



UNIMORE
UNIVERSITÀ DEGLI STUDI DI
MODENA E REGGIO EMILIA

**YOOX
NET-A-PORTER
GROUP**



VITON-GT: An Image-based Virtual Try-On Model with Geometric Transformations

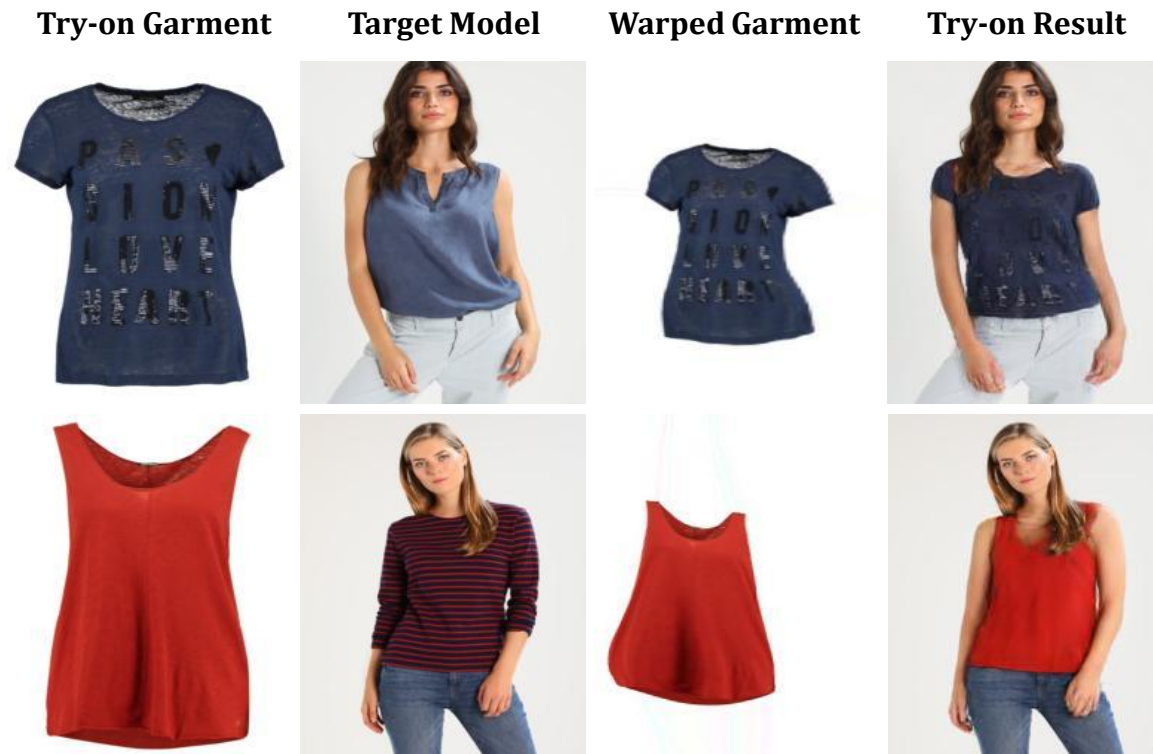


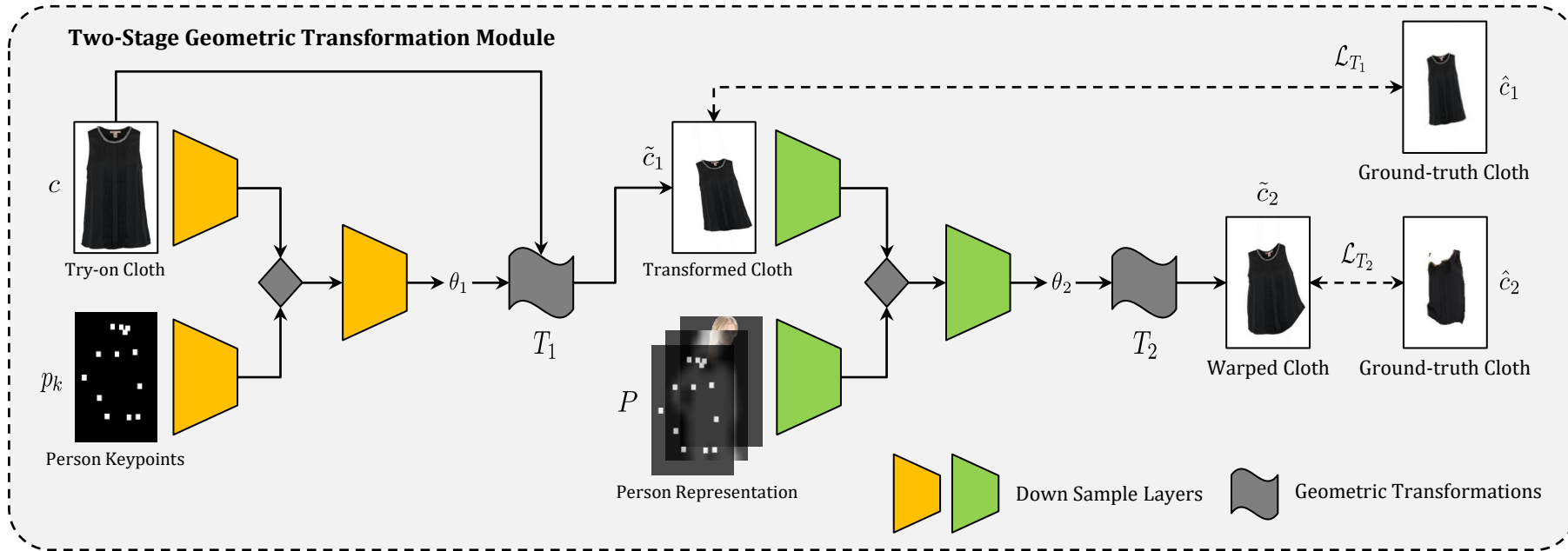
Matteo Fincato¹, Federico Landi¹, Marcella Cornia¹, Fabio Cesari², Rita Cucchiara¹

¹University of Modena and Reggio Emilia, ²YOOX NET-A-PORTER GROUP

Email: ¹{name.surname}@unimore.it, ²{name.surname}@ynap.com

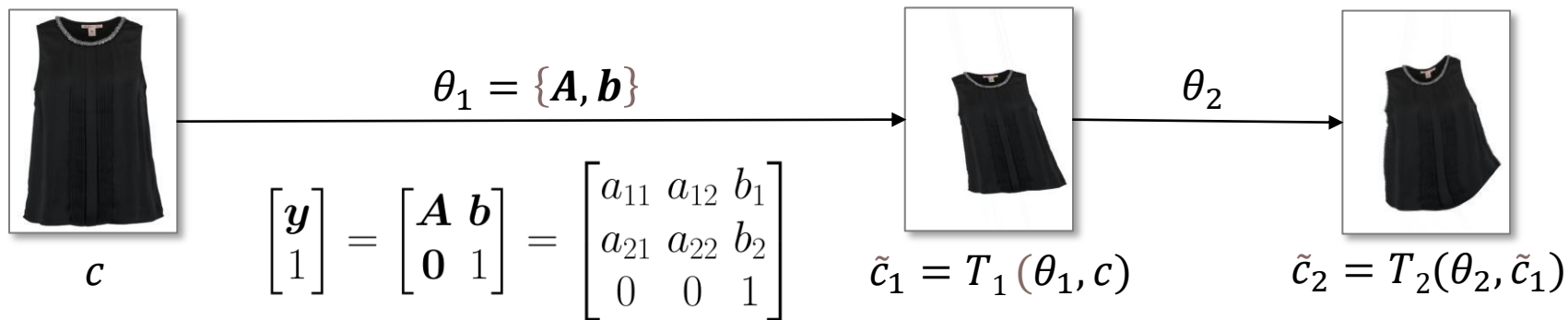
- We propose a novel architecture for image-based 2D single-pose virtual try-on called **VITON-GT** (Virtual Try-ON with Geometric Transformations).
- The proposed model includes two different components: **a two-stage geometric transformation module** and **a transformation-guided try-on module**.

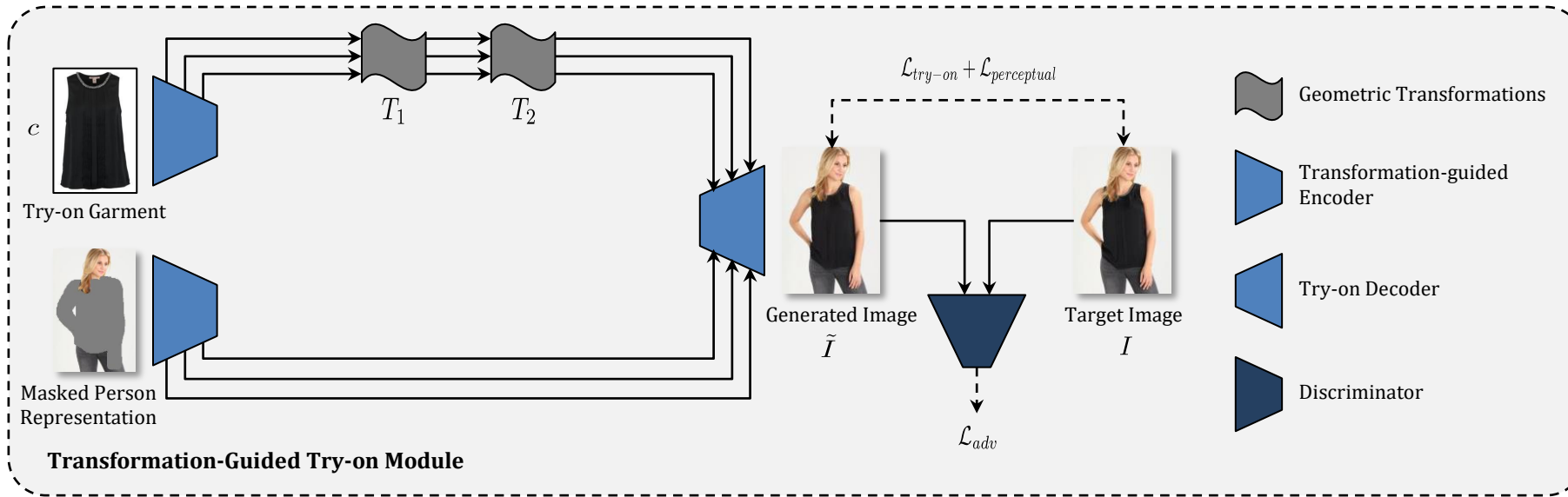




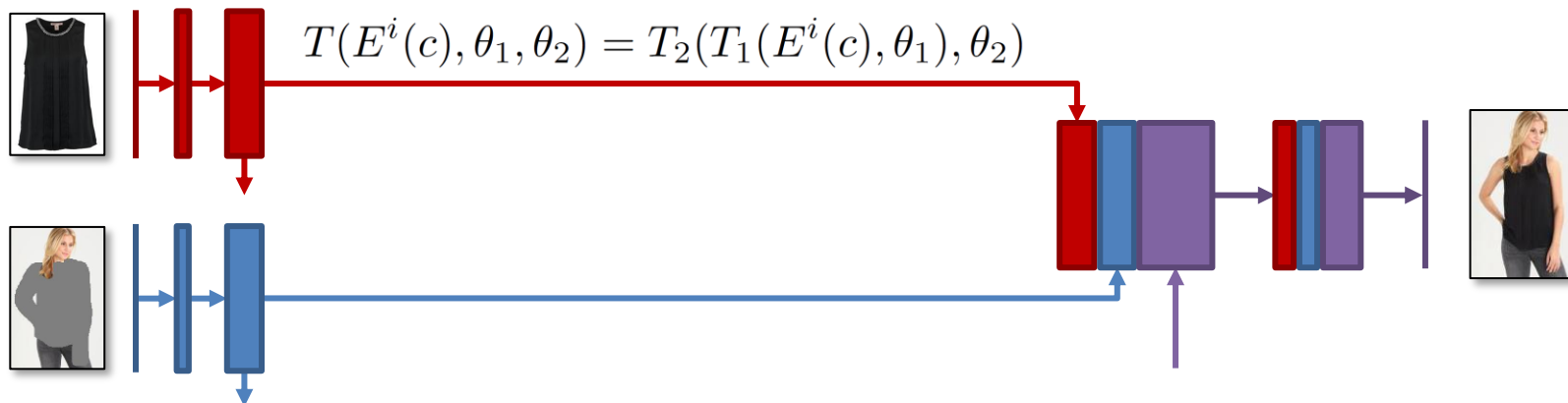
- We employ two different geometric transformations, namely **affine** and **thin-plate spline**, to warp the in-shop image c of a particular garment.
- We compute the parameters θ_1 for the affine transformation T_1
- We perform regression to predict the parameters θ_2 for the TPS transformation T_2
- Final loss:

$$L_{GT} = \lambda_1 L_{T_1} + \lambda_2 L_{T_2}$$





- We generate an output image \tilde{I} representing the reference person wearing c by employing a U-Net architecture
- We apply the previous learned spatial transformations in the clothes branch



- To guide the generation of the final image we implement three different losses: a L_1 loss, a perceptual loss and an adversarial loss
- $L_{TON} = \rho_1 L_{ton} + \rho_2 L_{vgg} + \rho_3 L_{adv}$

Try-on Garment



Target Model



CP-VTON [1]



VITON-GT



Model	FID	KID	IS
CP-VTON [1] (TPS only)	101.12	6.80±0.67	3.31±0.35
VITON-GT (Affine + TPS)	59.53	3.27±0.48	3.40±0.22

- Our model outperforms CP-VTON [1] on all the evaluation metrics.
- The affine transformation helps the TPS generating better warped clothes.
- Reduced artifacts and distortions.

[1] B. Wang, H. Zheng, X. Liang, Y. Chen, L. Lin, and M. Yang. "Toward characteristic-preserving image-based virtual try-on network". ECCV, 2018



Model	SSIM	MS-SSIM	FID	KID	IS
CP-VTON [1]	0.789	0.838	19.04	0.93±0.18	2.61±0.14
VITON-GT (no FT, no Adv. Loss)	0.879	0.919	15.32	0.58±0.19	2.72±0.14
VITON-GT (no Adv. Loss)	0.879	0.921	13.01	0.36±0.12	2.73±0.09
VITON-GT	0.886	0.925	12.45	0.32±0.12	2.76±0.11

- Reduced distortions while maintaining textures.
- Increased realism of the final images.
- Preserving details and characteristics of the original clothes.

[1] B. Wang, H. Zheng, X. Liang, Y. Chen, L. Lin, and M. Yang. "Toward characteristic-preserving image-based virtual try-on network". ECCV, 2018



Model	Short-Sleeve T-Shirts			Long-Sleeve T-Shirts			Sleeveless T-Shirts			Shirts			Sweatshirts		
	FID	KID	IS	FID	KID	IS	FID	KID	IS	FID	KID	IS	FID	KID	IS
CP-VTON [1]	23.81	0.86±0.16	2.41±0.21	31.92	1.85±0.33	2.66±0.18	31.50	1.98±0.34	2.36±0.20	35.38	2.33±0.38	2.43±0.14	31.89	1.57±0.28	2.63±0.15
VITON-GT (no FT, no Adv. Loss)	22.11	0.76±0.16	2.54±0.12	23.74	0.89±0.22	2.69±0.09	27.52	1.42±0.24	2.47±0.18	28.85	1.49±0.27	2.65±0.18	27.00	1.11±0.21	2.63±0.11
VITON-GT (no Adv. Loss)	20.95	0.61±0.16	2.63±0.17	20.02	0.62±0.16	2.79±0.16	24.30	1.16±0.30	2.47±0.10	25.67	1.18±0.27	2.60±0.17	24.30	0.90±0.17	2.70±0.14
VITON-GT	20.73	0.57±0.15	2.65±0.14	20.83	0.64±0.17	2.81±0.18	22.88	1.01±0.24	2.56±0.16	25.22	1.17±0.27	2.62±0.10	25.59	1.04±0.19	2.76±0.10



- We have presented VITON-GT, a new image-based virtual try-on model that integrates multiple geometric transformations of the input clothes during the generation of the try-on result.
- Through extensive experiments on two different datasets, we have demonstrated the effectiveness of our solution w.r.t. previously proposed methods.



UNIMORE
UNIVERSITÀ DEGLI STUDI DI
MODENA E REGGIO EMILIA

**YOOX
NET-A-PORTER
GROUP**



Thank you!



Matteo Fincato¹, Federico Landi¹, Marcella Cornia¹, Fabio Cesari², Rita Cucchiara¹

¹University of Modena and Reggio Emilia, ²YOOX NET-A-PORTER GROUP

Email: ¹{name.surname}@unimore.it, ²{name.surname}@ynap.com