Human pose estimation is defined as the localization of anatomical keypoints or landmarks (which we refer to as parts) and is tackled using various methods, depending on the final goals and the assumptions made:

1. The use of single or sequential images as input;
2. The use (or not) of depth information as input;
3. The localization of parts in a 2D or 3D space;
4. The estimation of single- or multi-person poses

The goal of this thesis is to localize the 2D feature points of the table tennis player as well as the racket. Such feature points need be manually labeled from a self-built dataset containing several parts of the existing datasets (COCO, MPII), YouTube videos and captured images in our table tennis room. Since table tennis is one of the fastest sports, an efficient approach, called *pose proposal networks*, should be implemented, which is a single-shot object detection paradigm using grid-wise image feature maps in a bottom-up pose detection scenario. The entire architecture is constructed from a single, convolutional neural network with relatively low-resolution feature maps and is optimized end-to-end directly using a loss function designed for pose detection performance.

Requirements:

- Python
- TensorFlow
- English speaking

**Kontakt**

Yapeng Gao  
Sand 1, Raum A312  
Tel.: 07071 29 78970  
yapeng.gao@uni-tuebingen.de

Dieses Angebot finden Sie auch auf unserem Webserver: http://www.cogsys.cs.uni-tuebingen.de