3D-Yoga: A 3D Yoga Dataset for Visual-based Hierarchical Sports Action Analysis

Jianwei Li[1], Haiqing Hu[1], Jinyang Li[1], and Xiaomei Zhao[2]

[1] Beijing Sports University, Beijing 100084, P.R. China
[2] Shandong Jianzhu University, Jinan 250101, P.R. China
{jianwei,hhq,ljy}@bsu.edu.cn,zhaoxiaomei20@sdjzu.edu.cn

1 3D-Yoga Dataset

1.1 Subjects information

There are 22 subjects (7 males and 15 females) participated in the capturing of yoga poses. Each subject has a different BMI and yoga level. The specific information of them is shown in Table 1 and Table 2.

Table 1: Basic information of the subjects

Body mass index	Male	Female	Number
BMI>26	2	4	6
19<BMI<23	3	8	11
BMI<18	2	3	5

Table 2: Clothing information of the subjects

Clothing types	Male	Female	Number
Sports wear	2	4	6
Yoga wear	1	6	7
Sleep wear	1	2	3
Leisure wear	3	3	6

1.2 Occluded joints analysis

We have calculated the average numbers of joints affected by self-occlusion in Classification I. Fig. 1 shows the comparison of 2.5D skeletons from the front and side cameras, and 3D fused skeletons. The vertical axis represents the average number of occluded joints under each category. It is clear that the number of occluded joints (red pillars) in 3D fused skeletons is lower than the original skeletons collected from the front and side cameras.

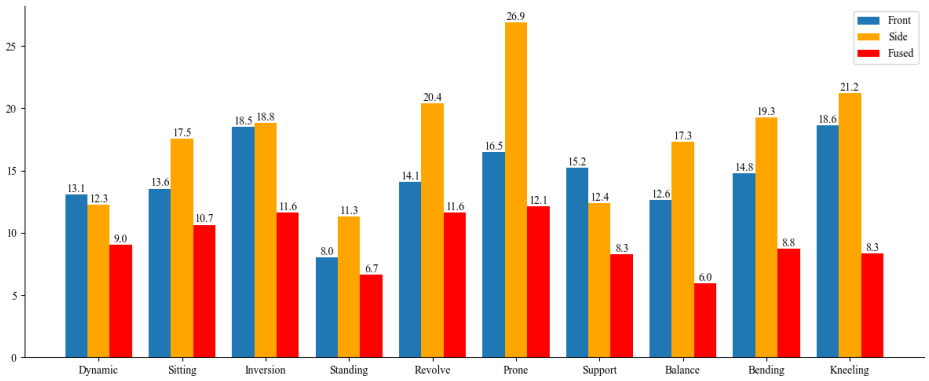

Fig. 1: The average numbers of occluded joints in Classification I.

1.3 Pose classification and assessment

Referenced to the exercises in *daily yoga*, there are 158 categories of yoga poses captured by us. In 3D-Yoga, the categories of yoga poses is adjusted to 117 to ensure that each label is different but covers all yoga formulas. Table 3, 4 and 5 show the two-level classification of these yoga poses. The criteria of completion score and difficulty coefficient for pose quality assessment are shown in Table 6 and Table 7 respectively.

2 Methods and Analysis

2.1 Cascade 2S-AGCN

Fig. 2 shows the architecture of the Cascaded 2S-AGCN, which consists of three levels. It is constructed by using 2S-AGCN models and fully connected layers with Dropout and ReLU. The trunk network and branches of Cascade 2S-AGCN are trained together. The input is the skeleton data, and the outputs of each level are the predicted results of Classification I, Classification II, and completion score respectively.

2.2 Experimental results

Due to the serious self-occlusion problem of yoga poses captured from a single view, the poor quality of original data may lead to poor fusion effect. According to the fusion quality, the fused 3D skeletons are divided into three parts: normal (N, 63.56%), poor (P, 15.75%) and bad (B, 20.69%). Where normal data is automatically fused with our method, poor data is fused with some manual corrections, and bad fused data has errors without corrections.

Table 8 shows the detailed quantitative analysis accuracies of yoga poses with Cascade 2S-AGCN. It can be seen that the analysis accuracies of most yoga poses (apart from *Revolve* and *Prone*) are higher when using fused 3D skeletons. Some accuracies of using combined data are also relatively high because the number of combined data is

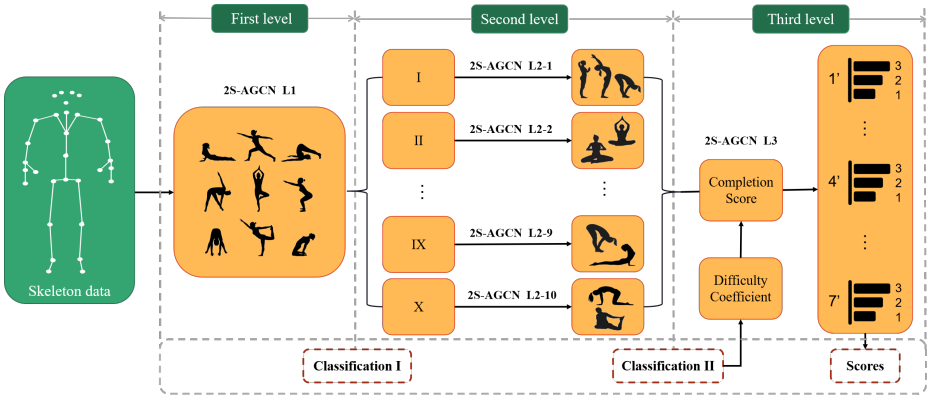

Fig. 2: Three-level cascaded 2S-AGCN networks.

larger than single view data and fused data. Table 9 shows detailed comparison results of the recognition accuracies (%) with ST-GCN, 2S-AGCN, and CTR-GCN. It is clear that the average accuracy of our method outperforms the state-of-the-art methods. The accuracies of ST-GCN, 2S-AGCN and our method are both improved by using the fused 3D skeletons in term of normal and poor data. Note that we interpolate the key frame data into sequence data with ten frames, so the accuracies of CTR-GCN are relatively higher by using single view data and combined data.

Fig. 3 shows visualization results of two yoga poses assessment for 4 subjects compared with the coach in the front view, and it can be seen that the performances and the scores are unified. The first subject is a professional yoga-trained practitioner, and her score is significantly higher than the other three amateur subjects.

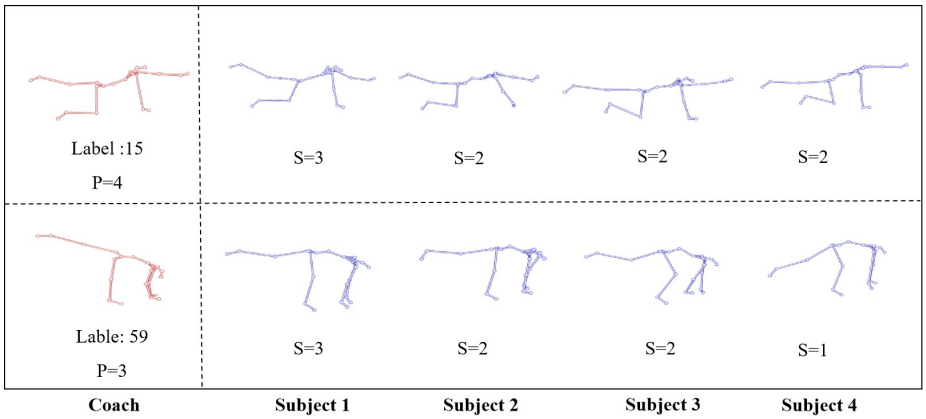

Fig. 3: Visualization examples of yoga pose assessment for 4 subjects.

Due to that most mainstream datasets for skeleton-based action analysis lack hierarchical semantic annotations, we cannot directly make experiments with our method. As shown in Table 10, we have clustered the semantic labels of NTU RGB+D 60 [2] into four categories. We tried to carry out the cross-subject experiment with our method by using the data of the C1 camera, and the coarse-gained and fine-gained accuracies (%) of single person poses are 87.97 and 76.03 respectively (learning rate = 1e-3, batch size = 25, and epoch = 50).

References

1. Chen, Y., Zhang, Z., Yuan, C., Li, B., Deng, Y., Hu, W.: Channel-wise topology refinement graph convolution for skeleton-based action recognition. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 13359–13368 (2021)
2. Shahroudy, A., Liu, J., Ng, T.T., Wang, G.: Ntu rgb+d: A large scale dataset for 3d human activity analysis. In: IEEE Computer Society. pp. 1010–1019 (2016)
3. Shi, L., Zhang, Y., Cheng, J., Lu, H.: Two-stream adaptive graph convolutional networks for skeleton-based action recognition. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 12026–12035 (2019)
4. Yan, S., Xiong, Y., Lin, D.: Spatial temporal graph convolutional networks for skeleton-based action recognition. In: Thirty-second AAAI conference on artificial intelligence (2018)

Table 3: Poses classification index I.

# Labels	Classification II (pose name)	Original Labels (1-158)	Classification I
1	cat pose to childs pose	134	
2	half monkey pose to dragon pose (L&R)	138	I.
3	standing apart to wreath pose	141	Dynamic pose
4	easy sun salutation	82	
5	angle seated	8,25,47,80,142	
6	bound angle pose/butterfly pose	65	
7	dynamic vizilla asana	36,51	
8	hero pose	144	
9	hunker pranayama	67,112,123	II.
10	sitting mountain pose	124	Sitting
11	sitting pranayama	12,145	
12	vizilla asana (L&R) alternate nostril breathing	64	
13	vizilla asana bird king pose (L&R)	63	
14	vizilla asana cow face pose (L&R)	53	
15	cow face pose (L&R)	99	
16	downward facing dog pose	16,127	III.
17	one-legged downward facing dog pose (L&R)	17,108	Inversion
18	chair pose/powerful pose	78,89,113	
19	high lunge (L&R)	22	
20	mountain pose	43,70,79,81	
21	pull back chair pose/powerful pose	19	
22	Standing pranayama	1	IV.
23	variant chair pose/powerful pose	18	Standing
24	variant chair bird King pose/powerful pose (L&R)	90	
25	warrior II pose (L&R)	84	
26	cervus (L&R)	101	
27	dynamic revolved pagoda	2	
28	dynamic revolved wide legged forward bend	4	
29	easy revolved pose (L&R)	125	
30	half lord of the fishes/seated spinal twist (L&R)	26	V.
31	prone arm stretching (L&R)	62	Revolve
32	reclined spine revolved (L&R)	10,27,35,153	
33	revolved high lunge (L&R)	7	
34	revolved lunge (L&R)	88	
35	revolved reclined (L&R)	97	
36	easy vinyasa	87	
37	pigeon pose (L&R)	74	
38	prone arm extending	61	
39	sleeping swan pose (L&R)	128	VI.
40	corpse	29,103	Prone pose
41	dynamic happy baby	28	
42	happy childs pose	111	
43	reclined butterfly pose	11,15	

Table 4: Poses classification index II.

# Labels	Classification II (pose name)	Original Labels (1-158)	Classification I
44	reclined butterfly warm uterus pose	66	
45	reclined feet-catching pranayama	94,95,104	
46	reclined warm uterus breathing	133	VI.
47	rock and roll	154	Prone pose
48	supported recline leg up (L&R)	98	
49	mermaid pose (L&R)	121	
50	cat pose	13,39,135,146	
51	dragon pose (L&R)	109	
52	easy dragon pose	5	
53	fishtailing cat pose	136	VII.
54	half monkey pose (L&R)	86	Support
55	quadruped support	60	
56	supported dragon pose (L&R)	137	
57	variant cat pose	14,147	
58	balancing gate pose (L&R)	37	
59	boat pose	156	
60	cross balancing I (L&R)	15	
61	cross balancing II (L&R)	56	
62	easy warrior III pose (L&R)	83	
63	goddess pose	45	VIII.
64	half boat pose	152	Balance
65	one-legged knee held (L&R)	114	
66	side plank pose (L&R)	30	
67	split boat	93	
68	squat pose	116	
69	tree balancing pose	44	
70	back-stretching pose	102,157	
71	elbow-holding seated forward bend	42	
72	feet-catching seated forward bend	126	
73	half bound extended angle pose/butterfly pose (L&R)	9	
74	one-legged back extended (L&R)	75,100	
75	reclined leg up	105	
76	reclined peak bending	96	IX.
77	seated forward bend	6,117,130	Bending
78	side extended wreath pose (L&R)	140	
79	sitting forward bending	68	
80	sitting leg up (L&R)	76	
81	standing forward bend	20	
82	strap-on wreath (L&R)	46	
83	straddle seated forward bend	91	
84	supported downward facing dog pose	107	
85	supported seated forward bend	110	
86	wide legged forward bend	3,115	

Table 5: Poses classification index III.

# Labels	Classification II (pose name)	Original Labels (1-158)	Classification I
87	wreath pose	77,92,139	
88	wreath pose II	24	
89	wide angle stretching	132	
90	wreath seated forward bend	48	
91	back bending Vizilla Asana	54	
92	camel pose	118	
93	crescent pose (L&R)	21	
94	half reclined hero pose	143	
95	locust pose	150	
96	seated butt lift back-bending (L&R)	69	
97	shoulder bridge pose	122	
98	sphinx pose	148	
99	upward facing dog pose	151	IX. Bending
100	back bending downward prayer warrior I pose (L&R)	71	
101	variant one-legged locust pose (L&R)	149	
102	extended side angle pose (L&R)	33,40,85	
103	extended triangle pose (L&R)	32,41	
104	gate pose (L&R)	34	
105	intense side stretch pose (L&R)	72	
106	kneel sideways (L&R)	31	
107	side bending gate pose (L&R)	38	
108	side bending half bound angle (L&R)	129	
109	side bending vizilla asana (L&R)	52	
110	arm-circling cat pose (L&R)	55	
111	equestrian pose feet-catching	59	
112	equestrian pose (L&R)	57	
113	baby downward prayer pose	50,158	X.
114	childs pose (pranayama)	49,73	Kneeling
115	dynamic equestrian pose (L&R)	58	
116	rabbit pose	120,131	
117	worship	23,106,119	

Table 6: Criteria description of completion score (S) for yoga pose quality assessment.

Scores	#Criteria	Descriptions
3	S	Keep flat when the limbs or body completes the support tasks; The support point consistent with the standard action.
	B	Support parts and body should be kept medium and stable; And parts that should not contact the ground leave the ground completely.
	P&T	Fully complete the bending angle in standard action.
2	S	Minor unreasonable bending of the limbs or body, Or support point are nuanced from the standard movement.
	B	Support parts and body produces a small displacement or the body has shaking; And parts that should not contact the ground leave the ground completely.
	P&T	Complete 75% of the bending angle in the standard action; Or beyond the standard action angle.
1	S	The limbs or body is greatly unreasonably bent when completing support tasks; And support points are different from the standard action.
	B	Support parts and body produce a small displacement or the body has shaking; Or parts of the body should not touch the ground partially contact the ground.
	P&T	Complete 50% of the bending angle in the standard action; Or the limbs should be kept straight have a small bend.
0	S	The limbs or body is greatly unreasonably bent when completing support tasks; And support points are more different from the standard action.
	B	Support parts and body produces a large displacement or the body has shaking; Or parts of the body should not touch the ground completely contact the ground.
	P&T	The completion range in the flexion action is very small; And the limbs that should be kept straight are greatly bent.

Table 7: Criteria of difficulty coefficient (P) for yoga pose quality assessment.

Scores	# Lables
1	9,10,11,12,20,22,38,40,41,42,46,50,55,75,76,112,114,117
2	1,7,13,14,16,18,23,26,27,28,29,39,47,53,57,63,64,84,93,98,102,103,109,110,113,115
3	21,24,25,30,31,32,35,36,37,49,52,54,56,59,62,65,69,73,74,83,85,97,99,111
4	2,5,15,17,19,43,44,45,48,51,60,66,68,70,77,81,91,96,100,101,116
5	6,8,33,34,58,67,71,72,78,79,87,88,92,94,95,105,106
6	3,4,61,82,86,6,90
7	80,104,107,108

Table 8: Analysis accuracies (%) of yoga poses in 3D-Yoga. The best results are in bold.

Stages	Categories	Number	Front	Side	Combined	Fused		
						N	N+P	N+P+B
I	Coarse-gained	10	75.60	71.87	77.18	**80.42**	76.98	74.64
	Dynamic	4	77.6	68.86	78.14	73.73	**78.79**	78.68
	Sitting	11	74.65	69.91	85.64	**87.29**	77.84	70.50
	Inversion	2	**100**	93.48	**100**	92.31	**100**	95.65
	Standing	8	83.77	72.08	81.81	**90.40**	90.07	84.41
	Revolve	10	76.98	**79.86**	77.70	73.56	71.06	76.98
II	Prone	14	61.68	65.87	**76.05**	60.68	63.09	66.64
	Support	8	54.17	62.50	60.41	53.84	**63.35**	61.81
	Balance	12	82.96	82.22	80.00	92.30	**92.68**	87.41
	Bending	40	64.09	62.19	67.68	**75.52**	66.19	62.34
	Kneeling	8	60.43	69.06	66.91	69.33	**72.38**	61.87
	Fine-gained	117	76.10	74.13	74.92	**82.94**	78.48	70.12
III	Score	4	71.40	70.97	65.43	**72.57**	71.48	70.12

Table 9: Comparison of recognition accuracies (%) with the state-of-the-art methods on 3D-Yoga. The best results are in bold.

Methods	Front	Side	Combined	Fused			Average
				N	N+P	N+P+B	
ST-GCN [4]	53.65	56.25	56.56	58.33	59.57	56.29	56.78
2S-AGCN [3]	58.90	55.82	56.86	66.84	60.33	58.23	59.50
CTR-GCN [1]	73.02	66.24	**76.40**	72.93	65.38	58.65	68.77
Ours	**76.10**	**74.13**	74.92	**82.94**	**78.48**	**74.36**	**76.82**

Table 10: Two-level labels for single person actions in NTU-60 dataset.

First level labels	Second level labels
I. Upper limb movement	A1,A2,A3,A4,A5,A7,A10,A11,A12,A13,A18,A19,A20,A21,A23, A25,A28,A29,A30,A31,A34,A37,A38,A39,A40,A44,A47,A49
II. Lower limb movement	A8,A9,A24,A26,A27,A42
III. Head movement	A35,A36,A41
IV. Head movement	A6,A14,A15,A16,A17,A22,A32,A33,A43,A45,A46,A48,A50