

Image Based Virtual Try-on Network from Unpaired Data Supplementary Material

Assaf Neuberger Eran Borenstein Bar Hilleli Eduard Oks Sharon Alpert
Amazon Lab126

{neuberger, eran, barh, oksed, alperst}@amazon.com

A Additional qualitative results

In the figures below, we include additional examples of our O-VITON virtual try-on method, both for a single garment (Fig.1) and two garments scenarios (Fig.2).



Figure 1: Our O-VITON virtual try-on results for pants: Each row shows the reference garment rendered on the query image for with (online) and without (feed-forward) the online optimization step. The feed-forward approach is able to accurately render solid patterns, however, for more complex textures (e.g. dots, stripes, etc.) the online texture optimization approach produces far better results.



Figure 2: Qualitative results for rendering two garments simultaneously: Our method effectively renders natural looking boundaries between the garments.

B Ablation study for spatial perturbations during shape generator train phase

In Sec.3.2.2 we described the train phase of the shape generation network $G_{shape}, D_{shape}, E_{shape}$. This phase includes an augmentation step that decouples the shape and pose of the body from the garment’s silhouette. This is enable the network to deal with large shape variations between the query and reference images. We apply a method similar to [1], that spatial perturbs each slice $s(\cdot, \cdot, c)$ of the segmentation map s , using a random affine transformation. Here we add an ablation study that demonstrates the contribution of this step using both qualitative (Fig.3) and quantitative results (Table 1).

	FID ↓	IS ↑	Human ↓
Tops			
O-VITON (w.o. SP)	35.95	2.87 ± 0.05	$94\% \pm 2\%$
O-VITON	16.63	3.02 ± 0.07	-
Single garment			
O-VITON (w.o. SP)	27.72	3.28 ± 0.07	$81\% \pm 2\%$
O-VITON	20.47	3.61 ± 0.09	-
Two garments			
O-VITON (w.o. SP)	30.52	3.42 ± 0.12	$95\% \pm 1\%$
O-VITON	28.52	3.51 ± 0.08	-

Table 1: Quantitative evaluation for the spatial perturbation (SP) step: The columns (left to right) show the results for the FID, IS and Human study scores respectively, for the scenarios of tops, single and two garment/s. A significant importance is shown with the spatial perturbation (SP) step.



Figure 3: The importance of the spatial perturbation step: using this step the network is able to better deal with large pose variations between the query and reference images.

C Additional failure cases

Here we present additional failure cases of two-garment virtual try-on that were mainly caused by significant pose differences.



Figure 4: Failure cases of two-garments virtual try-on.

References

- [1] Amit Raj, Patsorn Sangkloy, Huiwen Chang, Jingwan Lu, Duygu Ceylan, and James Hays. Swapnet: Garment transfer in single view images. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 666–682, 2018. 3