

2nd International Conference on Computer Engineering, Information Science & Application Technology (ICCIA 2017)

# **Outfit Recommendation System Based on Deep Learning**

Ying Huang<sup>a</sup>, Tao Huang<sup>b</sup>

School of Information Engineering, Wuhan University of Technology, Wuhan 430070, China; a969925862@qq.com, b 119072957@qq.com

**Keywords:** outfit recommendation, deep learning, dataset.

**Abstract.** In this paper, we propose an outfit recommendation system based on deep learning. Our goal is to use the system not only to judge an outfit if it is good or not but also to recommend good outfit to users when it is given a pool of cloth items. Our proposed model includes two parts: one is feature extractor based on ResNet-50, and the other is a binary classifier which is to classify the outfits into good ones and bad ones. Since our model is based on deep learning, it is necessary to use huge data to train the model. We collected a dataset which consists of 409,776 outfits with 644,192 items from the famous fashion website called Polyvore.com. With this dataset, we trained our model and the performance of it is over 84%. And our model can also recommend daily outfit to users.

### 1. Introduction

Fashion has a tremendous impact on our society [1]. Clothing is the most important thing that people should consider every day. People want to look good, as we can see that there are lots of online social sites such as Instagram or Facebook where people can show their fashion pictures to the world. Also people like to buy lots of clothes to try different styles to make them look more appealed, which is reflected in the growing online clothing sales, reaching 370 billion dollars in the US by 2017, and 191 billion euros in Europe[2], also over 103 billion RMBs just on one festival day in Taobao.com, an E-commerce website in China.

Since fashion is such a not only popular but also extremely complex topic in our daily life, it is important for us to study about it. However, we cannot make it manually. Recent years more and more researchers started to use computer vision to study on fashion [3, 4, 5, 6]. One of the main task is application domain of fashion recommendation [7,8]. The main focus is to parse clothing from photographs, which is widely used on online shopping sites. In addition, it will make people's life more convenient if there is a system which can suggest good outfit to users based on the clothing that the users already have. An occasion-oriented clothing recommendation is given to suggest the most suitable clothing from the user's own clothing photo album [9], which just put attention to the occasion not the outfit itself.

In this paper, our goal is to judge if an outfit is good or not. In our case, we mainly put attention to the outfit itself to see if the combination of kinds of clothing items is good or bad, rather than consider about the characteristics of the users or the scene behind the users, which makes our system so general that it can be used in every domain. What's more, our aim is to use the system to recommend good outfit to users based on their clothing items. The scenario illustration of our system is shown in Figure 1.

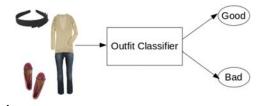


Fig. 1 An scenario illustration of the outfit classifier system



## 2. Related Work

What to wear is the most important thing people care about every day. Also fashion has become a hot topic in recent years. This is shown in the growing interest in fashion-related research.

Firstly, since fashion is a huge topic and it includes huge data, too. So models based on deep learning is the best choice to study fashion. Easy-to-use models [15, 16] are very popular because their computational power of graphical processing unit, also problem-oriented models [10] are welcomed since their performance are pretty high. In our work, we use ResNet-50 as our feature extractor, which is able to extract all information in the image.

Secondly, some outfit related applications such as identifying people based on their outfits then to predict occupation [11], finding outfit similarity [12], recommending good outfits to users [13], and predicting the existed outfit style [14] are the works related to our goal. However, most of the approaches here use annotated data, which is not suitable for general situations. Otherwise, in our case, the model we propose is more general, allowing user to judge the outfit itself.

Thirdly, there are a lot of fashion recommendation systems, like generating outfit according to the user's body color [15] and recommending outfit to user according to the occasion like wedding, dating, etc. [13], both of them use manually annotated data like above. But there are other systems use unannotated data like [5] uses data from Polyvore.com, just the same as ours, to recommend items to users according to their personal preference.

Finally, the attributes of cloth items like color, style [16] and texture [17], etc. are also considered to prove the outfit recommendation system performance good. As for color, O'Donovan P proposes a model that can rate the preference of people to 5 color combinations [18], which can be used in outfit to see if the color of the outfit is proved to be a good combination or not; And the same with style, which is used in [19] to recommend items according style. These attributes of items can be used as a method to prove that the judgement or the recommendation of our outfit is good or not.

#### 3. Dataset

#### 3.1 Dataset Creation.

We used a novel dataset from a clothing-oriented online website Polyvore.com, shown in Figure 2. In this website, user can create an outfit from a pool of items by himself/herself and post the outfit to share with others. Each shared outfit includes information such as outfit name, items in the outfit, similar styles, comments about the outfit, number of how many people like the outfit and so on.

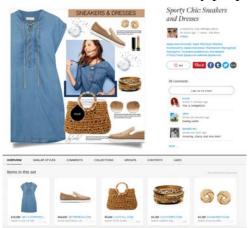


Fig. 2 An example of shared outfit

## 3.2 Training Dataset and Testing Dataset.

In order to train our model, we need training and testing dataset, and in both training and testing dataset, we need good samples and bad samples. Firstly, we select outfits which the number of likes is more than 1 as our good samples. Here we have 21,623 outfits to be good samples. Then, we create bad samples by randomly using the cloth items in good samples.



# 4. Proposed Model

## 4.1 Overview of the Model.

Our model is based on deep network, which mainly includes two parts: feature extractor and binary classifier, the architecture of our model is shown in Figure 3. Firstly, we use the feature extractor to extract all kinds of useful information like color, texture, etc. from the input image, then the extracted information is put into the binary classifier to get "1" or "0", "1" means that the input outfit is good, and "0" means bad.

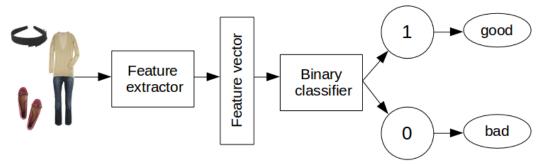


Fig. 3 The architecture of our model

## 4.2 Feature Extractor.

Since the training of the network is huge, we use ResNet-50 [14] as feature extractor in our model. The input of the network is image of cloth item in an outfit. In order to extract the feature of one image of item, firstly, we preprocess image to make the size of the item in the image consistent, then we resize the image into 224×224 pixels. We can get a 2048-float feature vector at the 5th pooling layer of the ResNet-50 for each cloth item in an outfit. In order to train the outfit classifier, we concatenate all the feature vectors of all the items in an outfit into one long feature vector as the input of the binary classifier described below.

# 4.3 Binary Classifier.

In our model, the binary classifier is a 2-layer MLP (Multi-Layer Perceptrons). The input of the network is the long feature vector of an outfit described above, and the output of the network is "0" or "1", here "0" means the outfit is a bad one and "1" means the outfit is good We trained our model with the dataset to form our binary classifier.

## 4.4 Outfit Recommendation.

After we have our outfit classifier model, which can judge if an input outfit is good or bad. Also we can use it to recommend outfit to users. The architecture of the system of outfit recommendation is shown in Figure 4. A pool of cloth items is given to the system, then the system can generate some outfits, input these outfits to the outfit classifier to get the judgement respectively, here we changed the output of the MLP to range from 0 to 1 instead of just "0" or "1". We sort the results of these outfits and choose the first one as our recommendation to the users.

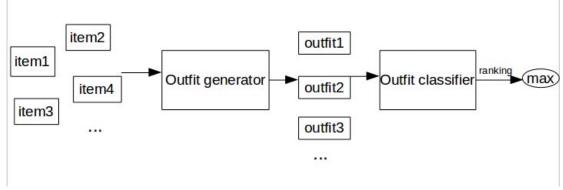


Fig. 4 The architecture of outfit recommendation system



# 5. Experimental Evaluation

#### 5.1 Performance of Outfit Classifier.

We trained our model with the training dataset in section 4. And here we use the testing dataset to test our model. We put these outfits into our model, and we can get the performance of accuracy about 84%, which proves that our model works well.

### 5.2 Performance of Outfit Recommendation.

In our outfit recommendation system, we give the system a pool of cloth items shown in Figure 5. Then the system generates several outfits randomly, and these generated outfits are put into our model to get the judgements ranging from 0 to 1. After sorting these judgements we can get the top 4 good outfits shown in Figure 6.



Fig. 6 The result of outfit recommendation system

## 6. Conclusion

In this work, we presented an outfit recommendation system based on deep learning. We created a novel dataset based on outfits rather than cloth items by selecting suitable outfits from the fashion website named Polyvore.com. We proposed a model that uses ResNet as a feature extractor and a 2-layer MLP as a binary classifier. Since the ResNet is already well trained, we put our attention to the training of our binary classifier. We tried several impossible models and found out that the 4096-fully-connected model performs best. Our model can judge if an outfit is good or bad with high performance. Also it is can be used in the outfit recommendation system, which is given a pool of cloth items it can recommend the users the best outfit generated from the input items.

#### References

- [1]. Simo-Serra E, Fidler S, Moreno-Noguer F, et al. Neuroaesthetics in fashion: Modeling the perception of fashionability[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015: 869-877.
- [2]. F. Magazine. US Online Retail Sales To Reach \$370B By 2017; e191B in Europe. http://www.forbes.com, 2013. [Online; accessed 14-March-2013].
- [3]. Yamaguchi K, Okatani T, Sudo K, et al. Mix and Match: Joint Model for Clothing and Attribute Recognition[C]//BMVC. 2015: 51.1-51.12.



- [4]. Liu Z, Luo P, Qiu S, et al. Deepfashion: Powering robust clothes recognition and retrieval with rich annotations[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016: 1096-1104.
- [5]. Hu Y, Yi X, Davis L S. Collaborative fashion recommendation: a functional tensor factorization approach[C]//Proceedings of the 23rd ACM international conference on Multimedia. ACM, 2015: 129-138.
- [6]. Jagadeesh V, Piramuthu R, Bhardwaj A, et al. Large scale visual recommendations from street fashion images[C]//Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2014: 1925-1934.
- [7]. Yamaguchi K, Berg T L, Ortiz L E. Chic or social: Visual popularity analysis in online fashion networks[C]//Proceedings of the 22nd ACM international conference on Multimedia. ACM, 2014: 773-776.
- [8]. Huang J, Feris R S, Chen Q, et al. Cross-domain image retrieval with a dual attribute-aware ranking network[C]//Proceedings of the IEEE International Conference on Computer Vision. 2015: 1062-1070.
- [9]. Chen H, Gallagher A, Girod B. Describing clothing by semantic attributes [J]. Computer Vision–ECCV 2012, 2012: 609-623.
- [10]. Jammalamadaka N, Minocha A, Singh D, et al. Parsing Clothes in Unrestricted Images[C]//BMVC. 2013, 1: 2.
- [11]. Liu S, Song Z, Liu G, et al. Street-to-shop: Cross-scenario clothing retrieval via parts alignment and auxiliary set[C]//Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. IEEE, 2012: 3330-3337.
- [12]. Wang N, Ai H. Who blocks who: Simultaneous clothing segmentation for grouping images[C]//Computer Vision (ICCV), 2011 IEEE International Conference on. IEEE, 2011: 1535-1542.
- [13]. Liu S, Feng J, Song Z, et al. Hi, magic closet, tell me what to wear![C]//Proceedings of the 20th ACM international conference on Multimedia. ACM, 2012: 619-628.
- [14]. He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016: 770-778.
- [15]. Abadi M, Agarwal A, Barham P, et al. Tensorflow: Large-scale machine learning on heterogeneous distributed systems [J]. arXiv preprint arXiv:1603.04467, 2016
- [16]. Jia Y, Shelhamer E, Donahue J, et al. Caffe: Convolutional architecture for fast feature embedding[C]//Proceedings of the 22nd ACM international conference on Multimedia. ACM, 2014: 675-678.
- [17]. Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015: 3431-3440.
- [18]. Song Z, Wang M, Hua X, et al. Predicting occupation via human clothing and contexts[C]//Computer Vision (ICCV), 2011 IEEE International Conference on. IEEE, 2011: 1084-1091.
- [19]. Vittayakorn S, Yamaguchi K, Berg A C, et al. Runway to realway: Visual analysis of fashion[C]//Applications of Computer Vision (WACV), 2015 IEEE Winter Conference on. IEEE, 2015: 951-958.