

*Monocular-to-3D Virtual Try-On using Deep  
Residual U-Net*

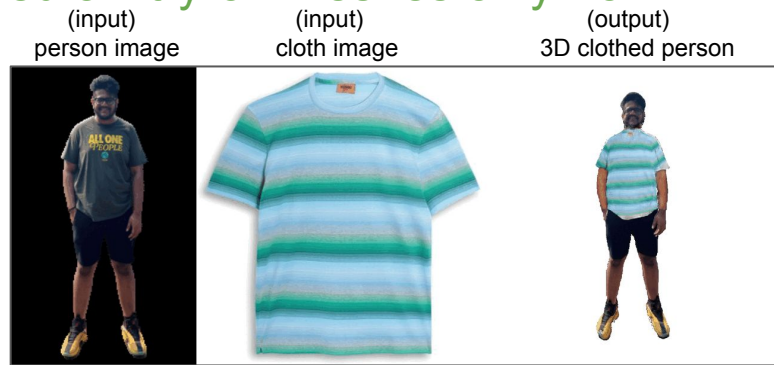
COMP 6381: Digital Geometric Modeling, Fall 2021  
Project Presentation

by

*Hasib Zunair, 40126681*

# Motivation

- Methods for 3D virtual try-on rely on **3D shapes** and **clothing templates** which is impractical for different clothing types (e.g. long/short sleeves)
- 2D try-on is another alternative, but cannot fully represent the human body
- A new method (M3D-VTON) generates textured 3D try-on meshes only from *person* and *clothing image*!



[1] Zhao. et al, <https://arxiv.org/abs/2108.05126>, ICCV 2021

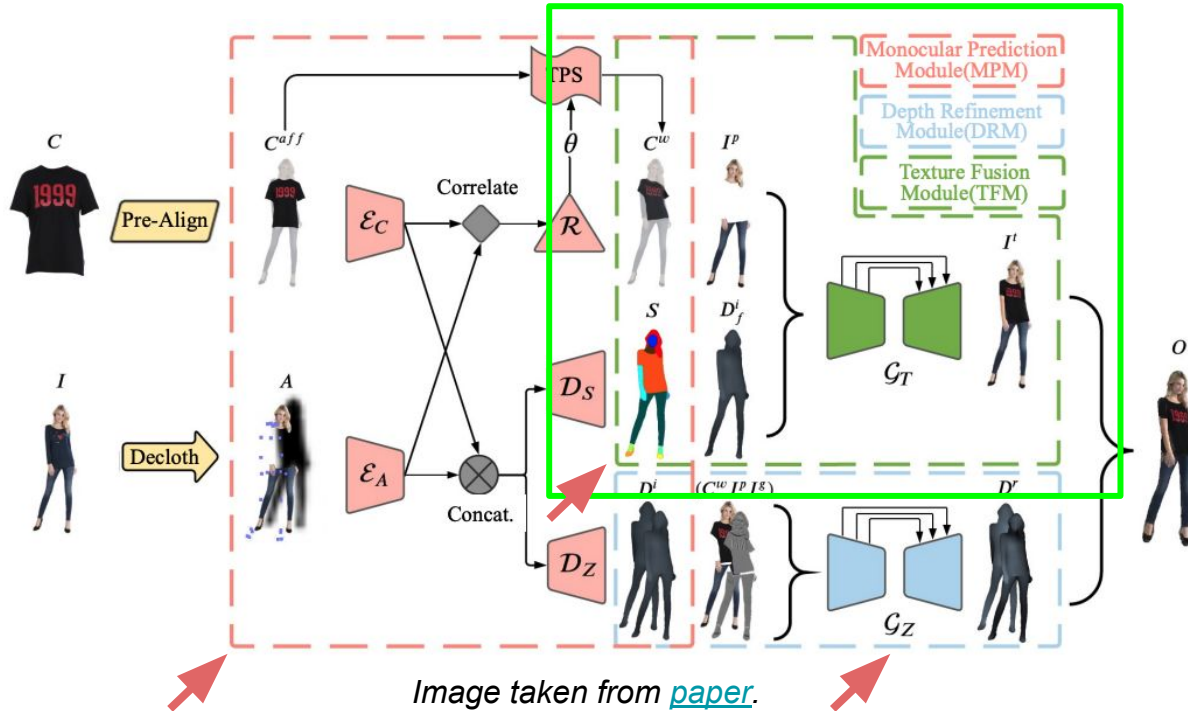
# Problems with M3D-VTON

- Simple U-Net [1] architecture is insufficient in *synthesizing body parts* and *capturing the complex relationship* of clothing (e.g front/back part)
- Leads to blurry and unrealistic outputs which negatively affects final 3D try-on result
- Residual connections [2] in the existing synthesis model to learn better representation of input data

[1] Ronneberger. et al, <https://arxiv.org/abs/1505.04597> , MICCAI 2015

[2] He. et al, <https://arxiv.org/abs/1512.03385> , CVPR 2016

# Approach

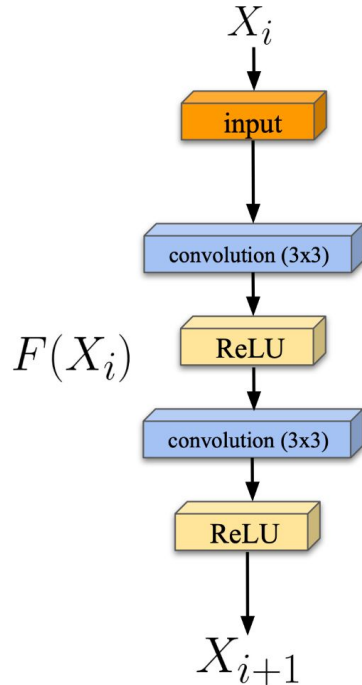


[1] Zhao. et al, <https://arxiv.org/abs/2108.05126>, ICCV 2021

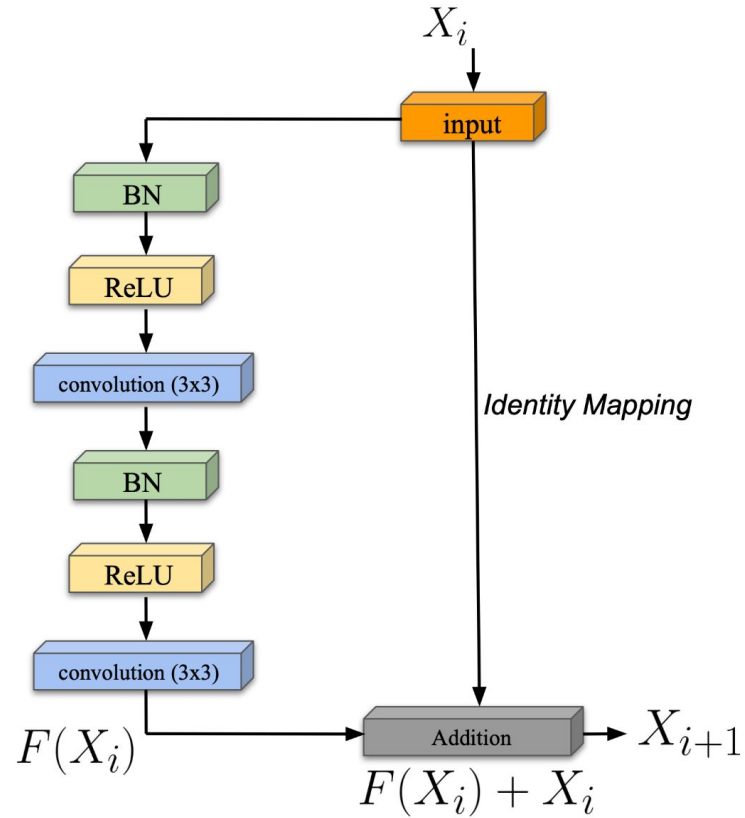
[2] Kazhdan. et al, <https://www.cs.jhu.edu/~misha/MyPapers/ToG13.pdf>, TOG 2013

[3] Telea. et al, <http://www.olivier-augereau.com/docs/2004JGraphToolsTelea.pdf>, J. Graphics 2004

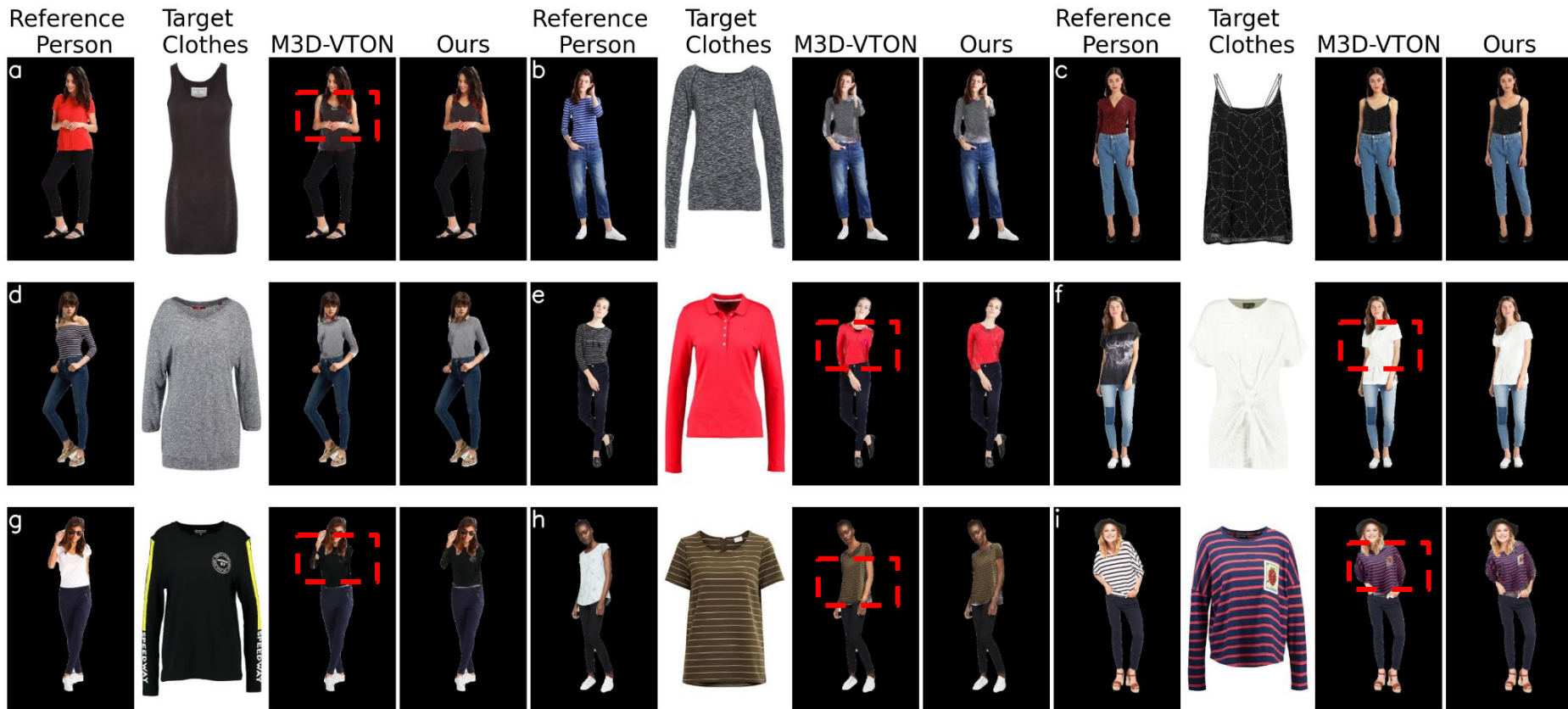
# Approach



*Plain connections in synthesis model*

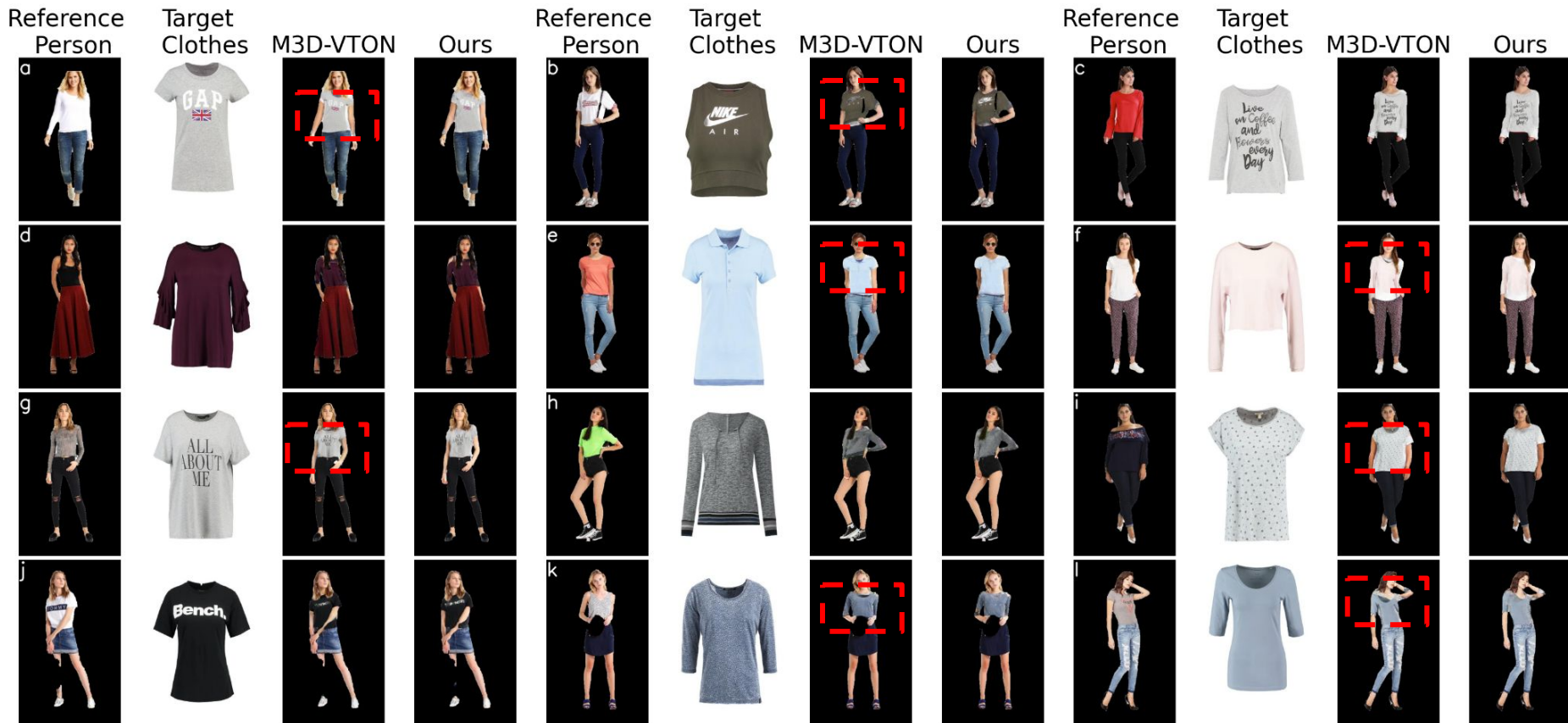


*Residual connections in synthesis model*



Visual comparison of 2D try-on outputs.

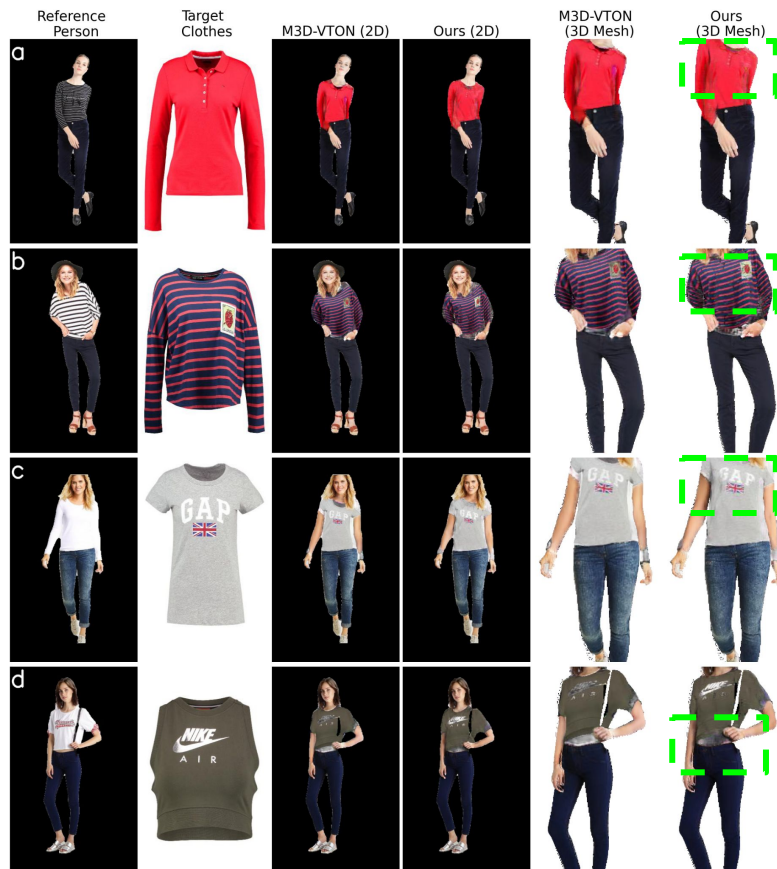
- Differentiate front and back part of clothing
- Preserves logo and texture of clothing
- Reduces artifacts in non-target body parts



Visual comparison of 2D try-on outputs.

- Differentiate front and back part of clothing
- Preserves logo and texture of clothing
- Reduces artifacts in non-target body parts

Better 2D try-on leads to better textured 3D mesh!



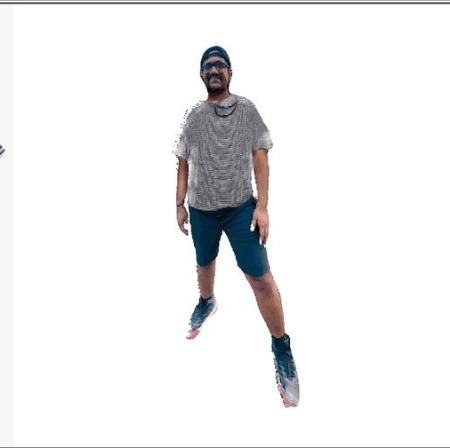
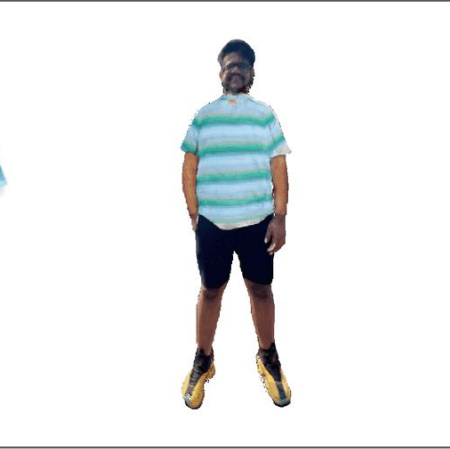
*Visual comparison of 3D try-on outputs.*



Person Image

Clothing Image

Textured 3D Mesh



*3D try-on outputs from our method using out-of-distribution images.*

Method	FID Score ↓	SSIM Score ↑
VITON, CVPR 2018	28.43	0.8807
CP-VTON, ECCV 2018	20.05	0.8503
CP-VTON+ CVPRW 2020	23.18	0.8782
ACGPN, CVPR 2020	20.19	0.8924
M3D-VTON, ICCV 2021	19.87	0.9725
Ours	<b>15.16</b>	<b>0.9814</b>

*FID and SSIM scores on the MPV3D test set.*

- Improvement on both metrics over M3D-VTON
- Significant improvement over other methods

## To conclude..

- Residual connections work better than plain connections
- Improves synthesis of 2D virtual try-on
- Eventually leads to better reconstruction of 3D virtual try-on meshes



[Project Website]