

# Fill in Fabrics: Body-Aware Self-Supervised Inpainting for Image-Based Virtual Try-On

Hasib Zunair, Yan Gobeil, Samuel Mercier, A. Ben Hamza

## Motivation

- Existing virtual try-on methods warp clothing item to fit human body and fuse warped clothing with human for synthesis.
- Loss functions are not perceptually motivated
- No robustness test
- Does not leverage context

## Failure Cases

- Complex **poses** and **shapes** of person
- Synthesize **skin color**
- **Occluded** body parts (e.g. long sleeve clothing)
- **Logo, texture** and **embroidery** of clothing.

Types of Pose



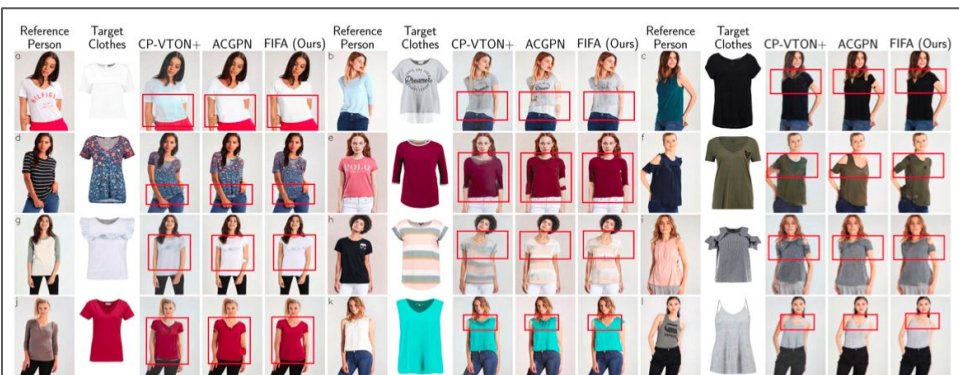
easy medium hard

These failure cases lead to blurry, unrealistic outputs as well as artifacts.



# Fill in Fabrics: Body-Aware Self-Supervised Inpainting for Image-Based Virtual Try-On

Hasib Zunair, Yan Gobeil, Samuel Mercier, A. Ben Hamza



Handles complex poses, retains logo, texture, embroidery, structure of clothing

Method	SSIM ( $\uparrow$ )				FID ( $\downarrow$ )
	VITON	VITON-E	VITON-M	VITON-H	
CA-GAN [15]	0.740	-	-	-	47.34
VITON [10]	0.783	0.787	0.779	0.779	55.71
CP-VTON [28]	0.745	0.753	0.742	0.729	24.43
VTNFP [31]	0.803	0.810	0.801	0.788	-
ClothFlow [9]	0.843	-	-	-	23.68
CP-VTON+ [20]	0.750	-	-	-	21.08
SieveNet [14]	0.837	-	-	-	26.67
ACGPN [30]	<b>0.845</b>	<b>0.854</b>	<b>0.841</b>	<b>0.828</b>	16.64
DCTON [8]	0.830	-	-	-	<b>14.82</b>
CIT [23]	0.827	-	-	-	-
<b>FIFA (Ours)</b>	<b>0.886</b>	<b>0.890</b>	<b>0.880</b>	<b>0.865</b>	<b>13.46</b>



MCM captures structure of clothing



MSC captures global context of clothing and shape of person

Method	SSIM ( $\uparrow$ )	FID ( $\downarrow$ )
ACGPN [30]	0.927	152.43
<b>FIFA</b>	<b>0.952</b>	<b>147.62</b>

FIFA is robust to in-the-wild images for virtual try-on