PIVTONS: Pose Invariant Virtual Try-on Shoe with Conditional Image Completion

Chao-Te Chou, Cheng-Han Lee, Kaipeng Zhang, Hu-Cheng Lee, Winston H. Hsu

National Taiwan University, Taipei, Taiwan

ACCV 2018 2 – 6 December 2018 Perth Western Australia

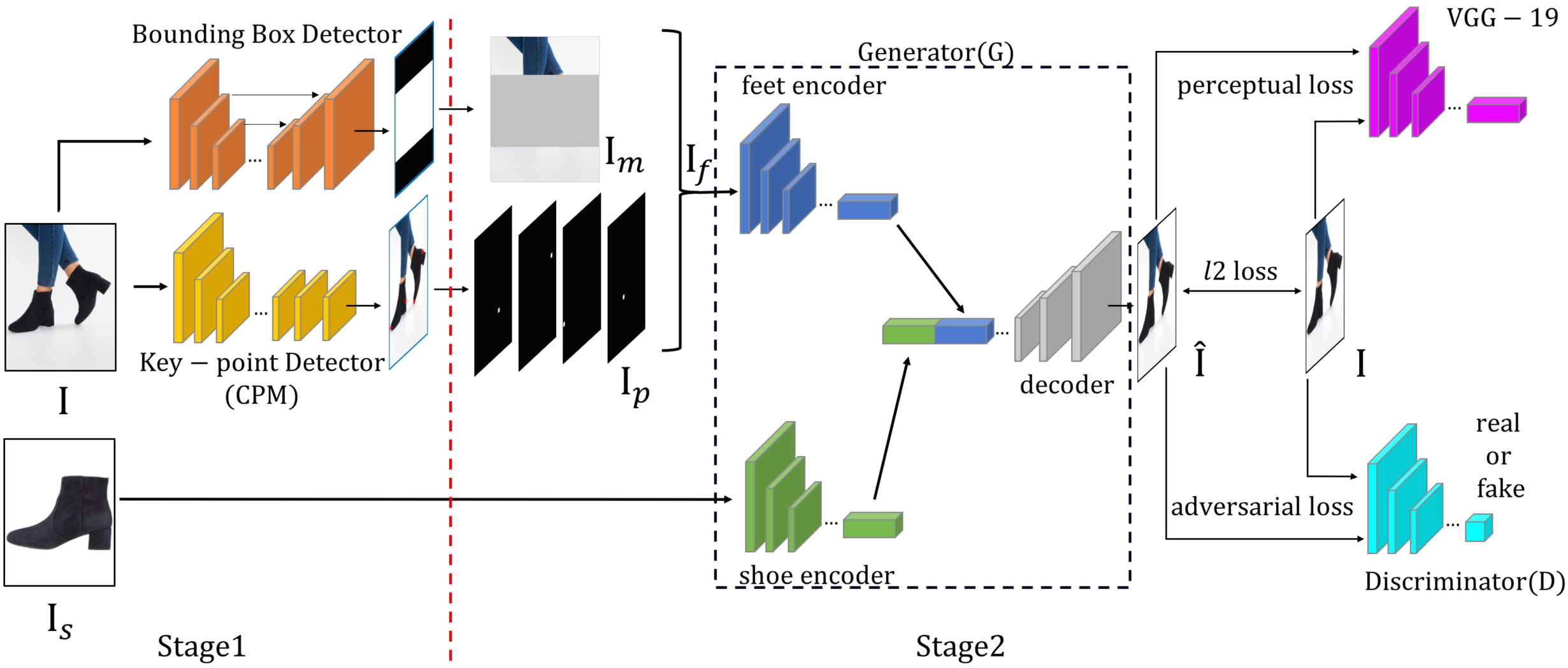
Contributions

- To the best of our knowledge, we are the first to cope with virtual try-on shoe with deep neural network.
- We present a conditional image completion method to deal with this problem.
- Through extensive experiments conducted in the collect images, we show promising results of our method.

Problem Definition

Source Image (User Input) | Target Item (From Website) | Target Image (Output)

Method



$L = L_{l2} + \alpha \cdot L_{perc} + \beta \cdot L_{adv}$, we choose $\alpha = 1$ and $\beta = 1$.

Experimental Results

The effectiveness of the masked source image I_m

Source Image | Target Item | Input I | Input I_m

The importance of key-points

Source Image | Target Item | w/o key-point | w/ key-point

The combination of losses

Source Image | Target Item | $\alpha = 0$ $\beta = 0$ | $\alpha = 1$ $\beta = 0$ | $\alpha = 1$ $\beta = 1$

Comparison with pix2pix-m

Source Image | Target Item | pix2pix-m | PIVTONS | Source Image | Target Item | pix2pix-m | PIVTONS

User Study

	pix2pix-m	w/o key-points	$\alpha = 0$ $\beta = 0$	$\alpha = 1$ $\beta = 0$
PIVTONS	0.9675	0.9775	0.9775	0.52

Reference:

[1] Isola et al., Image-to-Image Translation With Conditional Adversarial Network, CVPR17

[2] Han et al., An Image-Based Virtual Try-on Network, CVPR18

[3] Zalando: https://www.zalando.co.uk/