Deutsches Forschungszentrum für Künstliche Intelligenz
German Research Center for Artificial Intelligence

Towards Unconstrained 2D Pose Estimation of the Human Spine

Muhammad Saif Ullah Khan, Stephan Krauß, and Didier Stricker
German Research Center for Artificial Intelligence (DFKI)

CVPR Nashville JUNE 11-15, 2025

Introduction

Problem: Current human-pose datasets treat the spine as a single straight line, making it impossible to analyse posture, trunk flexion, or segmental motion in sports and healthcare. We propose:

- **SpineTrack**: first large–scale 2D dataset with **9 vertebral keypoints** + body + feet → **25 k** synthetic & **33 k** real images, 58 k persons
- **SpinePose**: lightweight teacher–student upgrade of RTMPose that learns the extra spine joints *without hurting* COCO/Halpe accuracy
- **Anatomical losses**: structure and smoothness terms keep the predicted spine physically plausible

Methodology

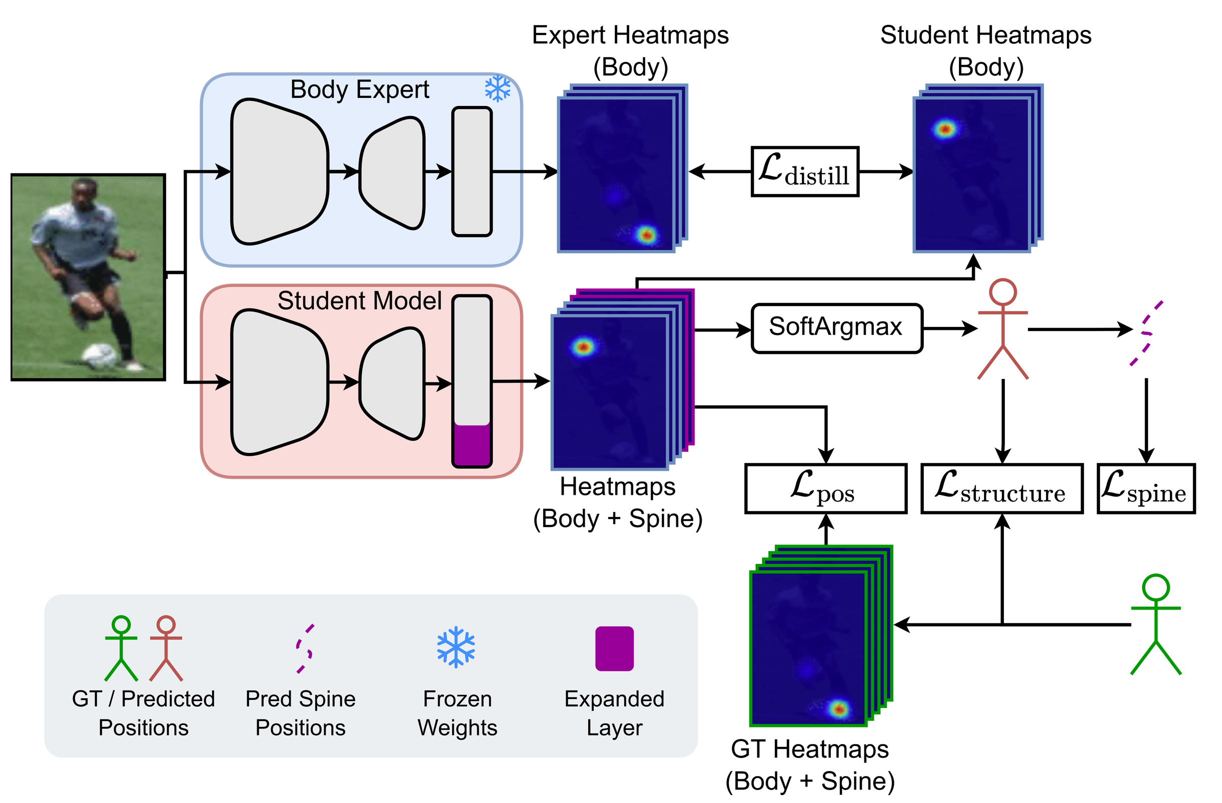

Figure 1: SpinePose: teacher–student distillation with anatomical priors.

$$(\mathcal{L}_{\text{total}} = \alpha L_{\text{pos}} + \beta L_{\text{distill}} + \gamma_1 L_{\text{struct}} + \gamma_2 L_{\text{smooth}})$$

- **Keypoint** L_{pos}: regress GT heat-maps
- **Distill** L_{distill}: maintain generalization
- **Structure** L_{struct}: bone-angle penalty
- **Smooth** L_{smooth}: curve regulariser along vertebrae

SpinePose-l variant	COCO AP	Halpe26 AP	Spine AP
Only Keypoint Loss	69.6	72.2	87.0
+ Distillation	70.8	73.2	87.3
+ Spine smoothness	74.7	76.6	**88.7**
+ Structure	71.2	73.4	87.7
Full (all losses)	**75.2**	**77.0**	88.4

- Distillation helps retain performance on standard real-world benchmarks
- Smoothness term brings largest spine gain
- Full recipe reaches the best cross-dataset balance

Synthetic Data Generation

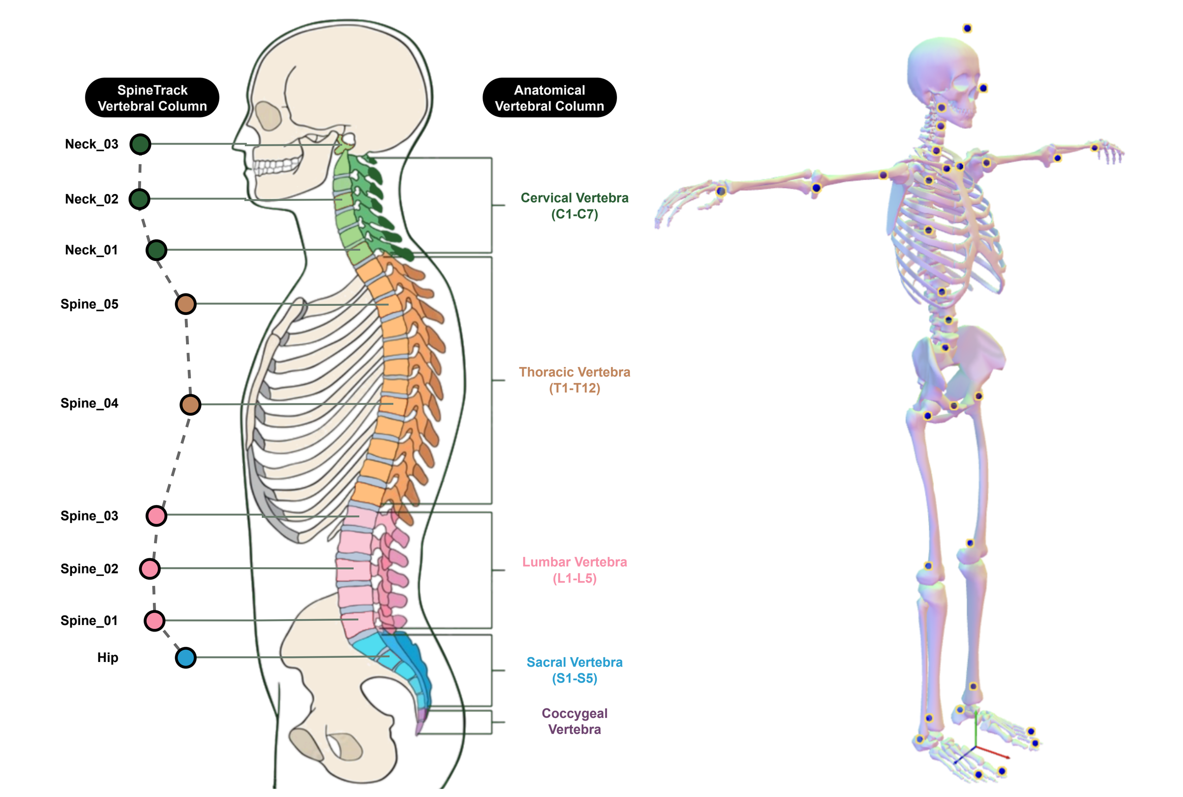

Figure 2: Skeleton definition & OpenSim alignment.

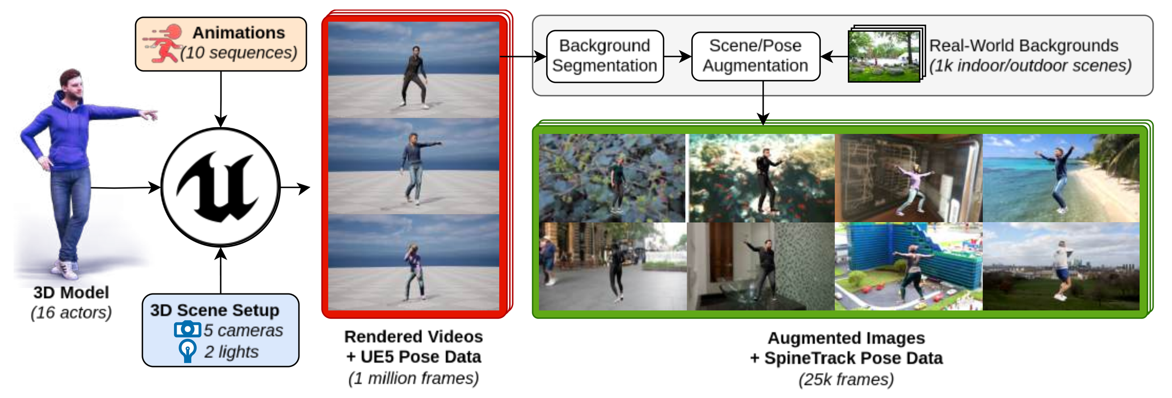

Figure 3: UE5 pipeline → 25 k labeled frames.

Contributions Summary

- **SpineTrack** dataset: 25 k synthetic + 33 k real images with biomechanically validated 9-vertebra annotations.
- Lightweight distillation upgrade adds spine to any RTMPose-style backbone (37 kpts) with *zero* extra backbone parameters.
- New **structure** and **smoothness** losses tailored to curved articulated segments.
- **+89.6 AP** on spine while retaining > 98% COCO/Halpe accuracy; effect consistent across S/M/L backbones.

Real Data Annotation

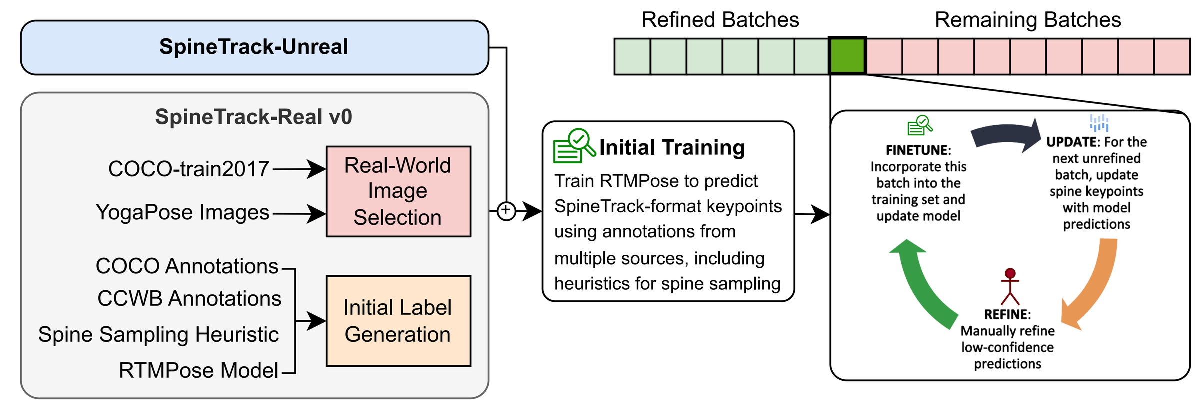

Figure 4: Active-learning loop for SpineTrack-Real with human-in-the-loop pipeline.

Figure 5: Example annotated in-the-wild images (33 k).

Qualitative Results

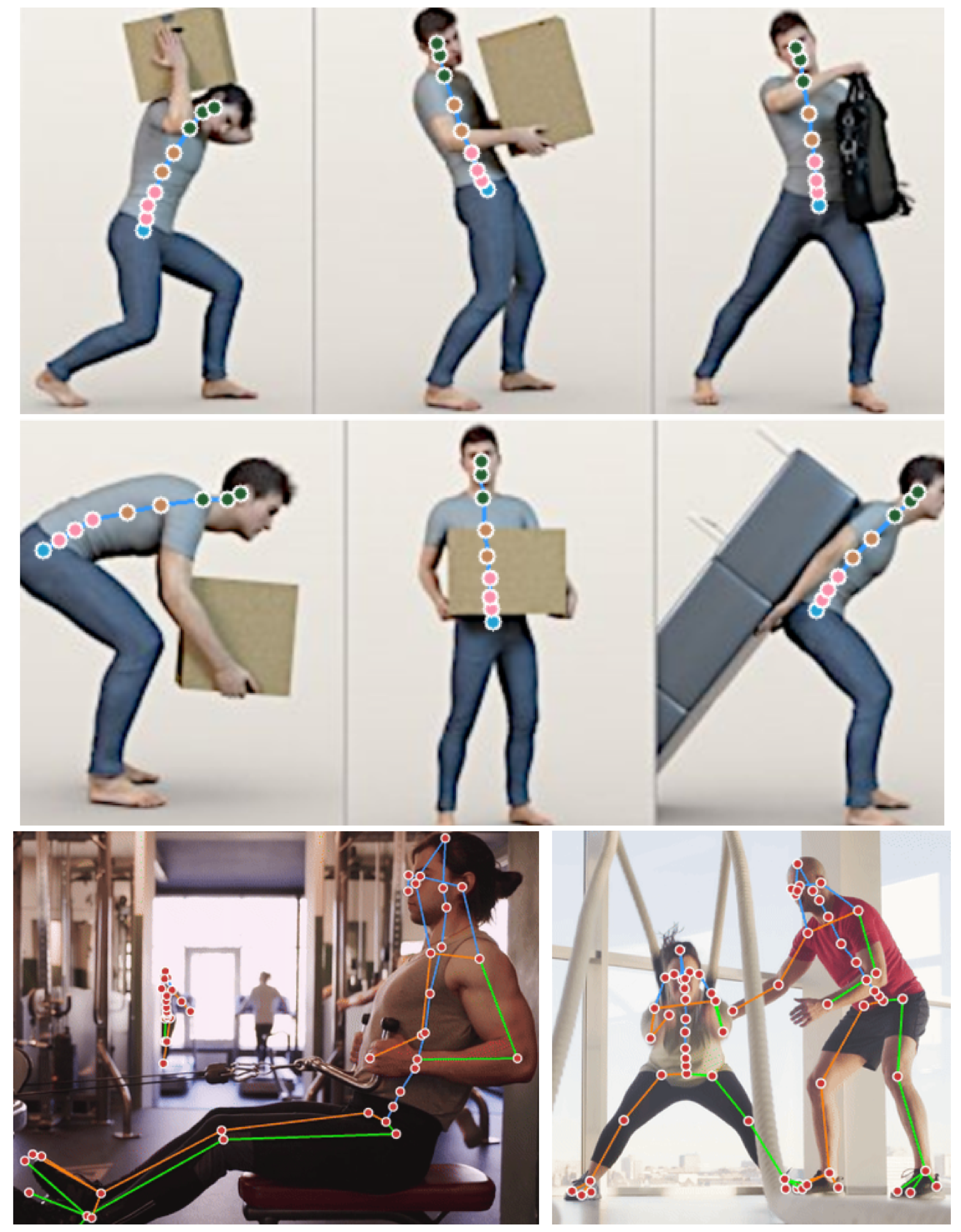

Figure 6: SpinePose accurately predicts extended spine joints in different everyday and sporting activities.

Detailed Results

Method	Train Data	Kpts	Body AP	Body AR	Feet AP	Feet AR	Spine AP	Spine AR	Overall AP	Overall AR	Params (M)	FLOPs (G)
RTMPose-t	Body8	26	76.9	80.0	74.1	79.7	0.0	0.0	15.8	17.9	3.51	0.37
RTMPose-s	Body8	26	**80.9**	**83.6**	**78.9**	**83.5**	0.0	0.0	17.2	19.4	5.70	0.70
SpinePose-s	SpineTrack	37	79.1	82.1	77.5	82.9	**89.6**	**90.7**	**84.2**	**86.2**	5.98	0.72
RTMPose-m	Body8	26	**85.5**	**87.9**	**84.1**	**88.2**	0.0	0.0	19.4	21.4	13.93	1.95
SpinePose-m	SpineTrack	37	84.0	86.4	83.5	87.4	**91.4**	**92.5**	**88.0**	**89.5**	14.34	1.98
RTMPose-l	Body8	26	**86.8**	**89.2**	**86.9**	**90.0**	0.0	0.0	20.0	22.0	28.11	4.19
RTMW-m	Cocktail14	133	84.3	86.7	83.0	87.2	0.0	0.0	6.2	7.6	32.26	4.31
SimCC-ResNet50	COCO	17	81.8	85.2	0.0	0.0	0.0	0.0	0.0	0.2	36.75	5.50
SpinePose-l	SpineTrack	37	85.4	87.7	85.5	89.2	**91.0**	**92.2**	**88.4**	**90.0**	28.66	4.22
SimCC-ResNet50*	COCO	17	83.2	86.2	0.0	0.0	0.0	0.0	0.0	0.3	43.29	12.42
RTMPose-x*	Body8	26	**88.6**	**90.6**	**88.4**	**91.4**	0.0	0.0	21.0	22.9	50.00	17.29
RTMW-l*	Cocktail14	133	86.0	88.3	85.6	89.2	0.0	0.0	6.5	8.1	57.20	7.91
RTMW-l*	Cocktail14	133	87.3	89.9	88.3	91.3	0.0	0.0	6.9	8.6	57.35	17.69
SpinePose-x*	SpineTrack	37	86.3	88.5	86.3	89.7	**89.3**	**91.0**	**88.3**	**89.9**	50.69	17.37

References

[1] Tao Jiang, Peng Lu, Li Zhang, Ningsheng Ma, Rui Han, Chengqi Lyu, Yining Li, and Kai Chen. Rtmpose: Real-time multi-person pose estimation based on mmpose, 2023.

[2] Ajay Seth, Michael Sherman, Jeffrey A Reinbolt, and Scott L Delp. Opensim: a musculoskeletal modeling and simulation framework for in silico investigations and exchange. *Procedia Iutam*, 2:212–232, 2011.

Learn More

Feel free to reach out at **mukh07@dfki.de** if you have any more questions.