Supplementary Materials for GC-VTON: Predicting Globally Consistent and Occlusion Aware Local Flows with Neighborhood Integrity Preservation for Virtual Try-on

A. Additional Qualitative Results

In this document, we further showcase the qualitative results of our method compared to the existing SOTA methods including VITON-HD [1], HR-VTON [4], and FS-VTON [3]. These results will serve as a further proof of the effectiveness of our method in various challenging scenarios.

In Fig 2 we show the advantages of occlusion handling around hand regions. It is clear that while other methods typically generate distortions or washed away textures around hand occlusions, GC-VTON successfully manages to preserve texture details. In Fig 3 we show the challenging tucked-in style results where most methods fail to properly preserve textures and run into squeezing artifacts. Our method on the other hand effectively guards against these scenarios – thanks to garment-garment occlusion handling and NIPR loss.

Finally, in Fig 4 we show the garment details and style preservation capabilities of GC-VTON. As other methods clearly fail to preserve key garment attributes, our method consistently generates realistic try-on results where all the components from the original garment are retained.

B. NIPR vs Second Order Constraint Qualitative Results

In Fig 7 of the main text, we show examples where a network trained with second-order constraint (SO) [2] fails to prevent squeezing and stretching artifacts. In contrast, model trained with our NIPR loss is free of such artifacts. Here in Fig 1, we present a few more such examples to strengthen the case for NIPR loss. The results are presented in a paired fashion where the goal is to warp the garment on the same person that is originally wearing it. This is done in order to have a target warp reference to compare the warps with SO and NIPR.

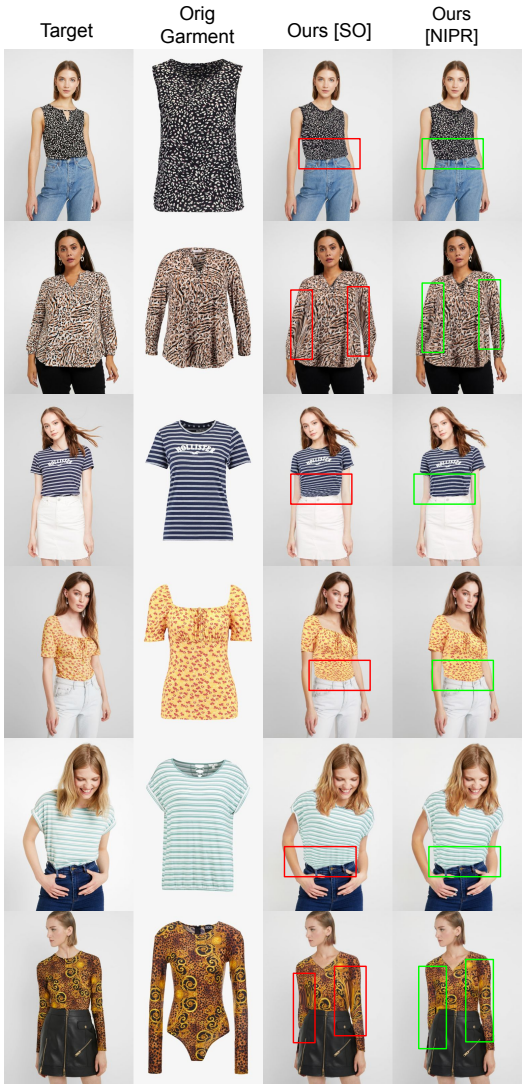

Figure 1. Qualitative comparison of NIPR vs second-order constraint (SO). NIPR is clearly better at warding off stretching and squeezing artifacts.

References

[1] Seunghwan Choi, Sunghyun Park, Minsoo Lee, and Jaegul Choo. Viton-hd: High-resolution virtual try-on via misalignment-aware normalization. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recog-*

Figure 2. Qualitative comparison of GC-VTON (ours) with existing benchmarks for hands related occlusions. The results show that other methods constantly run into generating artifacts around the regions occluded by hands. Our method is free of these discrepancies and generates reasonable warps in such challenging scenarios.

Figure 3. Qualitative comparison of GC-VTON (ours) with existing benchmarks for tucked-in style. Tucked-in style is challenging as some portion of upper garment is obstructed by bottom garment. Typically, existing methods warp the lower regions of upper garment to match the boundaries of the garments. This results in distorted regions around the boundaries. With our occlusion handling, we effectively resolve these issues.

nition, pages 14131–14140, 2021. 1

[2] Yuying Ge, Yibing Song, Ruimao Zhang, Chongjian Ge, Wei Liu, and Ping Luo. Parser-free virtual try-on via distilling appearance flows. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 8485–8493, 2021. 1

[3] Sen He, Yi-Zhe Song, and Tao Xiang. Style-based global appearance flow for virtual try-on. In *Proceedings of*

Figure 4. Qualitative comparison of GC-VTON (ours) with existing benchmarks for texture/style preservation capabilities. Global boundary misalignment, texture washway, unrealistic hair, and missing key garment attributes makes the existing solution unsuitable for complex scenarios. Our method is able to retain all the key garment attributes and preserves texture integrity to generate realistic results.

the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 3470–3479, 2022. 1

[4] Sangyun Lee, Gyojung Gu, Sunghyun Park, Seunghwan Choi, and Jaegul Choo. High-resolution virtual try-on with misalignment and occlusion-handled conditions. In *European Conference on Computer Vision*, pages 204–219. Springer, 2022. 1