Clothes Recognition and Retrieval

Wei-Hsiang, Shen
tommyrpg1010@gmail.com

Chong-Yao, Zhang Jian
jerry_zj@icloud.com

Yi-Lin, Hsu
lsst50213@gmail.com

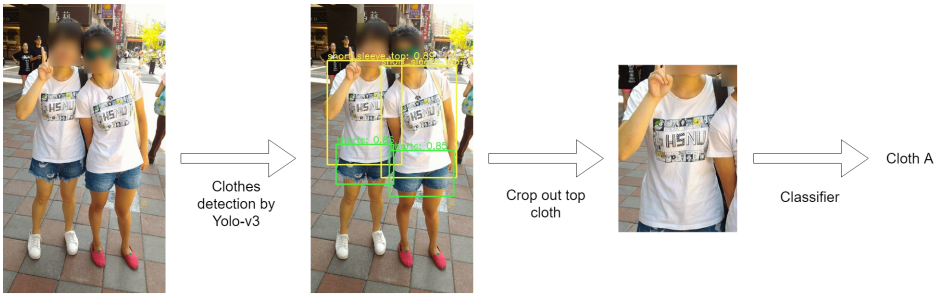

Figure 1: System pipeline.

1 Introduction

Recently, online retails stores are growing quickly and surpass traditional physical stores. Consumers can browse through thousands of merchandises at online retails stores. Sometimes, it is difficult to search the ideal item we want in the massive item choices offered by all online stores. For instance, it is time-consuming to search the ideal cloth since there are too many choices.

In this work, we propose a system (Fig [1]) to help users search their ideal clothes. Users can query the database by a arbitrary photo, and the system would recommend the most similar cloth in the store. We would mainly focus on:

- Clothes detection to crop out top short sleeve image
- Clothes classification to find the most similar cloth

2 Related Works

Kalantidis, et al. [1]. proposed an automatic clothing suggestions using arbitrary photos. They perform human pose estimation to segment the images in 26 body parts. Then, they segment and cluster the clothing using Approximate Gaussian Mixture clustering. After segmenting the clothing, they then perform classification to suggest the most similar clothing. However, our work make use of the newly release Deep Fashion 2 [2] dataset to train a object detection network, thus omitting the need of pose estimation and clothes segmentation.

3 Methods

Our method consists of two parts, clothes detection and clothes classification.

3.1 Clothes Detection

Clothes detection is an object detection network that allows the system to crop out the interest region (clothes) from the background. In our case, we want to crop our the bounding box of short top sleeves.

We choose Yolo-v3 [3] as our object detection network because of its high speed and simplicity. We trained the yolo-v3 network on Deep Fashion 2 [2] dataset that consists of 13 clothing categories (Fig [2]).

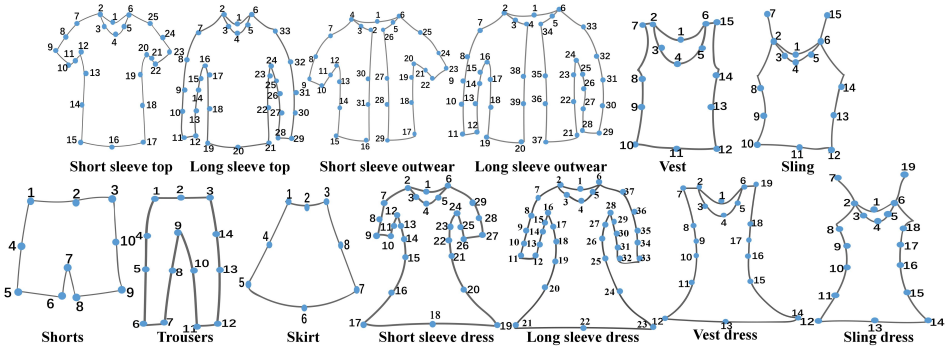

Figure 2: Clothing categories of Deep Fashion 2 dataset.

3.2 Clothes Classification

Clothes classification is the most difficult parts of our work, since we can only access to limited images on the online retail stores (only 3 to 5 images per clothing). Therefore, most learning-based methods (multilayer perceptron, convolutional neural networks, support vector machines) work poorly due to the lack of training images, even with heavy image augmentation techniques.

We would like to point out in real-life scenario, this should not be an issue, since we can easily obtain huge amount of images from the retail stores. However, in this project, we simplify our task to classify short top sleeves into two clothing categories, one with stripes and one without stripes (Fig [3]).

Figure 3: Clothing categories of our classifier. First row (class A) is the clothing without stripes, second row (class B) is the clothing with stripes.

We collect a dataset consist of 20 and 26 images for class A and class B. Since we only have access to small amount of image data, we decide to use hand-craft features with a minimum error classifier.

We notice that if the clothes have stripe, its intensity should vary along a vertical line. We take the mean projection of intensity along the vertical line (Fig [4]). We can see that clothes with stripes have obvious periodic peaks, while clothes without stripes would seem more like random noise. We think this would be a good feature to classify clothes with stripes and without stripes.

We train a minimum error classifier to decide the threshold count of the peaks and the threshold value of each peak. The minimum error classifier achieve 92% accuracy on the 46 images dataset correct,

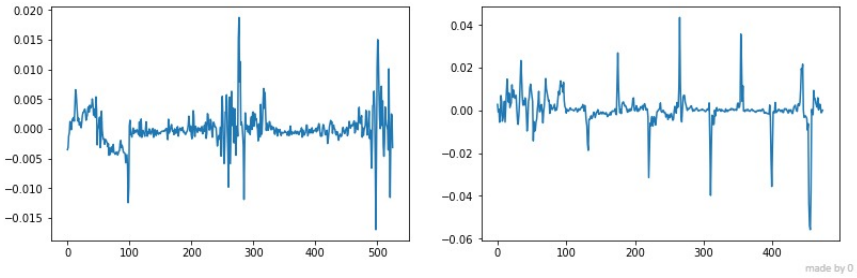

Figure 4: Left is clothing without stripes, right is clothing with stripes.

and it helps us decide the two threshold for the classifier. In summary, value above 0.01 would be considered as a peak, and images more than 0.02 peaks would be classified as a cloth with stripes.

4 Results

Our results are shown in Fig[5]. We can see that clothes detection produces good result, with tight bounding boxes and correct labels. This should be the effort of the huge data from DeepFashion2 [2].

Also, the second cloth is misclassified as clothes with stripes since it contains logo with horizontal-varying intensity, while the dataset we collected consist of mostly pure color clothes.

5 Conclusion and Future Works

In this project, we trained a object detection network that is capable of detect various class of clothing with great accuracy. While we did not actually retrieval clothing suggestion for online retail stores, we think that it would not be difficult if we have access to huge amount of training images.

The future works of this project includes:

1. Extend the clothes type to shorts, trousers, skirts, ... etc.
2. Collect enough image datas so that the classification process works and can give clothing suggestion to users.

Figure 5: Results of our system. (a) Input image, (b) clothes detected images, (c) crop image with classification result.

References

[1] Yannis Kalantidis, Lyndon Kennedy & Li-Jia Li. (2013) "Getting the Look: Clothing Recognition and Segmentation for Automatic Product Suggestions in Everyday Photos".

[2] Yuying Ge, Ruimao Zhang, Lingyun Wu, Xiaogang Wang, Xiaoou Tang & Ping Luo. (2019) "DeepFashion2: A Versatile Benchmark for Detection, Pose Estimation, Segmentation and Re-Identification of Clothing Images".

[3] Joseph Redmon & Ali Farhadi. (2018) "YOLOv3: An Incremental Improvement".