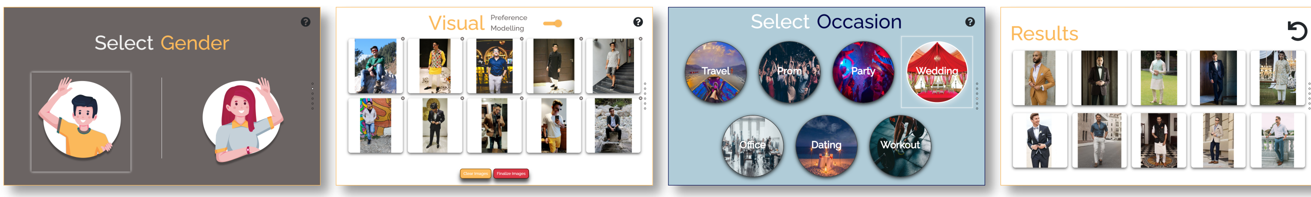
Fashionist: Personalising Outfit Recommendation for Cold-Start Scenarios

Dhruv Verma*, Kshitij Gulati*, Vasu Goel, Rajiv Ratn Shah
Indraprastha Institute of Information Technology, Delhi, India
{dhruv17046,kshitij17162,vasu18322,rajivratn}@iiitd.ac.in



(a) Select Gender

(b) Upload outfit images for visual preference modelling

(c) Select desired occasion for outfit recommendation

(d) Display recommendations based on *gender*, *occasion* and *personal preferences*

Figure 1: User interface of the Fashionist web application.

ABSTRACT

With the proliferation of the online fashion industry, there have been increased efforts towards building cutting-edge solutions for personalising fashion recommendation. Despite this, the technology is still limited by its poor performance on new entities, i.e. the cold-start problem. We attempt to address the cold-start problem for new users, by leveraging a novel visual preference modelling approach on a small set of input images. Additionally, we describe our proposed strategy to incorporate the modelled preference in occasion-oriented outfit recommendation. Finally, we propose *Fashionist*: a real-time web application to demonstrate our approach enabling personalised and diverse outfit recommendation for cold-start scenarios. Check out https://youtu.be/kuKgPCkoPy0 for demonstration.

CCS CONCEPTS

• **Information systems** → **Personalization**.

KEYWORDS

personalised outfit recommendation; cold-start problem; fashion concept prediction; multi-task learning

ACM Reference Format:
Dhruv Verma*, Kshitij Gulati*, Vasu Goel, Rajiv Ratn Shah. 2020. Fashionist: Personalising Outfit Recommendation for Cold-Start Scenarios. In *Proceedings of the 28th ACM International Conference on Multimedia (MM '20), October 12–16, 2020, Seattle, WA, USA.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3394171.3414446

*These authors contributed equally to this work.

ACM ISBN 978-1-4503-7988-5/20/10.
https://doi.org/10.1145/3394171.3414446

1 INTRODUCTION

Today, online social media is an integral part of peoples' lives. With its growing influence, people have a strong desire to remain connected. They share almost every aspect of their lives, including the outfits they are wearing. The notion that an individual's dressing style speaks volumes about his/her personality has led to an increased consciousness amongst people for the appropriateness of an outfit in a particular context. E-commerce for fashion is booming, but consumers still face issues while selecting fashion outfits that are appropriate in context and suit their liking. With this scenario in place, personalised fashion recommendation has gathered increasing interest from both academia and industry.

Various approaches have been adopted to personalise outfit recommendation [2, 3, 6, 12, 16]. However, most of these approaches fail to address the cold-start problem [7]. It is not only relevant to new users but also for users who prefer to experience the recommendation service in 'guest mode' or without getting profiled. Solving this issue forms the basis for our work. The optimal recommendation for a particular scenario would necessarily be suitable to that context while keeping the visual elements of fashion (e.g. colour, texture, style) as per the user's liking. These visual elements can be better characterised in terms of clothing *categories* (e.g. *tops:* t-shirt, blouse; *bottom:* pants, leggings) and *attributes* (e.g. *neckline:* round, halter; *sleeve length:* short, three fourth). Therefore, our approach focuses on the robust detection of these visual elements.

Currently, none of the publicly available fashion datasets [4, 8, 9, 14] has adequate representation for user-context as well as occasion. Hence, we curated our occasion-oriented fashion knowledge dataset constituting images scraped from Instagram and Pinterest, and manually annotated as described in [15]. In this paper, we attempt to tackle the cold-start problem for new users by taking a few outfit images depicting their fashion tastes as input. We extract visual fashion elements from these images and model them as preferences for personalising the outfit recommendation.

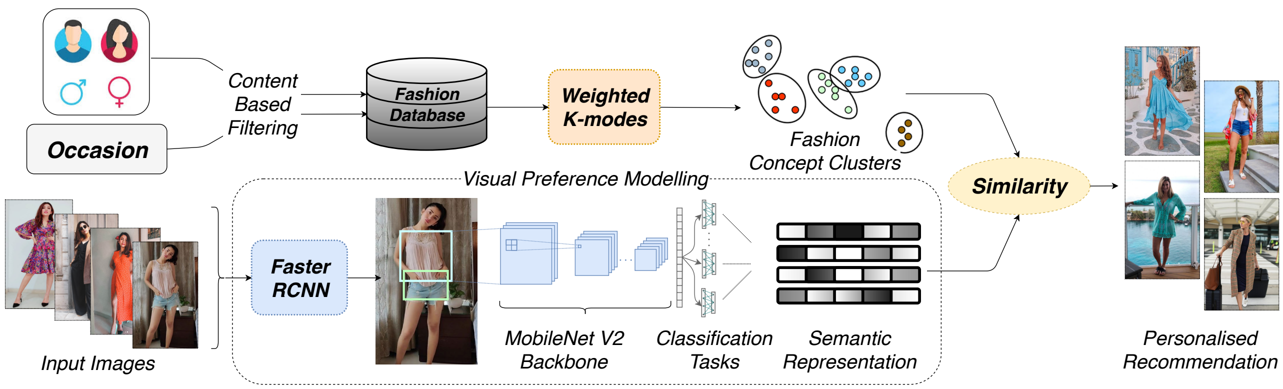


Figure 2: The proposed pipeline for personalised fashion recommendation in cold-start scenarios.

2 METHODOLOGY

The combined occurrences of visual fashion elements (clothing categories and attributes) in the input outfit images are a representation of the user's preference for different fashion choices. Therefore, the first portion of our work describes an approach to model the user's preferences based on these visual fashion elements. The second portion describes our strategy to incorporate these extracted visual preferences into outfit recommendation. We present the proposed pipeline in Figure 2.

2.1 Visual Preference Modelling

Fashion outfits are a combination of multiple clothing items (like tops and bottoms or full body dresses). We use a state of the art object detection technique; Faster RCNN [11] fine-tuned on ModaNet [18] dataset for robust detection of these clothing items. After extracting different clothing items as regions of interest, we apply *multi-task learning*, which has been used in various computer vision applications owing to its remarkable information-sharing capabilities [17], to extract category and attribute knowledge from these regions of interest. The multi-task classifier uses a MobileNet V2 [5] backbone, trained on DeepFashion [9] dataset and fine-tuned on our curated dataset. Hence, we can represent each region of interest in terms of category and attribute labels semantically similar to our dataset. Thereafter, the representations of individual regions of interest are pooled together, forming a semantic representation for every outfit in the set of input images. Table 1 shows the performance of the proposed multi-task learning approach evaluated against state of the art techniques for category and attribute prediction.

2.2 Personalised Outfit Recommendation

After extracting the semantic representations (as described in Section 2.1) for the given input images, our goal is to fuse this extracted preference knowledge into occasion-oriented outfit recommendation. We start with applying content-based filtering on the dataset to retrieve outfits relevant to the desired occasion and gender. These retrieved outfits represent various fashion concepts like style and seasons. We use *feature-weighted k-modes* clustering [1, 13] (with higher weights to the category information) to group the retrieved outfits into fashion concept clusters.

Each cluster is weighted according to its similarity with the set of input images (similarity is calculated between the semantic representation of cluster centroids and outfits). Subsequently, we use these cluster weights to decide the number of outfits to be drawn from that cluster in the final result, hence ensuring both diversity and incorporation of preference even in cold-start scenarios.

Table 1: Comparison with state of the art architectures

Model	Accuracy	
	Category	Attributes
FashionNet [9]	67.33%	65.6%
FashionKE [10]	**73.95%**	69.59%
MobileNet MTL (Our Method)	73.5%	**89.1%**

3 DEMONSTRATION

For demonstrating the proposed methodology, we built *Fashionist*: a web-based application enabling personalised outfit recommendation for new users. In this web application, users have to select their gender (see Figure 1a), upload a set of ten images representing their fashion tastes (see Figure 1b), select the occasion they wish to dress for (see Figure 1c) and Voila! Top ten occasion relevant outfits that incorporate the users' personal preferences (see Figure 1d) are recommended[1] . The users also have the option of skipping the visual preference modelling step and proceed without uploading images to retrieve outfits drawn uniformly from each fashion concept cluster. The implementation of Fashionist follows a server-client architecture, supporting multiple client requests and cross-platform support for users to use it anywhere and anytime.

4 LIMITATIONS AND FUTURE WORK

While the server implementation (with GPU support) helps achieve real-time performance for our application, it might pose privacy concerns to users uploading their own pictures. Client-side processing with mobile-deployable architectures can potentially address these concerns. Additionally, instilling certain high-level fashion concepts like body shape, seasonal variation and cultural dependence of certain clothing styles can help improve personalisation. We plan to investigate these avenues in the future.

[1] The recommendations follow no particular order or ranking.

REFERENCES

[1] Anil Chaturvedi, Paul E Green, and J Douglas Caroll. 2001. K-modes clustering. *Journal of classification* 18, 1 (2001), 35–55.

[2] Wen Chen, Pipei Huang, Jiaming Xu, Xin Guo, Cheng Guo, Fei Sun, Chao Li, Andreas Pfadler, Huan Zhao, and Binqiang Zhao. 2019. POG: Personalized Outfit Generation for Fashion Recommendation at Alibaba iFashion. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. ACM, Anchorage AK USA, 2662–2670. https://doi.org/10.1145/3292500.3330652

[3] Xu Chen, Hanxiong Chen, Hongteng Xu, Yongfeng Zhang, Yixin Cao, Zheng Qin, and Hongyuan Zha. 2019. Personalized Fashion Recommendation with Visual Explanations based on Multimodal Attention Network: Towards Visually Explainable Recommendation. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, Paris France, 765–774. https://doi.org/10.1145/3331184.3331254

[4] Xintong Han, Zuxuan Wu, Yu-Gang Jiang, and Larry S. Davis. 2017. Learning Fashion Compatibility with Bidirectional LSTMs. In *Proceedings of the 2017 ACM on Multimedia Conference - MM '17*. ACM Press, Mountain View, California, USA, 1078–1086. https://doi.org/10.1145/3123266.3123394

[5] Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam. 2017. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. *arXiv:1704.04861 [cs]* (April 2017). http://arxiv.org/abs/1704.04861 arXiv: 1704.04861.

[6] Yang Hu, Xi Yi, and Larry S. Davis. 2015. Collaborative Fashion Recommendation: A Functional Tensor Factorization Approach. In *Proceedings of the 23rd ACM international conference on Multimedia - MM '15*. ACM Press, Brisbane, Australia, 129–138. https://doi.org/10.1145/2733373.2806239

[7] Blerina Lika, Kostas Kolomvatsos, and Stathes Hadjiefthymiades. 2014. Facing the cold start problem in recommender systems. *Expert Systems with Applications* 41, 4 (March 2014), 2065–2073. https://doi.org/10.1016/j.eswa.2013.09.005

[8] Si Liu, Jiashi Feng, Zheng Song, Tianzhu Zhang, Hanqing Lu, Changsheng Xu, and Shuicheng Yan. 2012. Hi, magic closet, tell me what to wear!. In *Proceedings of the 20th ACM international conference on Multimedia - MM '12*. ACM Press, Nara, Japan, 619. https://doi.org/10.1145/2393347.2393433

[9] Ziwei Liu, Ping Luo, Shi Qiu, Xiaogang Wang, and Xiaoou Tang. 2016. DeepFashion: Powering Robust Clothes Recognition and Retrieval with Rich Annotations. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, Las Vegas, NV, USA, 1096–1104. https://doi.org/10.1109/CVPR.2016.124

[10] Yunshan Ma, Xun Yang, Lizi Liao, Yixin Cao, and Tat-Seng Chua. 2019. Who, Where, and What to Wear?: Extracting Fashion Knowledge from Social Media. In *Proceedings of the 27th ACM International Conference on Multimedia*. ACM, Nice France, 257–265. https://doi.org/10.1145/3343031.3350889

[11] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. 2016. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *arXiv:1506.01497 [cs]* (Jan. 2016). http://arxiv.org/abs/1506.01497 arXiv: 1506.01497.

[12] Dikshant Sagar, Jatin Garg, Prarthana Kansal, Sejal Bhalla, Rajiv Ratn Shah, and Yi Yu. 2020. PAI-BPR: Personalized Outfit Recommendation Scheme with Attribute-wise Interpretability. *arXiv:2008.01780 [cs]* (Aug. 2020). http://arxiv.org/abs/2008.01780 arXiv: 2008.01780.

[13] S.Aranganayagi and K.Thangavel. 2009. Improved K-Modes For Categorical Clustering Using Weighted Dissimilarity Measure. (March 2009). https://doi.org/10.5281/ZENODO.1070405 Publisher: Zenodo.

[14] Xuemeng Song, Fuli Feng, Jinhuan Liu, Zekun Li, Liqiang Nie, and Jun Ma. 2017. NeuroStylist: Neural Compatibility Modeling for Clothing Matching. In *Proceedings of the 2017 ACM on Multimedia Conference - MM '17*. ACM Press, Mountain View, California, USA, 753–761. https://doi.org/10.1145/3123266.3123314

[15] Dhruv Verma, Kshitij Gulati, and Rajiv Ratn Shah. 2020. Addressing the Cold-Start Problem in Outfit Recommendation Using Visual Preference Modelling. *arXiv:2008.01437 [cs]* (Aug. 2020). http://arxiv.org/abs/2008.01437 arXiv: 2008.01437.

[16] Ruiping Yin, Kan Li, Jie Lu, and Guangquan Zhang. 2019. Enhancing Fashion Recommendation with Visual Compatibility Relationship. In *The World Wide Web Conference on - WWW '19*. ACM Press, San Francisco, CA, USA, 3434–3440. https://doi.org/10.1145/3308558.3313739

[17] Yu Zhang and Qiang Yang. 2018. A Survey on Multi-Task Learning. *arXiv:1707.08114 [cs]* (July 2018). http://arxiv.org/abs/1707.08114 arXiv: 1707.08114.

[18] Shuai Zheng, Fan Yang, M. Hadi Kiapour, and Robinson Piramuthu. 2019. ModaNet: A Large-Scale Street Fashion Dataset with Polygon Annotations. *arXiv:1807.01394 [cs]* (April 2019). http://arxiv.org/abs/1807.01394 arXiv: 1807.01394.