

# Supplementary Material for the paper: Enhancing Monocular 3D Hand Reconstruction with Learned Texture Priors

Giorgos Karvounas<sup>1,3</sup>  
gkarv@ics.forth.gr

Nikolaos Kyriazis<sup>1</sup>  
kyriazis@ics.forth.gr

Iason Oikonomidis<sup>1</sup>  
oikonom@ics.forth.gr

Georgios Pavlakos<sup>2</sup>  
pavlakos@cs.utexas.edu

Antonis A. Argyros<sup>1,3</sup>  
argyros@ics.forth.gr

<sup>1</sup>ICS-FORTH, <sup>2</sup>University of Texas at Austin, <sup>3</sup>University of Crete

The qualitative results on in-the-wild images illustrate the generalization ability of our method. The examples cover diverse scenarios, including hands with different glove textures, colors, and materials, as well as varying poses, viewpoints, and complex backgrounds. While baseline predictions often struggle to adapt to such challenging conditions, our approach with learned texture priors produces reconstructions that better capture the underlying hand shape and pose, even in visually cluttered or unconstrained settings. While our method demonstrates strong performance in a variety of scenarios, it is not without limitations. There remain challenging cases where the predictions fail to capture key aspects of the hand model. In the last figure, we present examples that illustrate these limitations. Specifically, these cases highlight situations where our approach was unable to correctly predict either the geometry of the hand or the texture.





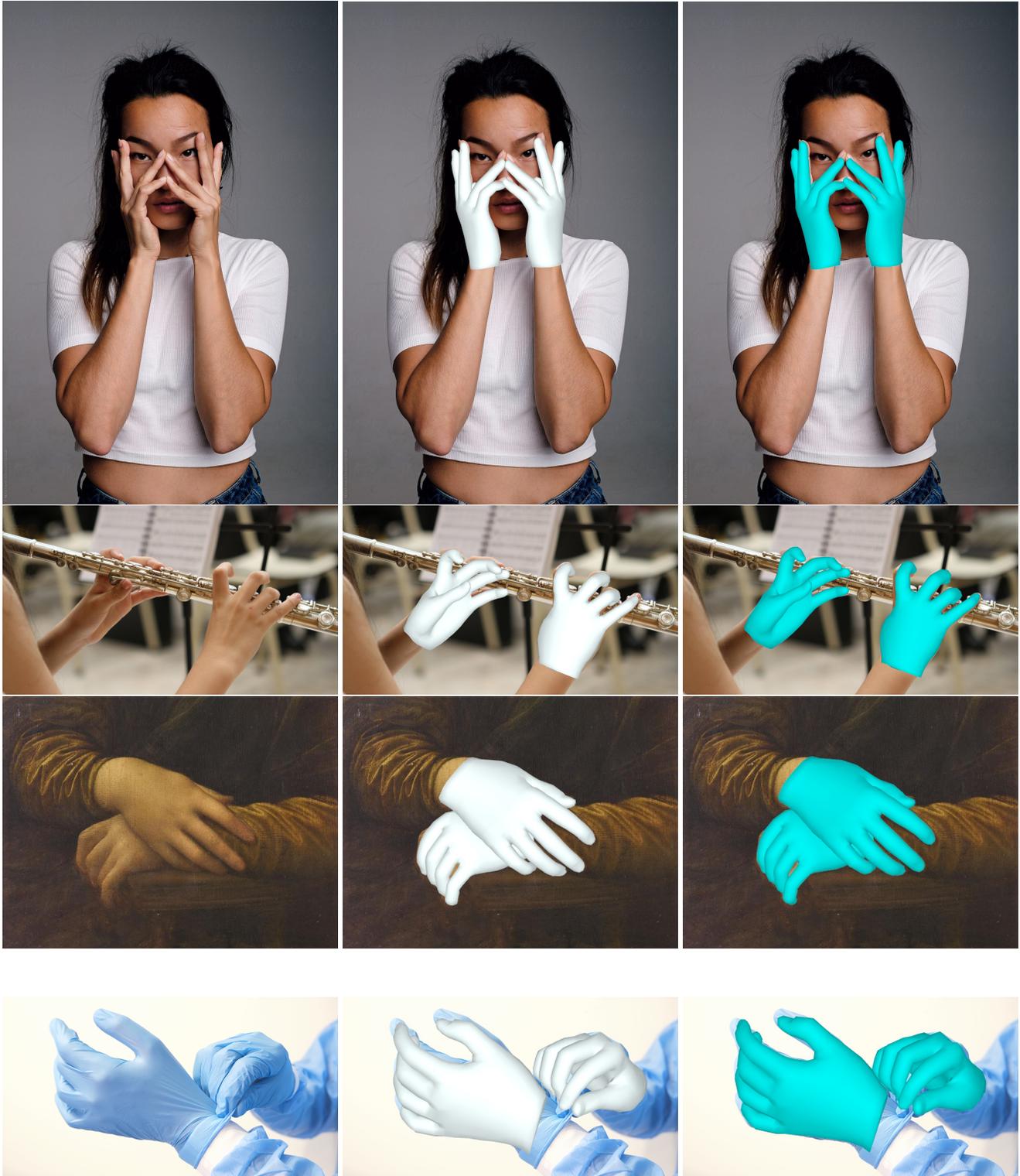


Figure 1. We provide indicative qualitative results on images in the wild to highlight the effectiveness of our approach. The first column shows the input RGB image, while the second column presents the reconstructed hand prediction by HaMeR. The third column illustrates the output from our method.

In the following figure, we present examples that illustrate the limitations of our method. Specifically, these cases highlight cases where our approach was unable to correctly predict either the geometry of the hand or the texture.



Figure 2. Examples of failure cases produced by our method. From left to right: input image, predicted hand geometry, and predicted geometry with texture generated from our module. These examples highlight scenarios where our approach struggles to accurately reconstruct either the hand geometry (bottom right) or the surface appearance (top right).