Appendix

Experiment	3DConv 1.1	3DConv 1.2	3DConv 1.3	DeciWatch 1.1	DeciWatch 1.2	DeciWatch 1.3	bi-ConvLSTM Model S 1.1	bi-ConvLSTM Model S 1.2	bi-ConvLSTM Model S 1.3	bi-ConvLSTM Model C 1.1	bi-ConvLSTM Model C 1.2	bi-ConvLSTM Model C 1.3	Total
Nose	38.6	50.1	44.5	76.8	76.7	68.3	51.2	23.2	34.7	31.0	23.1	26.7	45.4
Ear	33.4	40.3	37.4	77.3	77.2	69.1	40.0	28.5	35.9	44.3	31.1	55.3	47.5
Shoulder	51.4	62.1	59.9	77.5	77.7	69.7	42.0	34.8	28.6	34.6	45.6	33.0	51.4
Elbow	61.5	60.8	62.3	76.7	76.8	67.4	39.5	24.2	36.0	38.9	47.1	50.7	53.5
Wrist	44.0	46.5	44.7	75.9	75.9	67.0	30.4	34.1	40.0	34.9	39.5	52.4	48.8
Pinky	30.5	29.1	27.4	73.9	74.1	65.7	19.6	28.9	32.0	15.8	20.5	24.7	36.9
Index finger	39.1	38.9	40.4	74.6	75.0	66.4	32.4	33.0	40.8	34.0	40.3	34.3	45.8
Thumb	35.7	35.8	36.7	74.8	74.7	66.1	21.6	36.3	33.3	29.5	25.0	24.8	41.2
Hip	79.3	76.9	87.2	79.9	80.2	71.3	45.4	50.8	51.9	42.9	50.4	40.8	63.1
Knee	54.3	61.2	53.6	77.8	78.0	69.0	35.0	33.8	47.4	31.6	30.8	26.8	49.9
Ankle	51.2	57.6	55.5	78.7	78.9	67.8	49.9	43.7	43.5	50.2	57.2	40.8	56.3
Heel	64.4	62.8	69.0	77.9	77.5	67.0	34.3	33.3	38.0	37.8	36.5	36.2	52.9
Toes	57.9	57.0	61.2	76.4	76.4	68.6	57.2	43.8	44.0	40.7	51.6	53.0	57.3
Total	49.7	52.3	53.1	76.6	76.7	68.1	37.8	34.9	39.0	35.9	39.0	38.5	

Table 1: Keypoint-specific testing PCK@0.05-accuracies of the various models for shiting-scalar $s = 1$. All the accuracies are in percentage.

Experiment	3DConv			DeciWatch			bi-ConvLSTM Model S			bi-ConvLSTM Model C			Total
	1.1	1.2	1.3	1.1	1.2	1.3	1.1	1.2	1.3	1.1	1.2	1.3	
Nose	33.0	49.3	37.7	74.3	74.5	22.4	33.1	32.3	36.0	22.6	27.4	30.7	39.4
Ear	30.8	35.1	30.3	72.9	72.9	43.8	36.4	33.2	40.2	29.7	37.1	45.8	42.4
Shoulder	48.7	56.6	51.5	77.0	77.0	31.0	49.9	45.3	27.0	40.2	41.5	37.9	48.6
Elbow	56.8	58.8	54.4	76.6	76.6	30.2	34.2	47.8	37.1	54.8	45.6	32.6	50.5
Wrist	43.5	46.0	41.2	74.7	74.7	36.3	25.6	35.2	40.0	34.1	40.7	39.6	44.3
Pinky	28.6	29.5	28.2	73.6	74.0	30.6	27.5	20.6	28.0	21.2	24.8	31.3	34.8
Index finger	39.3	40.7	40.7	74.8	74.5	30.6	42.2	33.4	44.7	37.7	30.3	38.2	43.9
Thumb	36.0	35.3	36.6	74.9	74.4	28.1	28.1	28.5	16.1	24.2	24.3	24.6	35.9
Hip	78.2	75.7	82.6	80.0	80.1	42.7	55.3	27.6	53.9	53.9	59.9	8.22	58.2
Knee	39.7	57.5	37.5	78.3	78.1	24.8	44.8	26.1	17.4	47.1	42.8	50.7	45.4
Ankle	37.2	62.9	39.4	75.8	76.0	53.9	46.3	46.1	34.9	51.2	49.2	45.7	51.6
Heel	65.3	60.3	68.8	78.2	77.8	55.1	24.3	32.8	40.1	42.5	29.4	37.7	51.0
Toes	61.4	62.4	60.0	75.8	75.9	32.8	54.1	72.1	42.1	39.1	53.0	56.0	57.1
Total	46.5	51.6	47.3	76.0	75.9	36.8	38.8	37.4	35.9	39.2	39.5	37.1	

Table 2: Keypoint-specific testing PCK@0.05-accuracies of the various models for shiting-scalar $s = 2$. All the accuracies are in percentage.

Experiment	3DConv			DeciWatch			bi-ConvLSTM Model S			bi-ConvLSTM Model C			Total
	1.1	1.2	1.3	1.1	1.2	1.3	1.1	1.2	1.3	1.1	1.2	1.3	
Nose	100	100	100	99.8	99.8	99.8	100	100	99.9	100	99.7	99.9	99.9
Ear	100	100	100	99.8	99.8	99.8	97.7	99.9	100	100	100	99.9	99.7
Shoulder	99.9	100	99.9	99.8	99.8	99.8	100	100	99.9	100	100	100	99.9
Elbow	99.9	99.9	99.9	99.4	99.4	99.4	100	100	100	100	99.9	100	99.8
Wrist	99.8	99.9	99.9	99.1	99.2	99.1	100	99.9	99.8	100	99.9	100	99.7
Pinky	93.4	93.1	94.4	98.3	98.4	98.3	97.2	98.8	97.0	98.0	99.0	98.6	97.0
Index finger	99.0	98.8	98.8	98.2	98.2	98.2	99.5	98.7	97.0	99.6	99.4	99.4	98.7
Thumb	98.9	98.8	98.9	98.3	98.3	98.3	96.8	99.6	97.8	99.7	98.6	99.6	98.6
Hip	99.9	100	100	99.7	99.7	99.7	100	100	100	100	100	100	99.9
Knee	100	100	99.9	99.7	99.7	99.7	100	100	100	100	100	100	99.9
Ankle	100	100	100	99.5	99.5	99.5	100	100	99.9	100	100	99.9	99.9
Heel	100	100	100	99.2	99.2	99.2	99.3	99.9	99.9	99.9	100	99.8	99.7
Toes	99.9	100	100	99.1	99.0	99.1	99.6	99.8	99.4	99.8	100	99.8	99.6
Total	99.2	99.3	99.3	99.2	99.2	96.1	99.4	99.7	99.2	99.8	99.7	99.7	

Table 3: Keypoint-specific testing PCK@0.2-accuracies of the various models for shiting-scalar $s = 1$. All the accuracies are in percentage.

Experiment	3DConv			DeciWatch			bi-ConvLSTM Model S			bi-ConvLSTM Model C			Total
	2.1	2.2	2.3	2.1	2.2	2.3	2.1	2.2	2.3	2.1	2.2	2.3	
Nose	100	100	100	99.8	99.8	97.2	100	99.9	99.7	99.7	99.9	100	99.7
Ear	100	99.8	100	99.7	99.7	97.1	99.8	99.8	100	99.9	99.9	99.9	99.6
Shoulder	99.9	99.7	99.9	99.8	99.9	95.8	99.9	99.8	100	99.9	100	100	99.6
Elbow	99.8	99.9	99.9	99.5	99.4	90.8	100	99.5	100	100	100	100	99.1
Wrist	99.8	100	99.9	99.1	99.1	94.4	99.8	99.7	99.8	99.8	100	99.7	99.3
Pinky	93.1	93.7	93.9	98.4	98.4	86.5	97.7	98.0	97.9	99.1	96.0	98.2	95.9
Index finger	98.9	99.0	98.8	98.2	98.2	88.9	99.4	99.1	99.2	99.6	98.0	97.5	97.9
Thumb	98.6	98.6	98.6	98.3	98.2	90.1	95.7	96.6	98.6	97.5	98.3	99.6	97.4
Hip	100	99.9	100	99.7	99.8	95.1	99.9	99.8	99.9	99.8	100	99.9	99.5
Knee	100	100	99.9	99.6	99.7	95.1	100	99.9	99.9	99.9	99.9	100	99.5
Ankle	100	99.9	100	99.5	99.5	93.9	100	100	100	100	100	100	99.4
Heel	100	100	100	99.3	99.3	91.0	99.9	99.9	99.8	99.9	99.6	99.9	99.1
Toes	99.9	100	100	99.0	99.1	91.9	99.9	99.4	98.4	99.8	99.6	99.8	98.9
Total	99.2	99.3	99.2	99.2	99.2	92.8	99.4	99.5	99.3	99.6	99.3	99.6	

Table 4: Keypoint-specific testing PCK@0.2-accuracies of the various models for shiting-scalar $s = 2$. All the accuracies are in percentage.
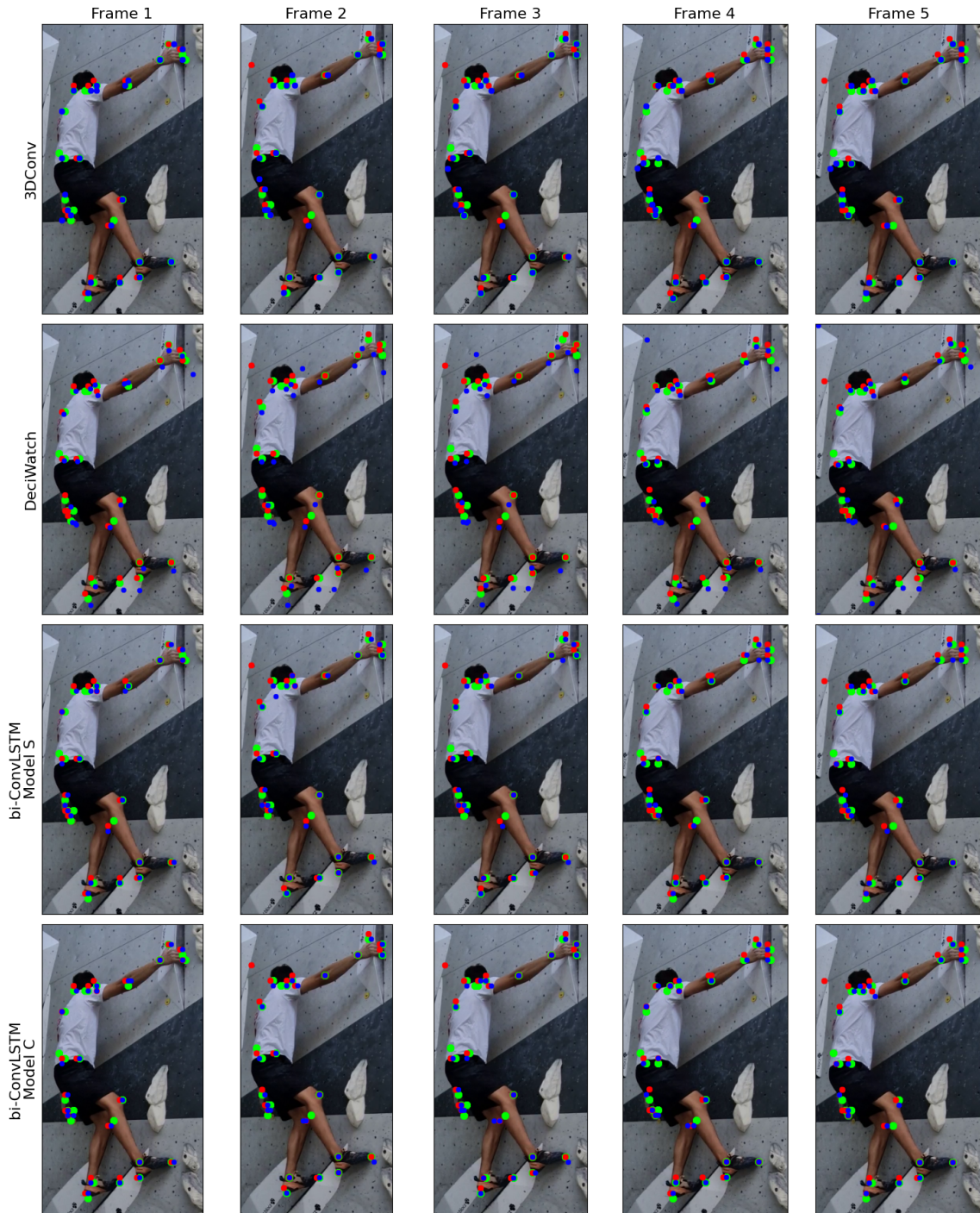
Figure 1: Ground truth keypoints (green), Mask R-CNN predictions (red) and predictions of the best setting of each model from section **??** on five consecutive frames from the single hold-out video.