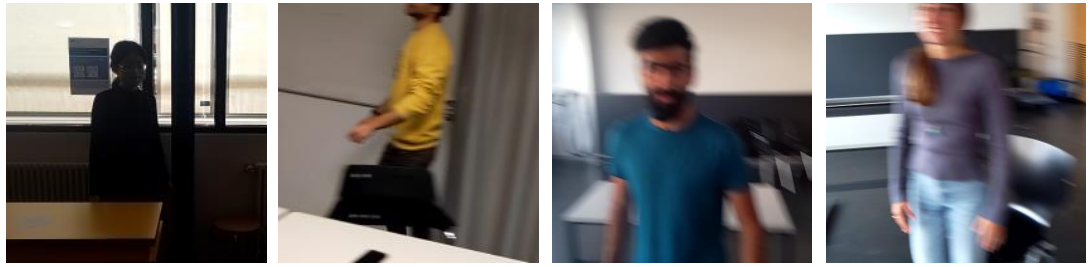
Exploring Effective Data Augmentations for 3D Human Pose and Shape Estimation from a Single Egocentric Image

Bowen Dang
dangbowen.bell@gmail.com

Motivation

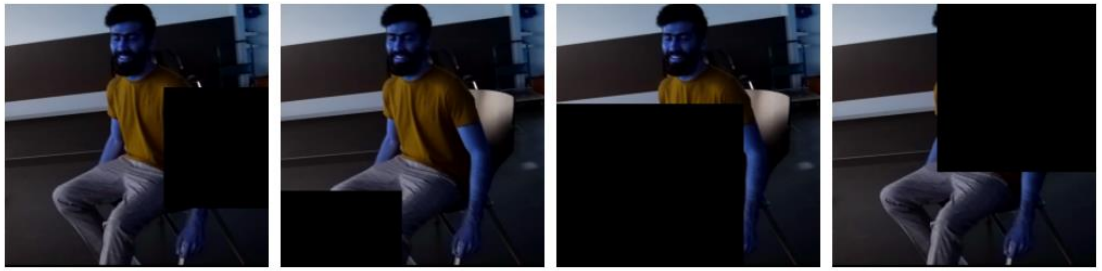
The occlusion and motion blur in the egocentric images make it hard to reconstruct accurate 3D human mesh.



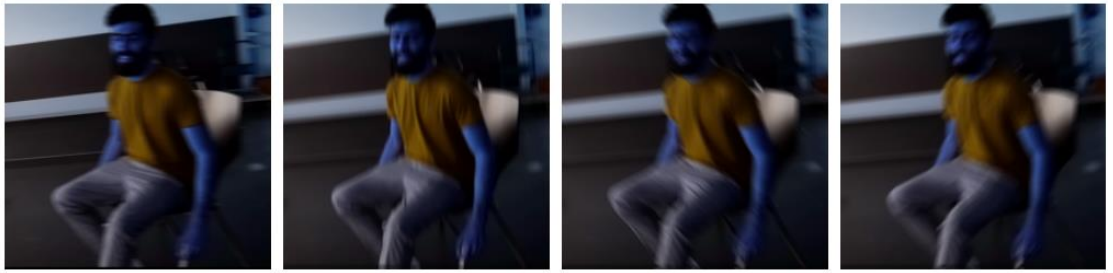
Method

We use SPIN[1] as our baseline and fine-tune the pre-trained model on EgoBody[2] dataset with CutOut[3] and Motion Blur.

CutOut: We use CutOut to simulate possible occlusion during shooting.



Motion Blur: We use Blur to simulate possible motion blur during shooting.



Results

The errors decrease after applying these data augmentations.

Model	MPJPE	PA-MPJPE	V2V	PA-V2V
SPIN	130.1220	81.9946	140.0008	90.0855
SPIN-ft	93.1091	61.5779	104.8394	70.5731
SPIN-ft(CO)	89.7161	58.6065	100.4049	66.2072
SPIN-ft(CO+MB)	87.8209	58.0566	98.3177	65.8668

Our best model ranks 3rd in the EgoBody challenge. Our code and model are publicly available at:
https://github.com/DangBowen-Bell/EgoBody-Challenge-Solution

Reference

[1] Kolotouros, et al, SPIN, ICCV 2019
[2] Zhang, et al, EgoBody, ECCV 2022
[3] DeVries, et al, CutOut, arXiv 2017