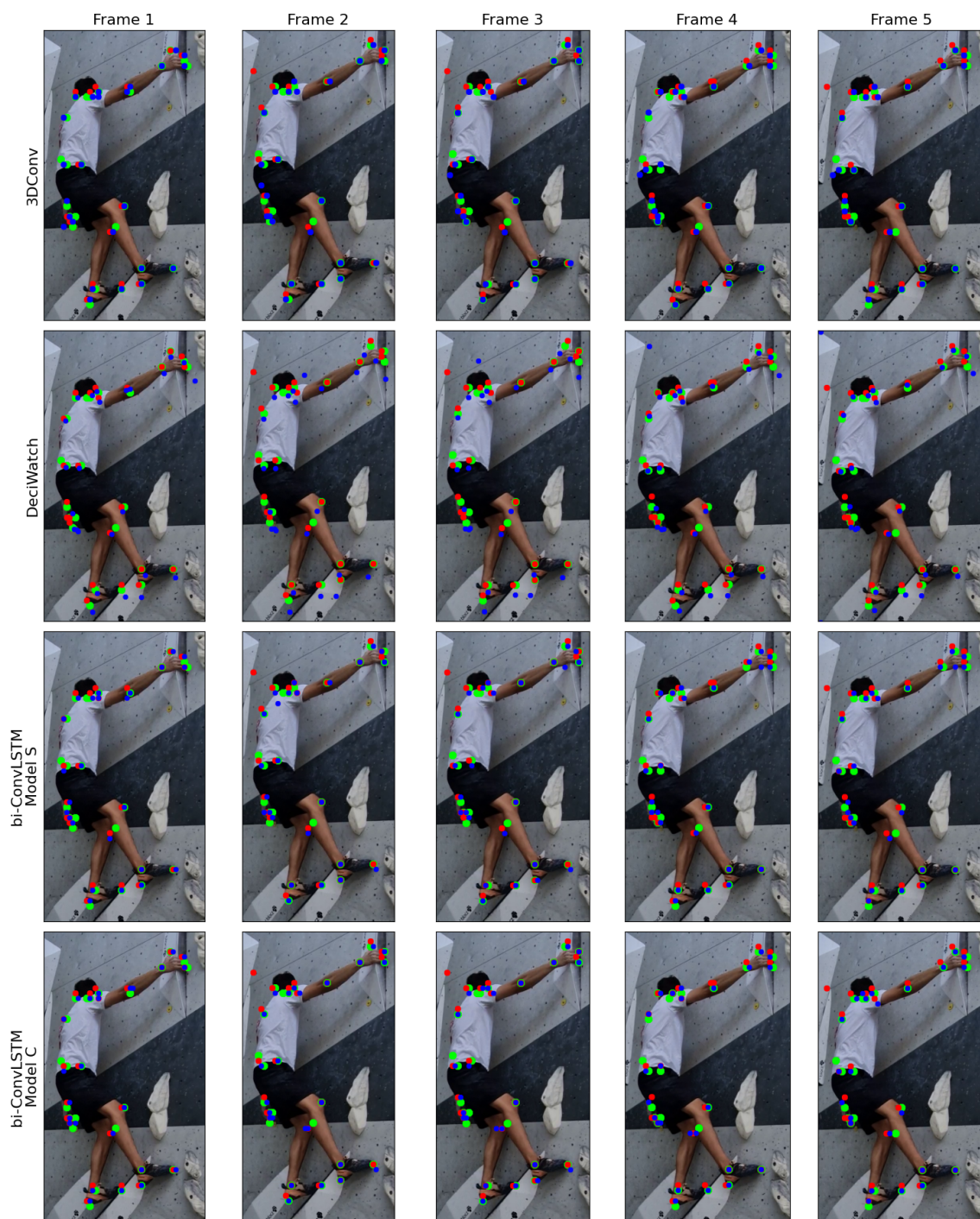# Appendix

Figure 1: Ground truth keypoints (green), Mask R-CNN predictions (red) and predictions of the best setting of each model from section **??** on five consecutive frames from the single hold-out video.