

Research using Deep learning to estimate 3D human pose applied in sports analysis and evaluation.

Lương Thị Thanh Minh

Abstract: This study investigates the application of deep learning for 3D human pose estimation to enhance sports performance analysis and evaluation. Traditional motion capture systems, while accurate, are often costly and require controlled environments, limiting their accessibility for routine training. We propose a monocular video-based framework leveraging convolutional neural networks (CNNs) and transformer architectures to reconstruct three-dimensional skeletal joint positions from single-camera footage. The model is trained on a combination of public datasets and a custom sports-specific dataset, incorporating data augmentation to handle occlusions and varied athletic movements. Quantitative evaluation demonstrates that our approach achieves competitive accuracy against marker-based systems, with a mean per-joint position error below 40mm for common sports actions. Qualitative analysis further validates its efficacy in identifying key biomechanical parameters, such as joint angles and limb symmetry, which are critical for technique correction and injury prevention. The findings suggest that deep learning-based 3D pose estimation offers a practical, scalable solution for real-time sports analytics, enabling detailed feedback without specialized hardware. This work contributes to the advancement of accessible, data-driven coaching tools.

Keywords: Deep learning, 3D human pose estimation, sports analysis, pose estimation