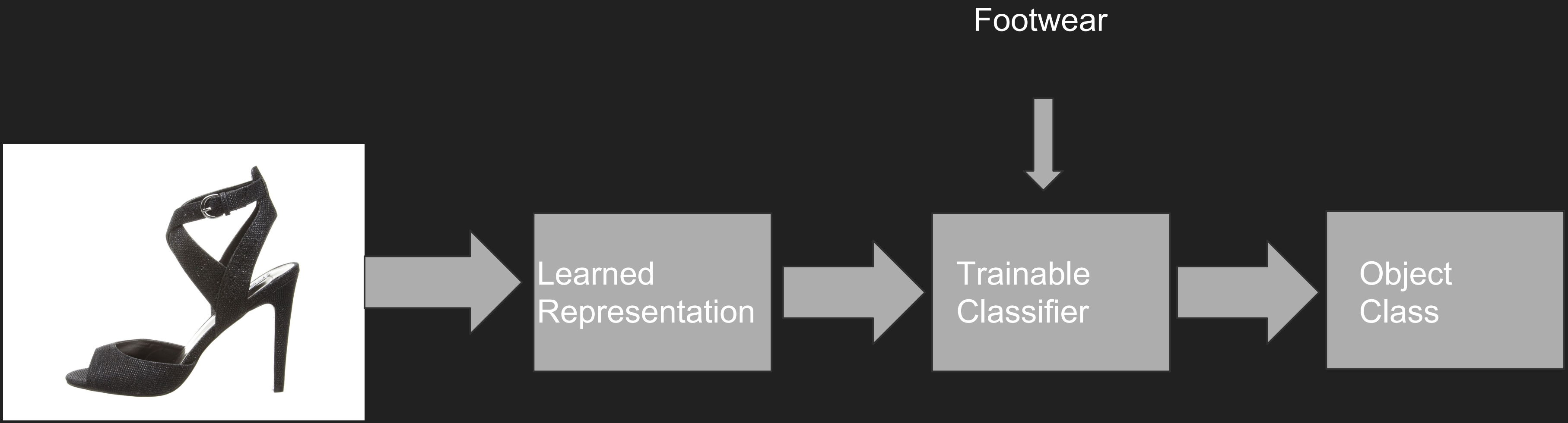



Model: CNN



CAR
TRUCK
VAN
BICYCLE
INPUT
CONVOLUTION + RELU
POOLING
CONVOLUTION + RELU
POOLING
FLATTEN
FULLY CONNECTED
SOFTMAX
FEATURE LEARNING
CLASSIFICATION

Image Classification Pipeline



Instead of creating hand designee feature extraction, the CNN provides learner known as kernels to learn and extract the features.

Experiments

Learning rate = 0.1

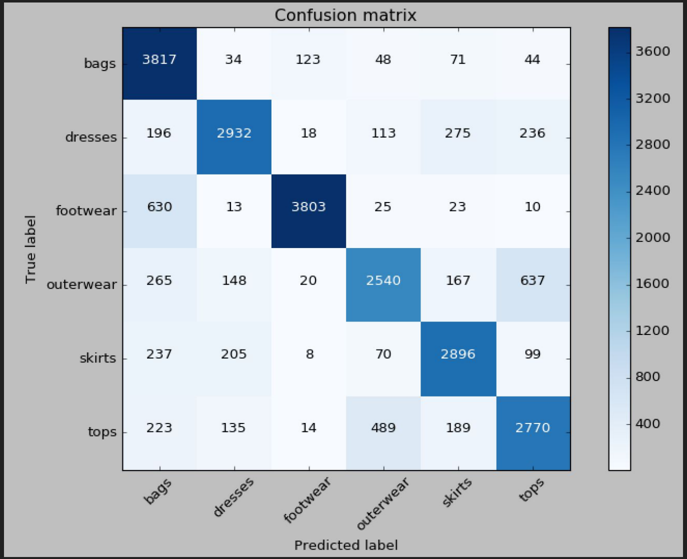
Epochs	Batch Size	Conv Layers	Number of Feature Maps	Kernel size	Maxpool Layers	Accuracy (Train)	Accuracy (Test)	Time (seconds)	Optimizer
5	64	5 (B)	32	5	4	65	72	717	Adam
5	64	5 (S)	32	5	4	67	76	795	Adam
10	64	7 (B)	32	5	4	64	71	1727	Adam
10	64	7 (S)	32	5	4	70	64	1654	Adam
10	64	5 (S)	64	5	4	80	79	1542	Adam

Experiments

	1st		2nd		3rd		4th		5th (Best)	
Category	Accuracy	F - Score	Accuracy	F - Score	Accuracy	F - Score	Accuracy	F - Score	Accuracy	F - Score
Bags	74	73	85	78	81	76	76	73	94	80
Dress	73	76	78	76	74	75	76	70	78	81
Footwear	86	84	88	88	76	82	85	83	86	90
Outerwear	65	63	59	65	63	62	58	55	69	72
Skirt	75	77	80	78	81	74	84	73	86	81
Tops	60	62	75	68	52	61	40	55	73	73

Results (Confusion Matrix and Classification Report)

- We can interpret from this results that this model classified 94% correctly out of all the footwear images that are labeled as footwear and out of all the footwear images, it classified 89% of images with accurate labels.



	precision	recall	f1-score	support
bags	0.84	0.85	0.85	4137
dresses	0.81	0.79	0.80	3770
footwear	0.94	0.89	0.92	4504
outerwear	0.69	0.77	0.73	3777
skirts	0.75	0.86	0.80	3515
tops	0.79	0.65	0.71	3820
avg / total	0.81	0.81	0.81	23523

Accuray rates on each class (SGD best model)

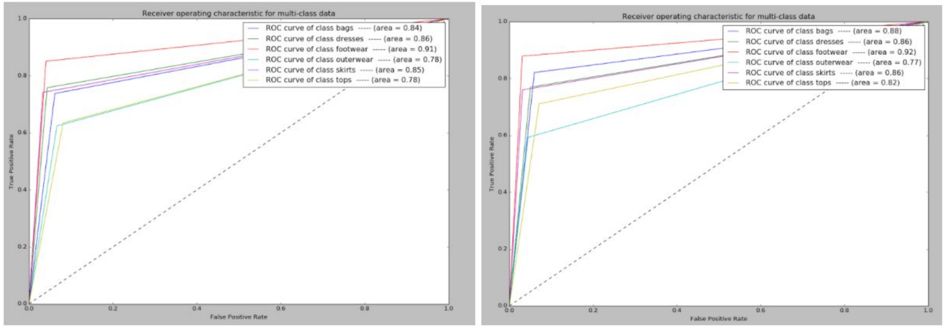
Test Accuracy of the model on the test images: 80 %

Accuracy of bags : 85 %
Accuracy of dresses : 83 %
Accuracy of footwear : 90 %
Accuracy of outerwear : 74 %
Accuracy of skirts : 88 %
Accuracy of tops : 62 %

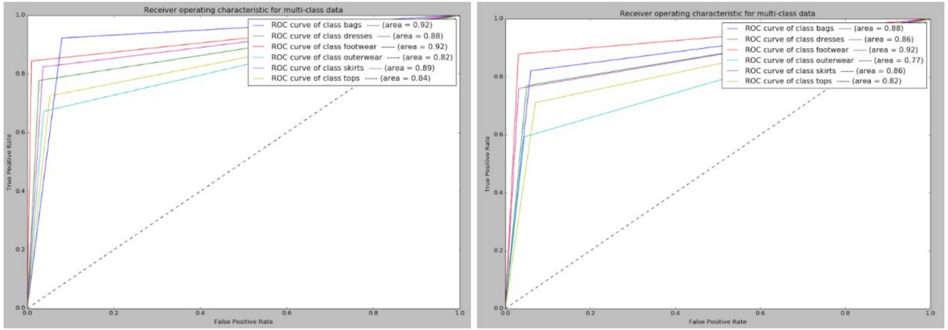
Results (ROC)

- straight technique has better accuracy than bottleneck technique
- SGD has better accuracy than Adam in this project.

(left-bottleneck, right-straight)



(left-SGD, right-Adam)



Discussion and Conclusion

- We found that footwear, bags and skirts consistently performed the best in all of the experiments we conducted out of the categories we chose.
- In conclusion, the model performed better than random guess; however, there are ways this model could be improved going forward
- One way to do this would be to simply gather more data and possessing more processing power.
- We could also try tuning the hyper-parameter differently.

Resources

https://forums.fast.ai/t/image-normalization-in-pytorch/7534
https://stackoverflow.com/questions/47850280/fastest-way-to-compute-image-dataset-channel-wise-mean-and-standard-deviation-in
http://acberg.com/papers/wheretobuyit2015iccv.pdf
https://github.com/flipkart-incubator/fk-visual-search
https://github.com/Airconaaron/blog_post_visualizing_pytorch_cnn/blob/master/Visualizing%20Learned%20Filters%20in%20PyTorch.ipynb
https://arxiv.org/pdf/1311.2901.pdf
https://stackoverflow.com/questions/47850280/fastest-way-to-compute-image-dataset-channel-wise-mean-and-standard-deviation-in
https://github.com/amir-jafari/Deep-Learning/blob/master/Pytorch_/6-Conv_Mnist/Conv_Mnist.py
https://discuss.pytorch.org/t/data-augmentation-in-pytorch/7925
https://stackoverflow.com/questions/2301789/read-a-file-in-reverse-order-using-python
https://pytorch.org/docs/stable/torchvision/transforms.html
https://stackoverflow.com/questions/12984426/python-pil-ioerror-image-file-truncated-with-big-images
http://www.apsipa.org/proceedings/2017/CONTENTS/papers2017/14DecThursday/Poster%204/TP-P4.14.pdf
https://www.quora.com/How-can-I-calculate-the-size-of-output-of-convolutional-layer
https://medium.com/@RaghavPrabhu/understanding-of-convolutional-neural-network-cnn-deep-learning-99760835f148