Supplementary Document for

Prior-guided Source-free Domain Adaptation for Human Pose Estimation

Dripta S. Raychaudhuri[1,2*] Calvin-Khang Ta[1] Arindam Dutta[1] Rohit Lal[1] Amit K. Roy-Chowdhury[1]
[1]University of California, Riverside [2]AWS AI Labs
{drayc001@, cta003@, adutt020@, rlal011@, amitrc@ece.}ucr.edu

A. Processing auxiliary dataset

In the experiment section of the main paper, we briefly described how to process the auxiliary dataset $\mathcal{D}_{\text{A}}$ to train the prior. Specifically, we generated a set of both plausible and implausible poses and calculated the corresponding distance values from the pose manifold $\tilde{\mathcal{D}} = \{(\boldsymbol{\theta}_i, d_i)\}_{i=1}^{M}$.

For poses derived from the auxiliary dataset, we label them as plausible poses and assign a distance value of $d = 0$. To generate an implausible pose $\boldsymbol{\theta_{nm}}$, we first randomly sample a pose $\boldsymbol{\theta_m} \sim \mathcal{D}_{\text{A}}$ and convert the sampled pose to polar coordinates,

$$u_1 = \arccos(\boldsymbol{\theta}^1_{\boldsymbol{m}})$$
$$u_2 = \arcsin(\boldsymbol{\theta}^2_{\boldsymbol{m}}).$$

Next, we sample noise from the Von-Mises distribution,

$$n_1, n_2 \sim f(n|\mu, \kappa) \quad \text{where } f(n|\mu, \kappa) = \frac{\exp\left(\kappa \cos(n - \mu)\right)}{2\pi I_0(\kappa)},$$

and add it to the coordinates to obtain the new pose $\boldsymbol{\theta_{nm}}$,

$$\boldsymbol{\theta}^1_{\boldsymbol{nm}} = \cos(u_1 + n_1)$$
$$\boldsymbol{\theta}^2_{\boldsymbol{nm}} = \sin(u_2 + n_2).$$

In our experiments we set $\mu = 0$ and sample κ randomly from the set $\{2, 4, 8\}$.

We employ the nearest neighbor strategy described in [1] to assign a distance value to each synthetically generated pose. To accomplish this, we first use FAISS [2] and L2 distances to approximate the k' nearest neighbors of the pose from the set of clean poses. From these neighbors, we identify the exact k nearest ones. In our approach, we set $k' = 500$ and $k = 5$. Finally, we determine the ground truth distance by calculating the average of the k smallest distances.

* Currently at AWS AI Labs. Work done while the author was at UCR.

B. Qualitative results

We present additional visual results on the SURREAL→BRIAR scenario in Figure A-1. Although our approach does not use any source data for adaptation, it is able to match the predictions produced by UDAPE [3], which uses source data.

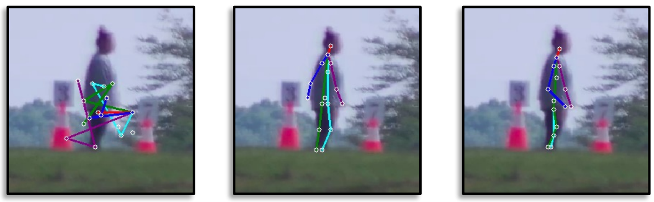
References

[1] Garvita Tiwari, Dimitrije Antic, Jan Eric Lenssen, Nikolaos Sarafianos, Tony Tung, and Gerard Pons-Moll. Pose-ndf: Modeling human pose manifolds with neural distance fields. In *ECCV*, 2022. A-1

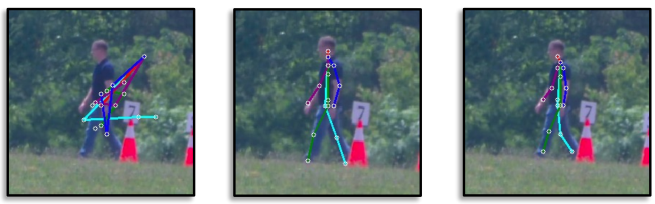
[2] Jeff Johnson, Matthijs Douze, and Hervé Jégou. Billion-scale similarity search with GPUs. *IEEE Transactions on Big Data*, 7(3):535–547, 2019. A-1

[3] Donghyun Kim, Kaihong Wang, Kate Saenko, Margrit Betke, and Stan Sclaroff. A unified framework for domain adaptive pose estimation. In *ECCV*, 2022. A-1, A-2

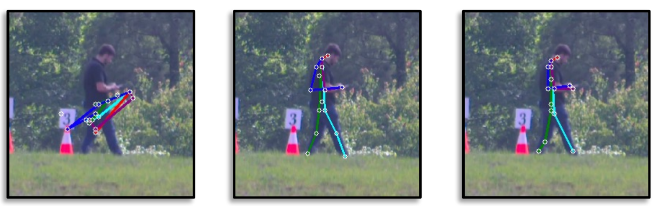
a) 100m



b) 200m

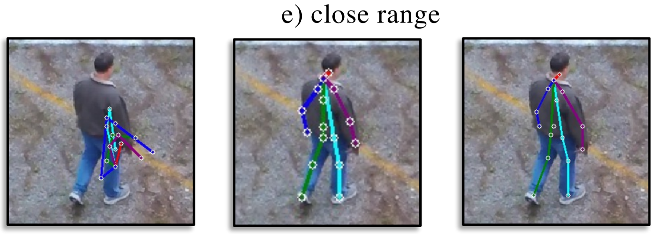


c) 400m



d) 500m

e) close range



f) uav

Figure A-1. **Qualitative results on SURREAL → BRIAR.** We demonstrate sample results on the BRIAR dataset at all ranges. For each range, we display three images: the leftmost shows the *Source only* prediction, the middle one shows the *UDAPE* [3] prediction, and the rightmost shows the prediction made by our framework.