

Transferring a Semantic Representation for Person Re-Identification and Search (Supplementary Material)

Zhiyuan Shi, Timothy M. Hospedales, Tao Xiang
Queen Mary, University of London, London E1 4NS, UK
{z.shi,t.hospedales,t.xiang}@qmul.ac.uk

1. Details on Supervision

In Sec. 5.2 of the paper, under the heading “Settings”, we have briefly mentioned that the supervision used for our model varies across the strongly and weakly annotated auxiliary sets. We now give the details on the different types of annotations and corresponding supervision used in our experiments in Table 1.

	Full / Fashion [6]			Weak / Clothing [1]	
Factor	Color	Category	BG	Texture	BG
FGP	Strong	Strong	$L_k = 0$	Weak	None
BGP	None	$L_k = 0$	None		

Table 1: Different types of supervision used by our models depend on how different attributes of each auxiliary dataset are annotated.

In our auxiliary datasets, there are various types of attributes (modelled as factors in our model) and each can be annotated in different ways (denoted as L_k , see Sec. 3.2 of the paper). More specifically, there are three factor types: Colour (12 factors), Category (22), Texture (6) and Background (BG) (20). The number of factors depend on the actual annotation from [6, 1]. These factors may be supervised strongly (L_k given by pixel level annotation), weakly (L_k image level annotation) or not at all (All $L_k = 1$).

Given strongly supervised data (Colourful-Fashion [6]), all superpixels/patches can be categorised into foreground patch (FGP) and background patch (BGP) (Table 1 rows). Foreground patches cannot contain background factors ($L_k = 0$). Colour and category data are strongly annotated so the learning of these factors are strongly supervised for the foreground (Strong). Colour factors can also occur on the background, so these are not supervised (None). Meanwhile, Category factors (e.g., shirt) cannot occur on the background, so $L_k = 0$ here. Finally, any background factors are free to be used on any background patch (None).

Given the weakly supervised dataset (Clothing-Attribute [1]), we do not know the location of foreground background

pixels, so there is no patch-type breakdown. Therefore the texture factors are weakly supervised (Weak); meanwhile the background factors are not supervised.

2. Re-identification Performance Measured by CMC Curves

In Table 1 and Table 2 (Sec. 5.2) of the submitted main paper, the proposed methods are compared with existing unsupervised and supervised learning approaches respectively. Due to space limitation, only the matching accuracy @ rank [1, 5, 10, 20] are reported. In this supplementary material, we provide the CMC curves over rank 1 to 25 obtained by different compared methods on the three datasets. Fig. 1 shows the CMC curves of the compared unsupervised learning based methods, including SDC [13], GTS [9], SDALF [3] and SDC_Final (eSDC) [13]. Fig. 2 compares the supervised learning based methods, including MLF [14], KISSME [4], LADF [5], LF [7], SCNCD_Final [11], PatMatch [12], SalMatch [12], KML [10] and MLF_Final [14]. Note that the CMC curves of some compared methods (i.e. KISSME [4], KML [10], SDC [13] and SDALF [3]) on some datasets are obtained based on our own implementation when their CMC curves were not reported in previous works. In addition, the CMC curves for some methods are not included in these plots if previous work provides neither the CMC curves nor the code for us to implement.

3. Per-query Person Search Results

In Fig. 4 of the submitted main paper, we reported the average precision-recall (PR) curves of all queries for various person search methods. Here we provide more details on their person search performance on the VIPeR-Tag [8] dataset in the form of the precision-recall (PR) curve of each query. There are 200 random queries used on the PETA [2] dataset; it is thus not possible to present the per-query PR curves on PETA here.

These per-query PR curves on VIPeR-Tag are shown in

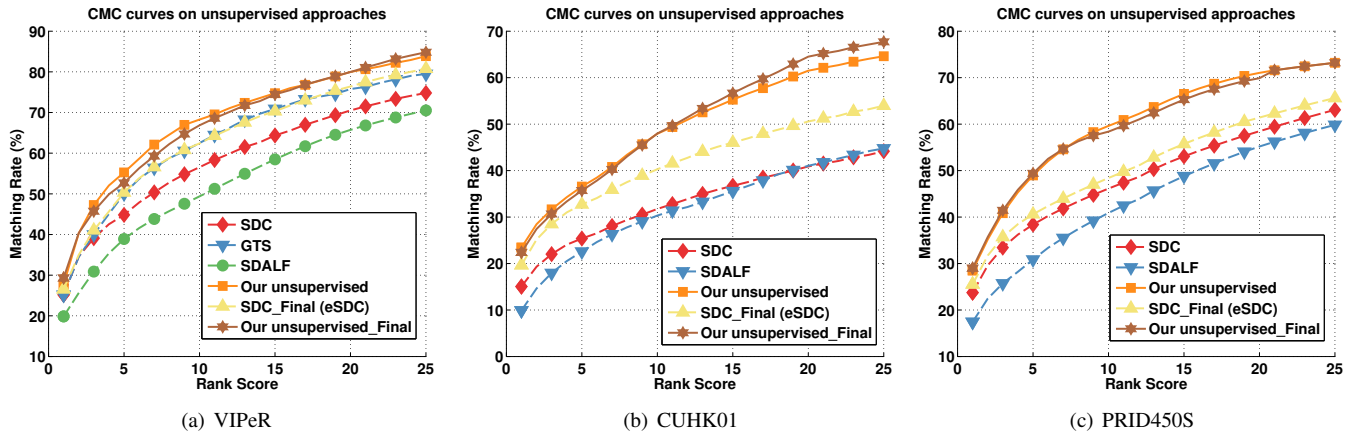


Figure 1: CMC comparison of unsupervised learning based approaches.

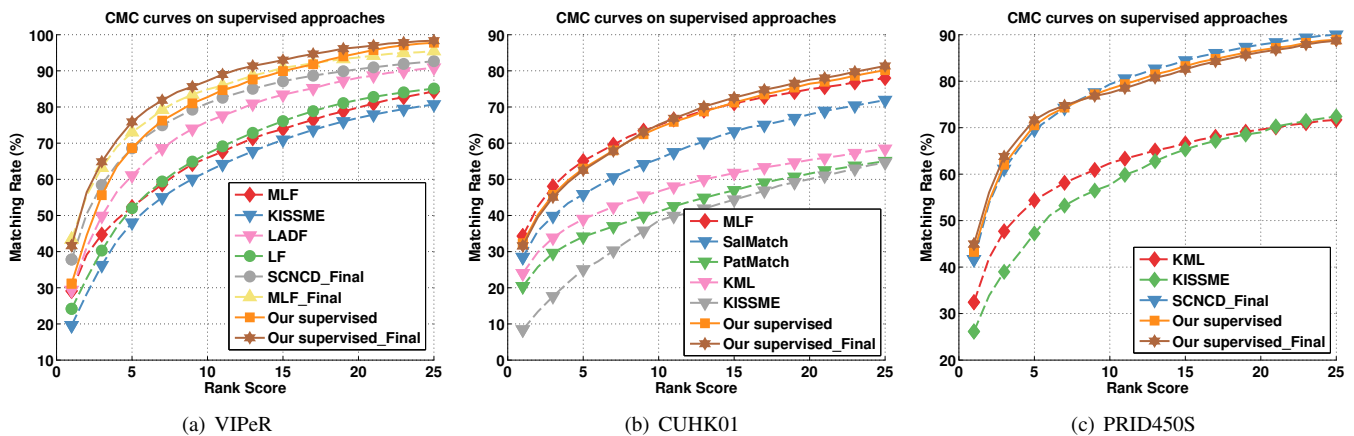


Figure 2: CMC comparison of supervised learning based approaches.

Fig. 3. Note that the queries used on this dataset as defined in [8] have different attribute names as those in our auxiliary Colourful-Fashion dataset. For examples, the attributes ‘UpperGarment’ and ‘LowerGarment’ appear in 13 out of the 15 queries. However, much finer-grained attributes are used in the Colourful-Fashion dataset; for instance, for ‘UpperGarment’, we have ‘blazer’, ‘T-shirt’, ‘blouse’, and ‘sweater’. These finer-grained attributes are thus merged to form the two coarser-grained attributes on VIPeR-Tag before person search.

References

- [1] H. Chen, A. Gallagher, and B. Girod. Describing clothing by semantic attributes. In *ECCV*, 2012. 1
- [2] Y. Deng, P. Luo, C. C. Loy, and X. Tang. Pedestrian attribute recognition at far distance. In *ACM Multimedia*, 2014. 1
- [3] M. Farenzena, L. Bazzani, A. Perina, V. Murino, and M. Cristani. Person re-identification by symmetry-driven accumulation of local features. In *CVPR*, 2010. 1
- [4] M. Koestinger, M. Hirzer, P. Wohlhart, P. M. Roth, and H. Bischof. Large scale metric learning from equivalence constraints. In *CVPR*, 2012. 1
- [5] Z. Li, S. Chang, F. Liang, T. S. Huang, L. Cao, and J. R. Smith. Learning locally-adaptive decision functions for person verification. In *CVPR*, June 2013. 1
- [6] S. Liu, J. Feng, C. Domokos, H. Xu, J. Huang, Z. Hu, and S. Yan. Fashion parsing with weak color-category labels. *Multimedia, IEEE Transactions on*, 2014. 1
- [7] S. Pedagadi, J. Orwell, S. Velastin, and B. Boghossian. Local fisher discriminant analysis for pedestrian re-identification. In *CVPR*, 2013. 1
- [8] R. Satta, G. Fumera, and F. Roli. People search with textual queries about clothing appearance attributes. In *Person Re-Identification*. Springer, 2014. 1, 2
- [9] H. Wang, S. Gong, and T. Xiang. Unsupervised learning of generative topic saliency for person re-identification. In *BMVC*, 2014. 1

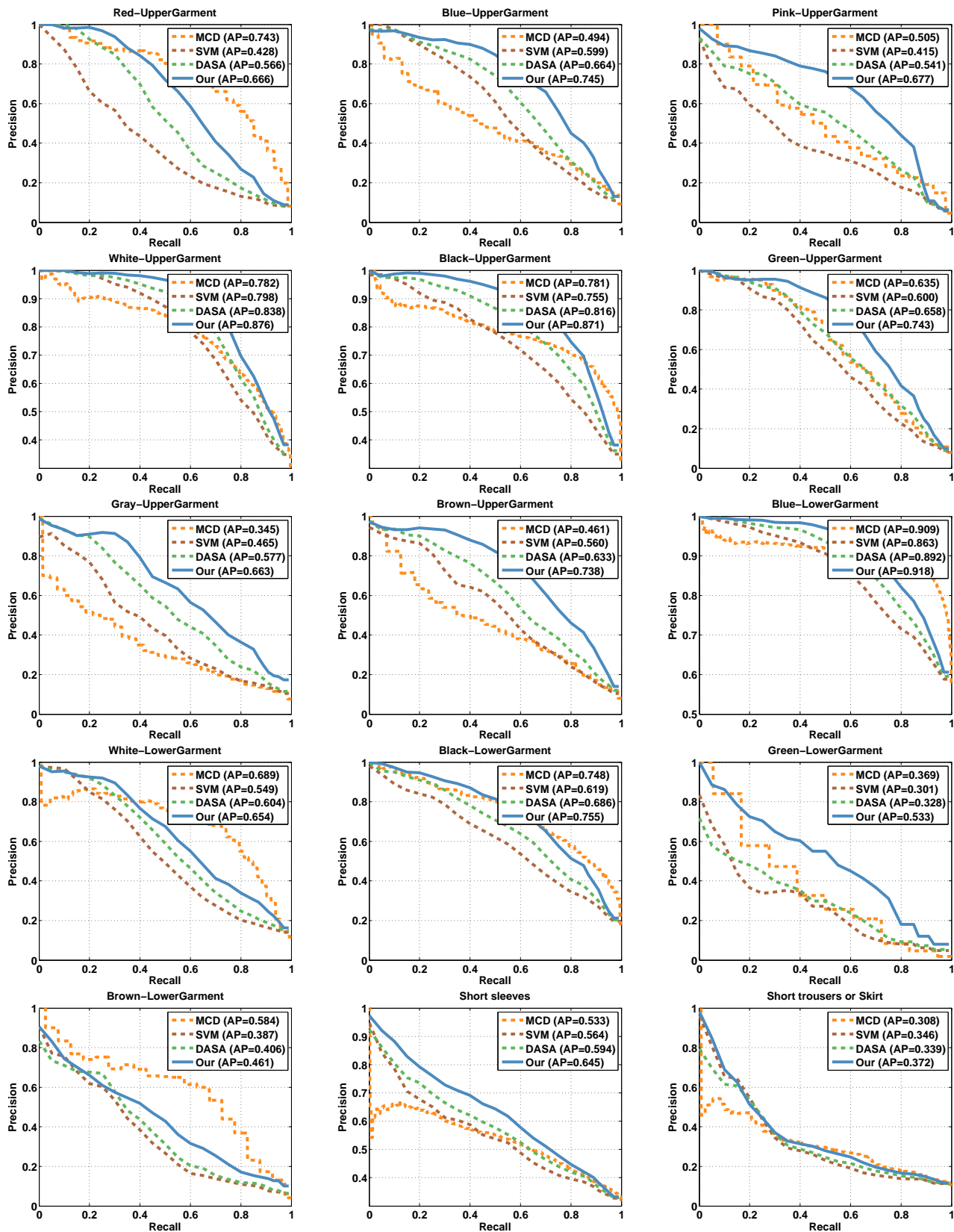


Figure 3: Person search performance on each object-attribute query

- [10] F. Xiong, M. Gou, O. Camps, and M. Sznajder. Person re-identification using kernel-based metric learning methods. In *ECCV*, 2014. [1](#)
- [11] Y. Yang, J. Yang, J. Yan, S. Liao, D. Yi, and S. Li. Salient color names for person re-identification. In *ECCV*, 2014. [1](#)
- [12] R. Zhao, W. Ouyang, and X. Wang. Person re-identification by saliency matching. In *ICCV*, 2013. [1](#)
- [13] R. Zhao, W. Ouyang, and X. Wang. Unsupervised saliency learning for person re-identification. In *CVPR*, 2013. [1](#)
- [14] R. Zhao, W. Ouyang, and X. Wang. Learning mid-level filters for person re-identification. In *CVPR*, 2014. [1](#)