

# Images Analysis: Clothing detection, classification and mix-match recommendation

\*Note: Sub-titles are not captured in Xplore and should not be used

1<sup>st</sup> Chau Minh Nguyen  
*chau.nguyenminh4@mail.dcu.ie*  
Dublin City University  
City, Country  
email address

2<sup>nd</sup> Dao Nguyen  
*dept. name of organization (of Aff.)*  
*name of organization (of Aff.)*  
City, Country  
email address

**Abstract**—This document is a model and instructions for  $\LaTeX$ . This and the `IEEEtran.cls` file define the components of your paper [title, text, heads, etc.]. \*CRITICAL: Do Not Use Symbols, Special Characters, Footnotes, or Math in Paper Title or Abstract.

**Index Terms**—component, formatting, style, styling, insert

## I. INTRODUCTION

In the field of fashion e-commerce, clothing image data analysis has many valuable applications on specific demand. One common problem for consumers is that they cannot find the exact items even though they already have the images of them, which often come from social media., sometime is hard to find where to buy the similar items, not to mention other items that often go together with. Hence, acknowledge the consumers needs from their favorite picture is an advantage to provide effective shopping recommendations, which increase interaction on the trade items and give benefits for both consumers and providers. Based on those practical needs, we will separate research question into two main tasks. First, we attend to concentrate on obtaining the method to apply deep learning model on outfit object detection and classification with given clothing images dataset. Then, assuming that the outfit combination from images of dataset are common, we will build a learning model to cluster the combination between items based on different perspectives to provide valuable mix and match recommendations. The project will have two contributions: (1) Develop a deep learning model for clothing object detection and clothing classification. (2) Develop a recommendation model for clothes mix and match.aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa

“LaTeX is a document preparation system and document markup language. LaTeX uses the TeX typesetting program for formatting its output, and is itself written in the TeX macro language. LaTeX is not the name of a particular editing program, but refers to the encoding or tagging conventions that are used in LaTeX documents”.

Identify applicable funding agency here. If none, delete this.

## II. RELATED WORK

Among a lot of algorithms that could be applied for image analysis, Convolutional Neural Networks (CNNs) were widely used in object recognition and classification because of its effective performance on large amount of training data, according to Li Deng and Dong Yu [1]. Since its first present in 90s, many researchers have been presenting different contributions on improving CNNs models. Alex Krizhevsky et al [2] introduced how to optimized a CNN model with a large learning capacity by different methods, including maximizing the size of the networks that can be trained by spreading the net from one to two GPU, which lead to the error rate reduction. They also recognized the relationship between number of convolutional layer- the depth and the networks performance. The depth aspect of CNNs architecture then was focused by [3]. Increasing number of convolutional layers with small convolution filters allowed the architecture to archive excellent performance and accuracy. In our practicum, we expect to observe those methods in our CNN model to solve the problems of image detection and classification. For the recommender system development, we acquired the fundamental knowledge of recommendation such as basic approaches of collaborative, content-based, knowledge-based filtering and key differences between these techniques [4], how to build recommenders using Python [5]. Among number of fashion recommendation methods that proposed, Tomoharu Iwata presented a three-stage method including: detecting top and bottom region from full-body photographs, using topic model to learn coordinate information, recommending a bottom (top) that has the closest topic proportions to those of the given top (bottom) [6]. M. Caron et al introduced the concept of using unsupervised training on convolutional neural networks with [4]. Our project is planned to absorb the methods and selective apply on our large-scale clothes dataset: DeepFashion [7].

### III. DATA AND RESEARCH METHOD

#### A. Data

Throughout this paper, we will use DeepFashion [8], a large-scale clothes dataset that has been effectively used in numerous research projects. It includes more than 300,000 images that were labelled into 50 categories, 1000 attributes, and clothing landmarks. Images of the dataset was collected from shopping websites and google images. The removal of duplicates and low resolution, image quality, or whose dominant objects are irrelevant to clothes was performed. The dataset is currently available for research community.

#### B. Research Method

For object detection and classification: we will apply different *CNNs* models and accomplish them with three steps process Training Validating Testing to improve the performance and the accuracy of each models. The results of FashionNet a novel deep learning structure [1], [3] will compared with our results considering the performance and the accuracy. For mix and match recommendation: An unsupervised learning recommender system is expected to develop. Top N accuracy will be used as the rate of recommending. Training and testing dataset will be generated by splitting the top and the bottom of the outfit from full-body outfit images randomly. The n-best accuracy represents the rate of recommending the correct bottom (top) photograph with n recommendations given a test top (bottom) photograph, following.

sdglksdglkj sdgjskldjg sdg sjdlkg j sg sdg sdg sdg sg sdg  
sdg sdg sdg sldk jsfd;kgjsklgj sldk jsldkgj ;sdkl gj  
4643 43643 6436 43 6

$$\sum_{n=1}^{100} (\alpha + n) \quad (1)$$

### IV. EXPECTED OUTCOME

As result, we would like to present a three-stage recommender framework. In stage one, we will propose a CNN model that could detect the bounding boxes that capture the clothing items in given full-body outfit images and classify them into labelled categories/ attribute with acceptable accuracy. A topic model based on item categories, attribute and relationship between items will be presented in stage two. In final stage, we will describe a mix and match recommendation method that uses the topic model.

### REFERENCES

- [1] L. Deng and D. Yu, "Deep learning: methods and applications," *Foundations and trends in signal processing*, vol. 7, no. 3–4, pp. 197–387, 2014.
- [2] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.
- [3] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
- [4] A. Felfernig, M. Jeran, G. Ninaus, F. Reinfrank, S. Reiterer, and M. Stettinger, "Basic approaches in recommendation systems," in *Recommendation Systems in Software Engineering*. Springer, 2014, pp. 15–37.
- [5] M. Caron, P. Bojanowski, A. Joulin, and M. Douze, "Deep clustering for unsupervised learning of visual features," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 132–149.
- [6] Z. Liu, P. Luo, S. Qiu, X. Wang, and X. Tang, "Deepfashion: Powering robust clothes recognition and retrieval with rich annotations," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 1096–1104.
- [7] A. M. Rashid, I. Albert, D. Cosley, S. K. Lam, S. M. McNee, J. A. Konstan, and J. Riedl, "Getting to know you: learning new user preferences in recommender systems," in *Proceedings of the 7th international conference on Intelligent user interfaces*, 2002, pp. 127–134.
- [8] T. Iwata, S. Wanatabe, and H. Sawada, "Fashion coordinates recommender system using photographs from fashion magazines," in *IJCAI*, vol. 22, no. 3. Citeseer, 2011, p. 2262.