CS230: Content-Based Image Retrieval System (CBIR) for eCommerce Using Deep Neural Networks  

Authors: Nicholas Sinthunont, Lee Reed

The Problem

Online retailers are failing to design products which cater to the behaviour of online shoppers in two ways:

- Search Limitations - eCommerce platforms typically allow for text-based inputs even though consumers typically rely on images for fashion inspiration
- “Not what I am looking for” syndrome - existing retail platforms require users to scroll through dozens of pages of products which may or may not be something they are interested in

Our project explores the following to solve the above problems:

- Is it possible to detect different fashion objects in a given image?
- Is it possible to classify these objects into different classes?
- Is it possible to find similar items to those objects within their respective classes?

Our Approach / Model

Object Detection

- Utilizes existing Masked R-CNN implementation utilized for detection of objects in the wild of multiple classes
- Avoids spurious edges of FCN’s & retains full image
- Object mask provided as an output to reduce noise & feed image matcher exact image

Image Matcher

- Input image fed into VGG16 model, with weights pre-trained on ImageNet, to extract image features
- Feature vectors fed into a ANN (Approximate Nearest Neighbors) implementation to identify top 5 Images

Results - Object Detection

Google Open Image

The result was that the Mask R-CNN was able to detect bounding boxes to a sufficient degree of accuracy when related to the “footwear” class but had high errors for others

Mechanical Turk

This was assumed to be the case because “footwear” typically appears in rectangular boxes but something like a hat has a more abstract shape thus requires more data to process

FCN Data Generations


Results - Image Matcher

Object Similarity Search - Dresses

Sample Similarity Matching

Looking Forward

As next step we are considering the following:

- Rigorous testing to validate and tune model against different fashion classes e.g. men vs. women jeans
- Further experimentation with similarity search algorithm to compare Amney performance with triplet loss approach
- Incorporating unsupervised techniques to automatically determine classes for boundary boxes
- Removal of background noise from detection output

Key References:
6. https://github.com/apark/tenney

Note: See Project Report for full list of citations