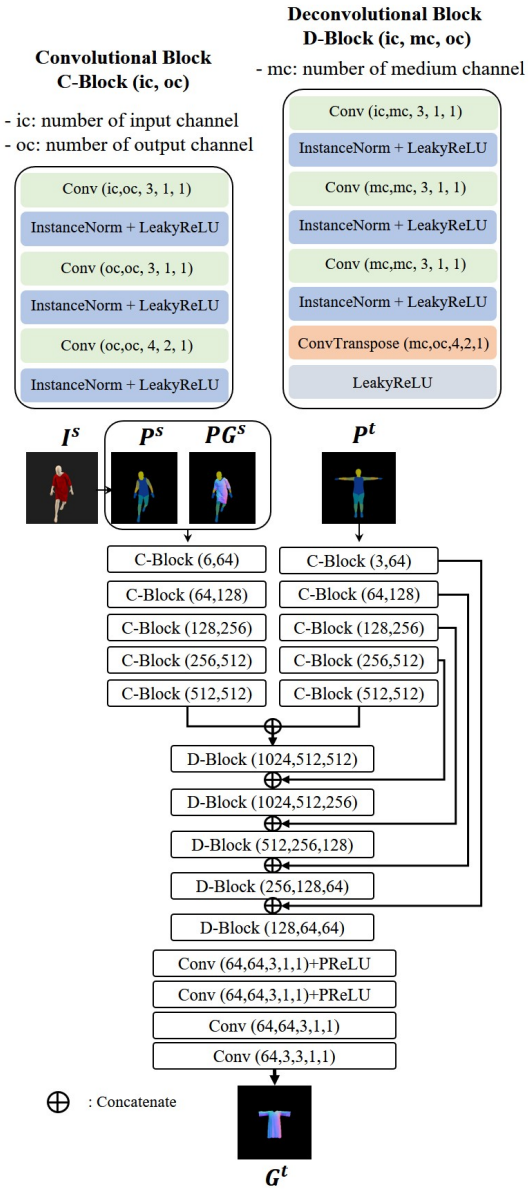

Figure 1: *The detailed structure of T-pose garment predictor. The architecture primarily consists of Convolutional Blocks (C-Block) and Deconvolutional Blocks (D-Block), with a LeakyReLU coefficient uniformly set at 0.2.*

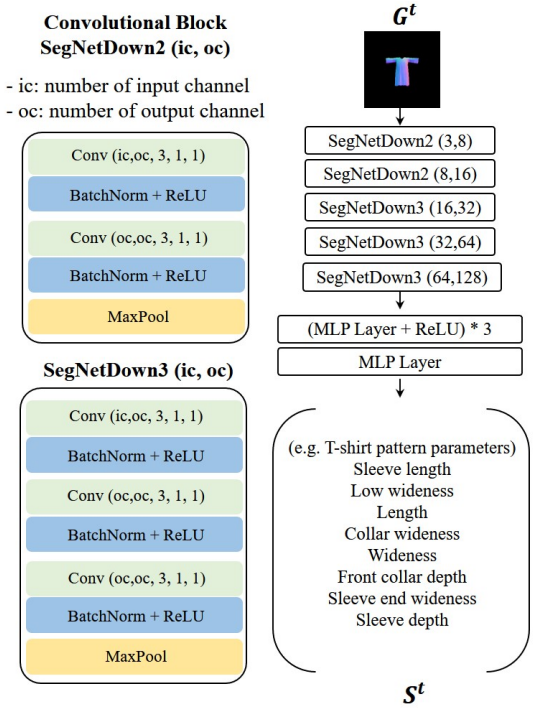

Figure 2: *The detailed structure of the sewing pattern parameter predictor.*

Appendix A: Network architecture

Figure 1 depicts the detailed architecture of the T-pose garment predictor. This architecture is inspired by the "SilNet" model from [YLG21]. Instead of predicting the combined silhouette of body and clothing at the target pose as they did, we focus on predicting the T-pose garment image G^t. The information obtained from concatenating the P^s and PG^s of Image I^s is passed through the C-Block, and the same process applies to the P^t image. As the information from these two C-Blocks integrates and passes through the D-Block, information about the T-pose is conveyed through a skip connection approach, thereby assisting the learning of clothing in T-pose.

Figure 2 presents the detailed architecture of the sewing pattern parameter predictor. To predict the parameters S^t, which consist of fewer than 10 parameters from high-dimensional images, a lighter feature that can effectively encapsulate image information is required. We initially train a SegNet [BKC17], an encoder-decoder architecture composed of convolutional layers. Subsequently, employing the learned filters, we decrease the dimension of G^t via SegNetDown blocks (2,3). Finally, we can predict the sewing pattern parameters S^t through a Multilayer Perceptron (MLP) process.

References

[YLG21] YOON J. S., LIU L., GOLYANIK V., SARKAR K., PARK H. S., THEOBALT C.: Pose-guided human animation from a single image in the wild. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2021), pp. 15039–15048. 3, 8

[BKC17] BADRINARAYANAN V., KENDALL A., CIPOLLA R.: Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence* 39, 12 (2017), 2481–2495. 4, 8

Supplementary Material
(3D garment reconstruction results from wild images)

Supplementary Material
(3D garment reconstruction results from wild images)