**2D/3D Pose Estimation and Action Recognition using Multitask Deep Learning**

1.what is the problem/topic?

🡪 2D/3D Pose Estimation and Action Recognition using Multitask Deep Learning.

2. why is it relevant?

🡪 The relevance of the paper is in to multitask 2D and 3D pose estimation using images and video sequence as action recognition and pose estimation are closely related to each other.

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

🡪 The previous methods used are two kinds , detection based and regression based methods. Detection based methods handle pose estimation as a heat map prediction problem, where each pixel in a heat map represents the detection score of a corresponding joint.

4.why is this not sufficient?

🡪 Detection approaches do not provide joint coordinates directly. To recover the pose in (x, y) coordinates, the argmax function is usually applied as a post-processing step.

5. what is the proposed solution?

🡪 The solution is to use a regression based method. Regression based approaches use a nonlinear function that maps the input directly to the desired output, which can be the joint coordinates.

6. why is the solution better?

🡪 The main advantage of regression methods over detection ones is that they often are fully differentiable. This means that the output of the pose estimation can be used in further processing and the whole system can be fine-tuned.

7. what is left/future work?

🡪 The problem of regression method is that regression function is frequently sub-optimal. In order to tackle this weakness, the Soft-argmax function has been proposed to convert heat maps directly to joint coordinates and consequently allow detection methods to be transformed into regression methods.