JULY 19–22
VANCOUVER, BC
2026 JOINT AAPM | COMP MEETING

Poserwgcn: Development and Preliminary Clinical Validation of a Real-Time 3D Human Pose Estimation Method for Radiotherapy Based on Rwkv and Graph Convolutional Networks

Lintao Song, MS[1]; Xinye Ni, PhD[1]

[1]Department of Radiotherapy, Changzhou No. 2 People's Hospital, Nanjing Medical University, Changzhou, Jiangsu, China

ABSTRACT

Patient pose variations during radiotherapy may reduce positioning accuracy and compromise treatment safety. We developed PoseRWGCN, a markerless and attention-free dual-stream framework for real-time 3D human pose estimation. The model combines an RWKV-based temporal stream for long-range motion modeling with a graph convolutional stream for learning skeletal topology, followed by adaptive feature fusion. It supports causal inference for real-time monitoring and bidirectional inference for offline analysis. A multi-scale alerting module was further designed to evaluate keypoint displacement, joint-angle variation, and projected skeletal-length changes. PoseRWGCN achieved MPJPEs of 41.9 mm and 42.2 mm under non-causal and causal settings on Human3.6M, respectively. On MPI-INF-3DHP, it achieved an MPJPE of 15.4 mm, a PCK of 98.6%, and an AUC of 85.6%. In radiotherapy-environment testing, the model obtained an MPJPE of 58.2 mm, a mean joint-angle error of 1.3987°, and an inference speed of approximately 20 fps. These results demonstrate the potential of PoseRWGCN for contact-free, real-time patient pose monitoring and deviation alerting during radiotherapy.

CONTACT

[Lintao Song]
[Department of Radiotherapy, Changzhou No. 2 People's Hospital, The Third Affiliated Hospital of Nanjing Medical University]
[Changzhou, Jiangsu, China]
[songlintao@stu.njmu.edu.cn]
[13753716971]

INTRODUCTION

Accurate and continuous monitoring of patient posture is essential for maintaining geometric precision and treatment safety during radiotherapy. Intrafraction pose variations may lead to deviations from the planned treatment position, potentially affecting dose delivery accuracy. However, conventional monitoring approaches often rely on patient-attached markers, contact sensors, or complex hardware configurations, which may limit patient comfort and clinical scalability.

Markerless 3D human pose estimation provides a promising contact-free alternative, but its application in radiotherapy requires simultaneous modeling of long-range temporal motion and spatial skeletal relationships while maintaining real-time causal inference. Therefore, we developed PoseRWGCN, an attention-free dual-stream framework integrating RWKV-based temporal modeling with graph convolutional skeletal learning. A multi-scale alerting strategy was further introduced to detect clinically relevant pose deviations and support real-time patient positioning surveillance.

RESULTS

PoseRWGCN achieved MPJPEs of 41.9 mm and 42.2 mm under non-causal and causal inference on Human3.6M, respectively. The difference of only 0.3 mm indicates that causal inference maintains comparable accuracy while enabling online monitoring. On MPI-INF-3DHP, the model achieved an MPJPE of 15.4 mm, a PCK of 98.6%, and an AUC of 85.6%.

In the radiotherapy-environment evaluation, PoseRWGCN achieved an MPJPE of 58.2 mm and a mean joint-angle error of 1.40°. The complete pipeline operated at approximately 20 fps, supporting real-time pose reconstruction and deviation alerting. Representative results demonstrated stable 3D skeletal reconstruction across different body postures.

Table 1. Comparison on Human3.6M under Protocol #1 and Protocol #2

Method	Inference	T	Params(M)	FLOPs(G)	MPJPE↓(mm)	P-MPJPE↓(mm)
MotionBERT-scratch	Causal	243	16.00	131.1	44.7	35.8
MotionAGFormer-S	Causal	81	4.80	6.6	47.7	37.7
HSTFormer	Bidirectional	81	22.70	2.1	42.7	33.7
MixSTE	Bidirectional	81	33.61	46.2	42.4	33.9
PoseRWGCN (Ours)	**Causal**	**81**	**4.09**	**2.2**	**42.2**	**34.6**

MATERIALS AND METHODS

Data Acquisition

Human3.6M and MPI-INF-3DHP were used for model development and benchmark evaluation. For preliminary validation in a radiotherapy environment, synchronized RGB images and reference 3D joint coordinates were collected from a volunteer using an RGB camera and an optical motion-capture system with 17 markers.

Figure 1. Radiotherapy-environment data acquisition setup: (a) placement of 17 motion-capture markers; (b) optical tracking and RGB camera configuration; and (c) representative RGB image.

PoseRWGCN Architecture

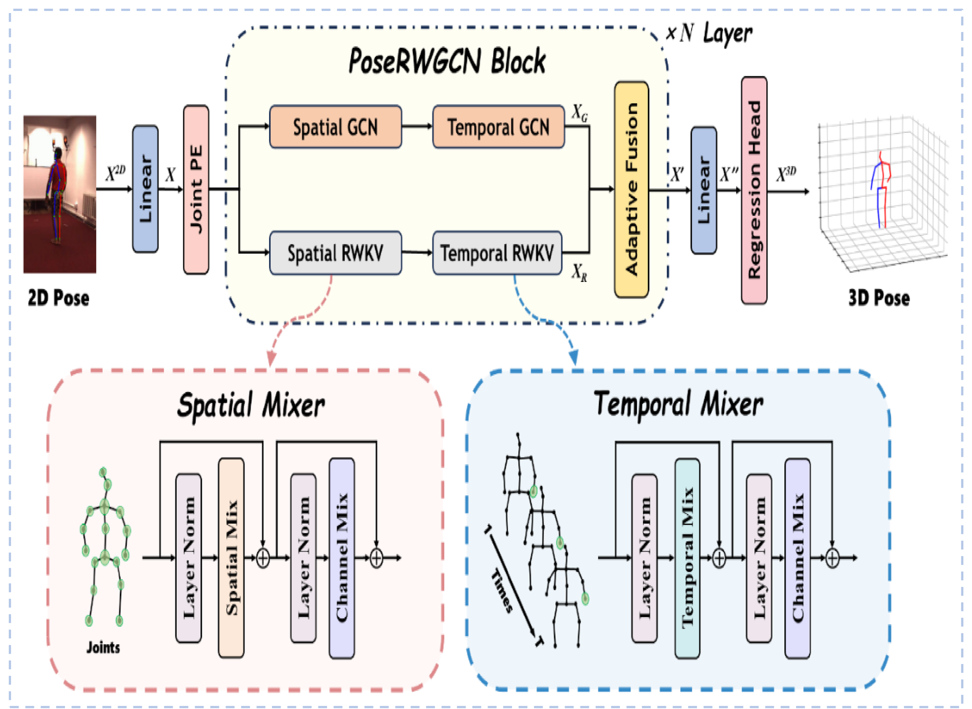

Figure 2. Overall architecture of PoseRWGCN, integrating RWKV-based temporal modeling and GCN-based skeletal topology learning.

- PoseRWGCN is an attention-free dual-stream network for 3D human pose estimation.
- The RWKV stream models long-range temporal dependencies through recursive gating.
- The GCN stream captures spatial relationships constrained by skeletal topology.
- Temporal and spatial features are integrated through an adaptive fusion module.
- Causal inference supports real-time online monitoring, whereas bidirectional inference is used for offline analysis.

Real-Time Alerting

- Pose deviations were quantified using keypoint displacement, joint-angle changes, and projected segment-length changes.
- Alerts were triggered when predefined deviation thresholds were exceeded.

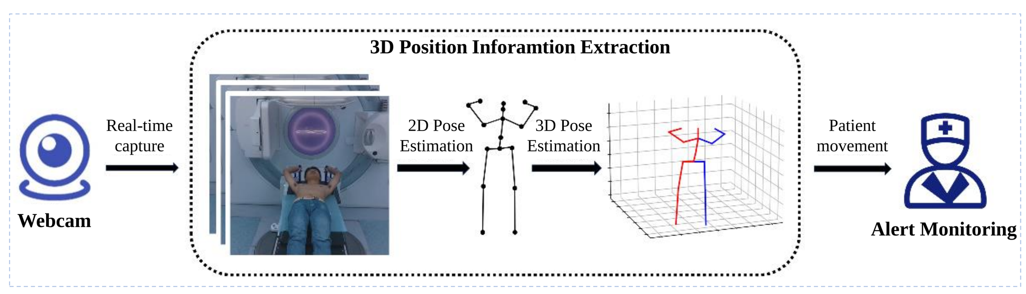

Figure 3. End-to-end workflow for real-time patient pose estimation and deviation alerting.

RESULTS

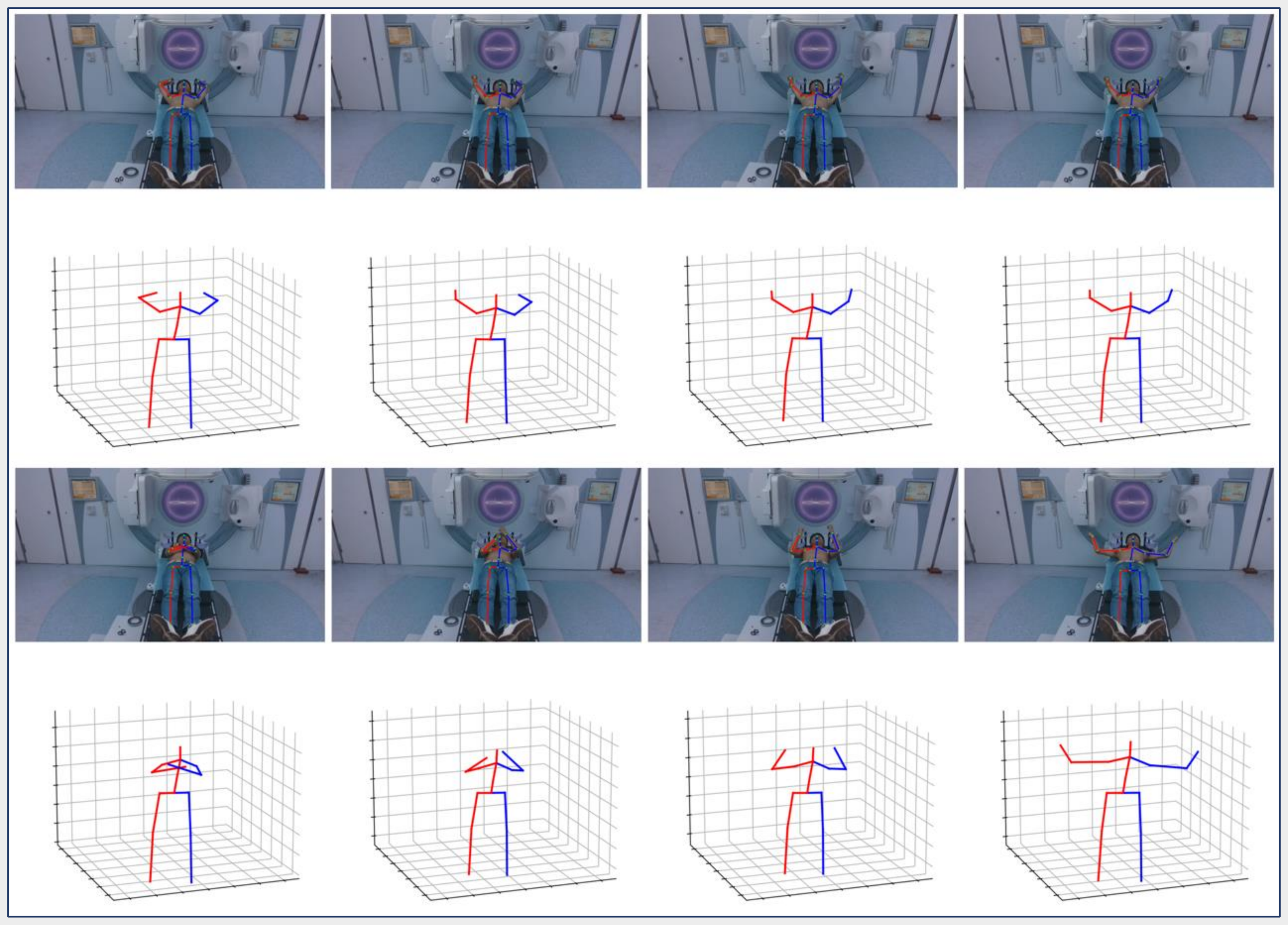

Figure 4. Representative RGB inputs and corresponding 3D pose reconstructions under different body postures in the radiotherapy environment.

CONCLUSIONS AND DISCUSSION

- PoseRWGCN showed comparable performance under causal and non-causal inference, with only a 0.3-mm increase in MPJPE in the causal setting, supporting real-time monitoring with minimal accuracy loss. By combining RWKV-based temporal modeling with GCN-based skeletal learning, the framework balances accuracy and computational efficiency.

- In the radiotherapy environment, PoseRWGCN achieved an MPJPE of 58.2 mm, a mean joint-angle error of 1.40°, and a processing speed of approximately 20 fps, demonstrating the feasibility of markerless, contact-free pose monitoring and deviation alerting.

- However, validation was limited to a small volunteer dataset. The system should therefore be considered a supplementary warning tool, and further patient-based clinical validation is required.