**3D Human Pose Estimation Using Convolutional Neural Networks with 2D Pose Information**

1.what is the problem/topic?

🡪 3D Human Pose Estimation Using Convolutional Neural Networks with 2D Pose Information

2. why is it relevant?

🡪It is relevant because it uses 2d pose information to map 3d pose estimates, which is not studied in detail and is important to real life applications of all kinds like self-driven cars.

3. what have other people done to solve the problem?

🡪Early works for 2D human pose estimation which are based on deformable parts model, pictorial structure, or poselets train the relationship between body appearance and body joints using hand-crafted features. Similar to the 2D case, early stage of 3D human pose estimation is also based on the low-level features such as local shape context or segmentation results.

4.why is this not sufficient?

🡪 The previous models used like select vector and random forest do not give desired results in terms of efficiency and 2d to 3d translations. Hence we use CNN to do the above task and improve efficiency of the model.

5. what is the proposed solution?

🡪The solution in this paper is to provide an end-to-end learning framework to estimate 3D structure of a human body from a single image. 3D and 2D pose information are jointly learned in a single CNN. 2D classification results directly propagate to 3D pose regressors inside the CNNs. Using additional information such as 2D classification results and the relative distance from multiple joints, improve the performance of 3D human pose estimation over the baseline method.

6. why is the solution better?

🡪 By reusing 2D joint classification result, the relationship between 2D pose and 3D pose is implicitly learned during the training phase. Multiple regression results with different root nodes gives an effect of ensemble learning. When both strategies are combined, 3D pose estimation results are significantly improved and show comparable performance to the state-of-the-art methods without exploiting any temporal information of video sequences.

7. what is left/future work?

🡪 The performance can be further improved by incorporating temporal information to the CNN by applying the concepts of recurrent neural network or 3D convolution. Efficient aligning method for multiple regression results may boost the accuracy of pose estimation.