

## Spatiotemporal Neural Network for Video-Based Pose Estimation

**Authors :** Bin Ji, Kai Xu, Shunyu Yao, Jingjing Liu, Ye Pan

**Abstract :** Human pose estimation is a popular research area in computer vision for its important application in human-machine interface. In recent years, 2D human pose estimation based on convolution neural network has got great progress and development. However, in more and more practical applications, people often need to deal with tasks based on video. It's not far-fetched for us to consider how to combine the spatial and temporal information together to achieve a balance between computing cost and accuracy. To address this issue, this study proposes a new spatiotemporal model, namely Spatiotemporal Net (STNet) to combine both temporal and spatial information more rationally. As a result, the predicted keypoints heatmap is potentially more accurate and spatially more precise. Under the condition of ensuring the recognition accuracy, the algorithm deal with spatiotemporal series in a decoupled way, which greatly reduces the computation of the model, thus reducing the resource consumption. This study demonstrate the effectiveness of our network over the Penn Action Dataset, and the results indicate superior performance of our network over the existing methods.

**Keywords :** convolutional long short-term memory, deep learning, human pose estimation, spatiotemporal series

**Conference Title :** ICIP 2021 : International Conference on Image Processing

**Conference Location :** Paris, France

**Conference Dates :** September 20-21, 2021