1 Introduction

- Imaged-based:
    - Geometry between joints in the target image:
        * Poselet conditioned pictorial structures
        * Exploring the spatial hierarchy of mixture models for human pose estimation
        * Articulated human detection with flexible mixtures of parts
    - Convolutional Pose Machine
    - Stacked Hourglass
    - OpenPose
    - HRNet

1.1 Related Work

2-dimensional pose estimation can be divided into either being image-based or video-based. Image-based methods [**MANGLER**]... . Video-based methods commonly use the correlating information among the frames of the video to perform the pose estimation. Early video-based methods used 3-dimensional convolutions to capture the correlating information between neighboring frames [5, 2]. Other methods use LSTM's [3] to capture the correlating information among the frames [4, 1]. Recently, transformers [6] have started to being used as a way of capturing the correlating information among the frames [7].

1.2 Problem Definition

1.3 Reading Guide