PoseGuru: Landmarks for Explainable Pose Correction using Exemplar-Guided Algorithmic Recourse

Supplementary Material

A. More Implementation Details and Ablation Studies

Implementation Details. As stated in the main paper, we employ the Blazepose [19] pre-trained pose estimation model to obtain the pose estimates from both datasets. Additionally, for training pose classifiers, we utilize individual shallow, fully connected neural networks dedicated to each dataset. The architectures for YogaHPC and Pilates32+P are [24, 128, 64, 31] and [24, 256, 128, 64, 32] respectively. These models are trained using a learning rate of 0.01, 1000 steps, and the NAdam optimizer. The architecture for the ANN Regressor baseline is [24, 64, 32, 24]. This model is trained using a learning rate of 0.001 and the Adam optimizer.

Additional Ablation Studies. *Effect of Coefficients of Terms in Objective Function.* As stated in the main paper, in our overall objective function in Sec 3, $\lambda_c, \lambda_s, \lambda_l$ and λ_a were chosen to be 0.1, 0.2, 0.3 and 0.4 respectively. An analysis of these choices are included here in Table 3 for the YogaHPC dataset. As observed, the best choices for the cost coefficients are 0.1, 0.2, 0.3 and 0.4 respectively. Key observations from this study include: (i) the framework, expectedly, performs the worst when the landmark and angle cost terms are completely ignored; (ii) While a higher weight for the landmark cost generally improves performance, this is not necessarily true for the angles cost term. This corroborates our decision to choose landmarks as a key component of our framework (unlike CARE which relies on angles); (iii) Higher weights for both prediction and stick length cost terms do not necessarily improve performance, and these coefficients need to be judiciously chosen.

Effect of Choice of Threshold Values. Table 5 shows the effect of different thresholds on Pilates-32+P dataset. We conducted our experiments using thresholds incremented by 0.01, ranging from 0.0 to 0.12. The table reports thresholds with 0.02 increments. Given that there can be multiple correct poses for a single incorrect Yoga or Pilates pose—requiring both precision and flexibility—we selected 0.10 as the final threshold for reporting results in the main paper. Lower thresholds demand higher precision from users, which may be impractical. A 10% tolerance provides an acceptable balance between accuracy and ease of correction.

Effect of Procrustes Transformation of Exemplar. Table 4 shows effect of Procrustes Transformation of Exemplar for various objective functions. Note that the Procustes distance is defined as follows. Let $X, Y \in \mathbb{R}^{n \times d}$ be two configurations of points in d-dimensional space. The Procrustes distance is the residual sum of squared differences between the optimally transformed configurations:

$$d(X, Y) = \|X_{\text{scaled}} - Y_{\text{scaled}} R\|_F$$

Thus, the Procrustes distance is calculated by centering and scaling the two configurations, aligning them through the optimal rotation R (determined via SVD), and then computing the residual difference between them. Whenever an exemplar is involved in a cost term, such as in C_{land}, transforming the exemplar into the Procrustean space of the candidate counterfactual (x') has consistently demonstrated improved results across all methods. This transformation allows for better alignment between the exemplar and the candidate, leading to more accurate outcomes as evaluated by our metrics.

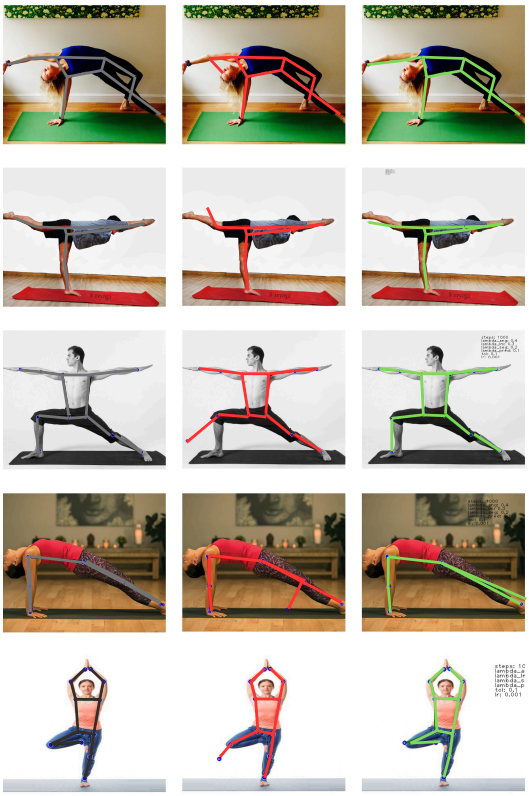

Figure 9. YogaHPC Additional Qualitative Results. Ground truth in black, incorrect pose in red and correct pose in green.

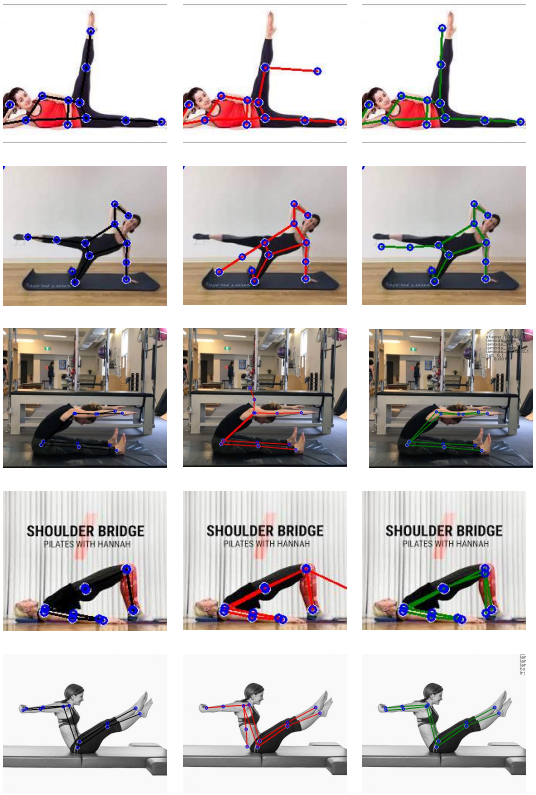

Figure 10. Pilates-32+P Additional Qualitative Results. Ground truth in black, incorrect pose in red and correct pose in green.

Coefficients				Metrics	
λ_c	λ_s	λ_l	λ_a	PCP (MPIJAD ↑)	PCIK ↑
0.3	0.3	0.2	0.2	0.66	**83**
0.3	0.3	0.3	0.1	0.69	**83**
0.4	0.4	0.2	0.0	0.58	80
0.2	0.4	0.4	0.0	0.62	80
0.4	0.4	0.0	0.2	0.59	80
0.5	0.5	0.0	0.0	0.22	75
0.2	0.4	0.0	0.4	0.63	80
0.4	0.3	0.2	0.1	0.65	**83**
0.3	0.2	0.2	0.3	0.67	82
0.4	0.4	0.2	0.0	0.58	80
0.5	0.5	0.0	0.0	0.22	75
0.4	0.4	0.2	0.0	0.58	80
0.1	0.2	0.3	0.4	**0.71**	82

Table 3. Performance across different choices of coefficients. Wining configuration is in blue.

B. Additional Qualitative Results

Figure 9 shows additional qualitative results for YogaHPC and Figure 10 shows additional qualitative results for the Pilates-32+P dataset. Note how the poses shown in green (rightmost image in each row) corrects the appropriate body segment from the poses shown in red (middle image in each row). These results are shown across a variety of poses across both datasets.

Metric	Method	Thresholds (T)				
		0.02	**0.04**	**0.06**	**0.08**	**0.10**
PCP (MAD ≤ T)	Pred+Seg+Land	0.43	0.83	0.95	0.99	1.00
	Proc. Pred+Seg+Land	**0.56**	**0.85**	**0.95**	**0.99**	0.99
	Pred+Seg+Angles+Land	0.46	0.81	0.93	0.98	1.00
	Proc. Pred+Seg+Angles+Land	**0.67**	**0.86**	**0.97**	**0.98**	0.99
PCP (MPIJAD ≤ T)	Pred+Seg+Land	0.03	0.13	0.28	0.43	0.52
	Proc. Pred+Seg+Land	**0.13**	**0.27**	**0.41**	**0.49**	**0.58**
	Pred+Seg+Angles+Land	0.04	0.16	0.34	0.46	0.53
	Proc. Pred+Seg+Angles+Land	**0.15**	**0.39**	**0.56**	**0.66**	**0.71**
PCK	Pred+Seg+Land	0.71	0.74	0.77	0.80	0.84
	Proc. Pred+Seg+Land	**0.73**	**0.78**	**0.82**	**0.86**	**0.88**
	Pred+Seg+Angles+Land	0.72	0.74	0.78	0.82	0.85
	Proc. Pred+Seg+Angles+Land	**0.75**	**0.79**	**0.84**	**0.87**	**0.89**
PCIK	Pred+Seg+Land	0.48	0.53	0.59	0.65	0.73
	Proc. Pred+Seg+Land	**0.52**	**0.62**	**0.69**	**0.75**	**0.80**
	Pred+Seg+Angles+Land	0.50	0.54	0.60	0.68	0.73
	Proc. Pred+Seg+Angles+Land	**0.54**	**0.63**	**0.72**	**0.78**	**0.82**

Table 4. Procrustes (Proc.) consistently performs better across different methods.

Metric	Method	Thresholds (T)				
		0.02	**0.04**	**0.06**	**0.08**	**0.10**
PCP (MAD ≤ T)	Centroid	0.29	0.54	0.80	0.92	0.98
	Medoid	0.33	0.61	0.82	0.92	0.97
	Exemplar	0.48	0.78	0.90	0.94	0.99
	Proc. Exemplar	0.61	0.83	0.87	0.92	0.93
	ANN Regressor	0.05	0.31	0.67	0.89	0.99
	CARE [27]	0.56	**0.95**	**1.00**	**1.00**	**1.00**
	PoseGuru	**0.66**	0.89	**1.00**	**1.00**	**1.00**
PCP (MPIJAD ≤ T)	Centroid	0.02	0.08	0.21	0.32	0.40
	Medoid	0.02	0.11	0.23	0.38	0.46
	Exemplar	0.07	0.18	0.38	0.57	0.65
	Proc. Exemplar	**0.11**	0.25	**0.47**	**0.66**	0.71
	ANN Regressor	0.00	0.02	0.05	0.13	0.21
	CARE [27]	0.00	0.03	0.13	0.27	0.32
	PoseGuru	0.10	**0.26**	0.45	**0.66**	**0.77**
PCK (within T)	Centroid	0.69	0.70	0.72	0.74	0.76
	Medoid	0.70	0.71	0.73	0.76	0.78
	Exemplar	0.71	0.73	0.76	0.78	0.82
	Proc. Exemplar	0.73	0.76	0.79	0.82	0.85
	ANN Regressor	0.64	0.65	0.66	0.67	0.68
	CARE [27]	**0.81**	**0.83**	**0.85**	**0.87**	0.89
	PoseGuru	0.74	0.78	0.82	0.85	**0.90**
PCIK (within T)	Centroid	0.01	0.05	0.12	0.18	0.26
	Medoid	0.01	0.06	0.15	0.23	0.31
	Exemplar	0.02	0.10	0.18	0.28	0.41
	Proc. Exemplar	0.52	0.57	0.63	0.68	0.74
	ANN Regressor	0.38	0.39	0.41	0.43	0.45
	CARE [27]	0.00	0.00	0.04	0.06	0.19
	PoseGuru	0.49	**0.58**	**0.67**	**0.73**	**0.81**

Table 5. **Threshold ablation study on Pilates dataset**: PoseGuru consistently performed well against all baselines as we vary thresholds. **Best results** are in bold and second best results are underlined.

C. Sample Dataset Images

Figure 13 shows samples of images from the YogaHPC dataset. Out of total 31 pose classes, five classes have significant occlusions (shown in red box), 15 classes have partial occlusions (shown in yellow box) and 11 classes are occlusion-free (shown in green box). Figure 14 shows images from the Pilates-32+P dataset. One RGB pose example for each of the Pilates-32+P pose classes is shown.

D. Sample Questions from User Study

As mentioned in the main paper (Sec. 5), we conducted a user study with 8 experts to rate pose corrections from both datasets. For completeness of understanding, Figures 11 and 12 show sample questions (out of a total of 8 questions for each dataset) from our user study. The assessment provided by Yoga and Pilates experts using the 5-point Likert scale enables us to gauge the extent to which they agree with our framework's pose correction approach.

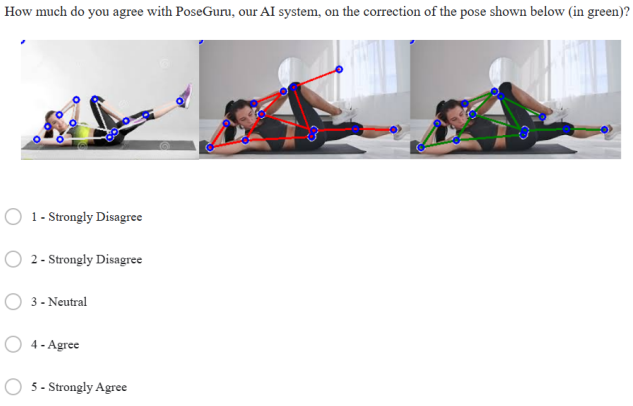

Figure 11. Sample question from User study on Pilates-32+P

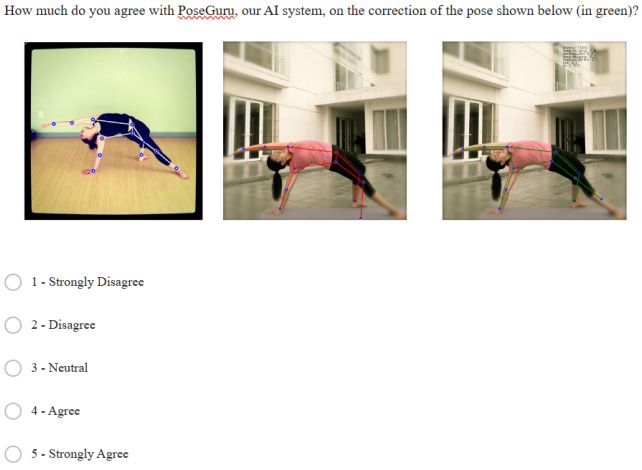

Figure 12. Sample question from user study on YogaHPC

Figure 13. Sample images from the YogaHPC dataset

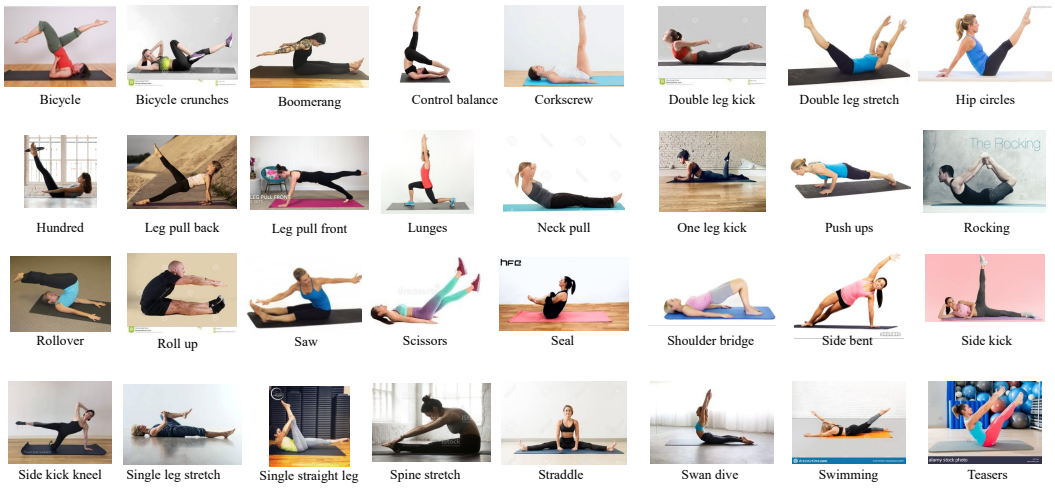

Figure 14. Sample images from the Pilates32 dataset