Supplementary Materials for
LucidAction: A Hierarchical and Multi-model Dataset for Comprehensive Action Quality Assessment

Linfeng Dong[1,2], Wei Wang[2], Yu Qiao[2], and Xiao Sun[2]

[1]Zhejiang University
[2]Shanghai Artificial Intelligence Laboratory
{donglinfeng, wangwei, sunxiao}@pjlab.org.cn, yu.qiao@siat.ac.cn

A.   Experiment

A.1   Datasets

AQA-7 [4] contains actions from seven different sport events, including gymnastic vaulting, big air skiing, big air snowboarding, synchronous diving 3m springboard, synchronous diving 10m platform and trampoline. Following the setting in [4, 1], we excluded the trampoline event with much longer videos than the other categories, resulting in 803 training videos and 303 testing videos. We use final score as assessment ground truth.

MTL-AQA [5] contains 1412 samples of diving covering individual, synchronous, 3m springboard and 10m platform events. Following the setting in [4, 1], we split the dataset into 1059 training samples and 353 testing samples. We use performance score as assessment ground truth. The performance score is given by final score divided by the difficulty value of actions in the sample. Difficulty values are known information during both training and inference.

FineDiving [8] contains 3000 samples of diving, covering 52 action categories, we split the dataset into 2251 training samples and 749 testing samples following the setting in [8]. We use performance score as assessment ground truth. The performance score is given by final score divided by the difficulty value of actions in the sample. Difficulty values are known information during both training and inference.

LucidAction is our newly proposed dataset which contains 6702 samples across 8 sport events including men's/women's floor exercise (MFE, WFE), vault (MVT, WVT), men's parallel bars (MPB), horizontal bars (MHB), women's uneven bar (WUB) and balance beam (WBB). We split the dataset into 5026 training samples and 1676 testing samples. Specially, we provide at least one perfect training sample which is free of penalty for each action category. The purpose of this split is to ensure that when using contrastive regression models for assessment, the reference exemplar itself is error-free. This avoids the ambiguity that execution errors in the exemplars might introduce into the scoring process. The ground truth used in the experiments is the performance score, which is obtained by subtracting all penalty scores from a perfect score of 5.0. More details of penalty scores is provided in section B.2.

A.2 Evaluation Metrics

We employ two metrics in our experiments: Spearman's rank correlation (ρ) and relative L2 distance(R-ℓ_2). Spearman's rank correlation assesses the rank correlation between predictions and ground-truth scores, defined as:

$$\rho = \frac{\sum_i \left(p_i - \bar{p}\right)\left(q_i - \bar{q}\right)}{\sqrt{\sum_i \left(p_i - \bar{p}\right)^2 \sum_i \left(q_i - \bar{q}\right)^2}} \tag{1}$$

where p and q represent the rankings of predictions and ground-truth scores, respectively. The relative L2 distance focuses on the numerical scoring difference between predictions and ground-truth scores. Given the highest and lowest scores of an action, denoted as $y_{\max}$ and $y_{\min}$, respectively, R-ℓ_2 is defined as:

$$\text{R-}\ell_2 = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{|y_i - \hat{y}_i|}{y_{\max} - y_{\min}}\right)^2 \tag{2}$$

where y_i and $\hat{y}_i$ indicate the ground-truth and predicted scores for the i-th sample, respectively.

A.3 Baselines

USDL[7] is a direct regression approach that uses a 3-layer MLP as action decoder and a temporal pooling layer followed by a softmax layer as score regressor.

CoRe [10] combines feature of target and exemplar from I3D with exemplar score as pair encoder, then pass to a group-aware regression tree as score regressor.

TSA [8] uses temporal segmentation modules composed by convolution layers and linear block as action decoder, a multi-head transformer block as pair encoder and a 2-layer MLP as score regressor.

TPT [1] adopt learnable queries to decode action sequences into a fixed number of temporal-aware part representations. It uses a 2-layer transformer block as action decoder, a 2-layer MLP as pair encoder and another 2-layer MLP as score regressor.

A.4 Implementation Details

We adopt I3D pretrained on Kinetics [2] as video backbone with 1024 output dimension and ST-GCN [9] pretrained on NTU RGB+D[6] as pose backbone with 256 output dimension. Following [4, 1], we evenly sample 103 frames for each video or pose sequence, split them into 20 clips with length 8 frames and stride 5 frames, then fed them into backbone. The action decoder is a transformer block with 2 layers, 2 query clusters and 512 feature dimension. The pair encoder is a 2-layer MLP with dimension (512, 256), (256, 256). The regressor and extra penalty heads are 3-layer MLP with dimension (256, 256), (256, 64), (64, 1). We set the initial learning rate to 10^{-4} for action decoder and pair encoder, 10^{-3} for regressor and penalty heads. We utilize Adam optimizer [3] and set weight decay to 0. In main experiments, we set the number of exemplar for voting as 2 for LucidAction and 10 for AQA-7, MTL-AQA, FineDiving during inference. The training epoch is 250 for AQA-7 and MLT-AQA, 170 for Finediving and 150 for LucidAction. We use multiple GeForce RTX 4090 GPUs for a total of about 3840 GPU-hours for data pre-process and experiments. We run each experiment three times with different random seed to account for variability and ensure the robustness of the results. We will release the LucidAction dataset and provide the code of baselines soon to promote future research on action quality assessment.

A.5 Ablation Settings

In the multi-view ablation study, we compare the single-view input baseline model with multi-view input models employing batch or channel strategies on LucidAction. The batch strategy places multiple views in the batch dimension, treating them as separate samples, while the channel strategy positions different views on the channel dimension within one sample. We also investigate the effects

of various positions for channel fusion after each module in the baseline architecture and different fusion options (Opt), namely concatenation (*Cat*) and averaging (*Avg*). To ensure a fair comparison, we expand the dimensions of the corresponding modules in base, batch, and channel-average models to match those in channel-concatenate models. For instance, when fuse position is after I3D backbone, the input dimension of action decoder will double from 1024 to 2048, while others remain unchanged.

In the curriculum learning ablation study, model using the mixed training strategy is trained for 150 epochs on the shuffled mixture of training data from up to the test level. In contrast, model using the curriculum learning strategy is trained for 150 epochs on the training data for each level, sequentially from low to high up to the test level. For example, when test level is level 5, the curriculum learning model is trained in the order of levels 3-4-5. By this ablation setting we ensure that both models use an equal amount of training data, thereby guaranteeing the comparability of the results.

A.6 Abaltion on Context Quality and Quantity

A key design concept of contrastive regression is to provide exemplars with known scores as context. In this section, we designed two sets of ablation experiments to examine how the quality and quantity of provided context from LucidAction affect the results of AQA.

We first investigate the impact of the candidates pool of exemplar selection on the effectiveness of AQA. In Table 1a, method "*None*" is the no-context baseline that performs direct regression on target performance score. Other models predict the relative score with different candidate pools of exemplars. method "*All*" randomly selects exemplars from all training samples, while method "*Level*" and method "*Act*" require exemplars to have the same curriculum level and same action category respectively with the target. Due to the penalty-based annotation, LucidAction has at least one perfect training sample without any execution errors for each action category. Therefore, we also compared the "*Act+Perf*" method that requires exemplars to be both error-free and within same action category as the target. In this ablation. As shown in Table 1a, "*Act*" elevates model performance from 0.586 in "*All*" to 0.672, achieving a 14.7% improvement, demonstrating that providing context from the same action category anchors the model's sensibility into a more accessible differential evaluation of similar action sequence, thereby reducing the complexity of model learning. Furthermore, the performance of "*Act+Perf*" improves further to 0.701. Perfect examples provide unambiguous references for the model, maintaining consistency in the mapping from execution differences to relative scores across different target-exemplar pairs, thus avoiding confusion in quality assessment standards caused by errors inherent in exemplars. It shows the superiority of stable but diverse action quality facilitated by the systematic curriculum-based data collection methods in LucidAction compared to collecting one-shot performances in public competitions.

Table 1: Ablation studies regarding the quality and quantity of contrastive regression contexts.

(a) Ablation study of exemplar candidate pool.

Candidates	$\rho \uparrow$	R-$\ell_2 (\times 100) \downarrow$
None.	0.547	0.704
All.	0.586	0.681
Level.	0.614	0.673
Act.	0.672	0.633
Act.+Perf.	**0.701**	**0.624**

(b) Ablation study of voting exemplars number.

#Exemplar	$\rho \uparrow$	R-$\ell_2 (\times 100) \downarrow$
0	0.547	0.704
1	0.681	0.658
2	**0.701**	0.624
3	0.698	**0.620**
5	0.696	0.623

Then, We investigate the model's performance concerning the number of voting exemplars during inference, where the final score is the average of the predicted scores from all target-exemplar pairs. The exemplar candidate pool is "*Act+Perf*" in this ablation study. If the number of perfect training samples in the action category is less than the desired number of exemplars, repeated sampling is performed from the perfect samples. As shown in Table 1b, providing one exemplar results in a 24.5% improvement in model performance compared to direct regression method with no exemplar, increasing from 0.547 to 0.681 in correlation, thus demonstrating the effectiveness of contrastive

(a) Camera layout for men's horizontal bars (MHB).

(b) Camera layout for women's balance beam (WBB).

(c) Camera layout for men's parallel bars (MPB).

(d) Camera layout for women's uneven bar (WUB).

(e) Camera layout for men's vault (MVT).

(f) Camera layout for women's vault (WVT).

Figure 1: Camera layout and corresponding frames for each event.

regression. However, the improvement becomes less pronounced as the number of exemplars reaches two, suggesting that optimal results can be attained in practical inference without a large number of exemplars.

B. The LucidAction Dataset Details

B.1 Multi-View Motion Capture

Figure 1 are the illustration of camera layout and corresponding frames for event vault (MVT, WVT), men's parallel bars (MPB), horizontal bars (MHB), women's uneven bar (WUB) and women's balance beam (WBB).

B.2 Detailed Penalty Items Annotation.

One open question of AQA is whether the model reasoning ability towards concept of action quality can improve when specific execution errors with semantic meaning, rather than a final score, is provided. The annotations of LucidAction contains detailed penalty items and penalty scores for execution errors in actions as extra supervision: All actions share an error set $E = \{e_i\}_{i=1}^N$, each error item has a corresponding penalty set $P_i = \{p_i^1, \cdots, p_i^M\}$ consisting of the admissible penalty

4

score w.r.t error e_i. penalty score $p_i^j \in [0, 1]$ denotes the severity of the error e_i from mild to severe. During training and inference, model predict the penalty score p_i for each error item e_i given a paired target action with a non-error exemplar. The error reasoning paradigm is consistent with both the human referee assessing process and the score annotation method of the proposed LucidAction dataset. For example, in action *Forward Roll*, an error e_i could be *Rolling not smooth* with penalty set $P_i = \{0, 0.1, 0.3, 0.5, 1\}$, 0 indicate the roll is perfectly smooth and 1 as not smooth at all. In action *Side Kick*, error *Angle deviation* with penalty set $P_i = \{0, 0.1, 0.3, 0.5, 1\}$ corresponding to {no deviation, <30°, >30°- 60°, >60°- 90°, >90°}.

B.3   Privacy and Ethical Considerations

During the data collection process, in order to ensure the integrity of the captured actions in the frame, it is inevitable to capture the athletes' faces. To this end, we apply automatic anonymization of faces on all videos in our proposed LucidAction dataset with open-source toolkit deface [1], and a consent form is provided by all the participants agreeing on being recorded and making the captured video publicly available for research purposes.

B.4   Descriptions of Action Categories

Table 2 shows the detailed descriptions of action categories within each curriculum level and sport event.

Table 2: The detailed descriptions of action categories.

Index	Event	Level	Action Code	Description
0	MFE	3	MFE3X00	Forward Roll
1	MFE	3	MFE3X01	Turn Left 180°
2	MFE	3	MFE3X02	Shoulder Elbow Stand
3	MFE	3	MFE3X03	Hollow Body Single Leg Sit-up (Left and Right Leg Each)
4	MFE	3	MFE3X04	Forward Bend from Sitting Position (2")
5	MFE	3	MFE3X05	Backward Roll
6	MFE	3	MFE3X06	Jump Turn 180°
7	MFE	4	MFE4T00	Forward Roll
8	MFE	4	MFE4T01	Swallow Balance
9	MFE	4	MFE4T02	Front Flip
10	MFE	4	MFE4T03	Body Leap
11	MFE	4	MFE4X00	Horizontal Split
12	MFE	4	MFE4X01	Vertical Split (Left Leg in Front)
18	MFE	4	MFE4X02	Vertical Split (Right Leg in Front)
19	MFE	4	MFE4X03	Turn Right 180° to Horizontal Split
20	MFE	4	MFE4X04	Horizontal Split to Prone Position
21	MFE	4	MFE4X05	Prone Support with Arms Sideways to Body Side
22	MFE	4	MFE4X06	Prone Knee Bend to Kneel Stand
23	MFE	4	MFE4X07	Forward Roll
24	MFE	4	MFE4X08	Jump Turn 180°
25	MFE	4	MFE4X09	Single Leg Balance (2")
13	MFE	4	MFE4X10	Backward Roll from Hollow Body Position
14	MFE	4	MFE4X11	Shoulder Elbow Stand (2")
15	MFE	4	MFE4X12	Forward Roll to Stand Up
16	MFE	4	MFE4X13	Side Flip
17	MFE	4	MFE4X14	Turn Left 90° and Stand Up
26	MFE	5	MFE5T00	Forward Roll Split Stand
27	MFE	5	MFE5T01	Forward Roll to Right Angle Sit
28	MFE	5	MFE5T02	Backward Shoulder Elbow Stand
29	MFE	5	MFE5T03	Single Shoulder Backward Roll
30	MFE	5	MFE5T04	Kneeling Balance on One Knee
31	MFE	5	MFE5T05	Kneel to Sit Jump Up
32	MFE	5	MFE5X00	Swallow Balance (2")
33	MFE	5	MFE5X01	Horizontal Split
38	MFE	5	MFE5X02	Vertical Split (Left and Right Each)
39	MFE	5	MFE5X03	Horizontal Split to Prone Position
40	MFE	5	MFE5X04	Prone Straight Leg to Stand Up
41	MFE	5	MFE5X05	Step Forward Handstand
42	MFE	5	MFE5X06	Handstand Forward Roll to Split Support (2")
43	MFE	5	MFE5X07	Straddle Push-up; Draw Legs; Kneel Push-up to Stand
44	MFE	5	MFE5X08	Forward Roll Straight Leg Stand
45	MFE	5	MFE5X09	Headstand (2")
34	MFE	5	MFE5X10	Forward Bend to Push-up

[1]https://github.com/ORB-HD/deface

35	MFE	5	MFE5X11	Push-up Turn Right 180° to Right Angle Sit
36	MFE	5	MFE5X12	Backward Roll Straight Leg Stand
37	MFE	5	MFE5X13	Step Side Flip (2 Times)
46	MFE	6	MFE6T00	Side Flip with 90° Turn
47	MFE	6	MFE6T01	Forward Roll Split Stand
48	MFE	6	MFE6T02	Split Stand Slow Start Handstand
49	MFE	6	MFE6T03	Forward Roll to Right Angle Sit
50	MFE	6	MFE6T04	Backward Roll Straight Leg Stand
51	MFE	6	MFE6X00	Step Forward Handstand
52	MFE	6	MFE6X01	Single Leg Swing to Handstand
56	MFE	6	MFE6X02	Handstand Forward Roll to Split Support
57	MFE	6	MFE6X03	Split Support Slow Start to Handstand
58	MFE	6	MFE6X04	Left Leg Forward Kick Right Leg Exchange Jump
59	MFE	6	MFE6X05	Step Side Flip (Consecutive 2 Times)
60	MFE	6	MFE6X06	Forward Bend to Single Leg Lift to Prone Position
61	MFE	6	MFE6X07	Straddle Push-up; Turn Right 180° to Right Angle Sit
62	MFE	6	MFE6X08	Forward Bend Backward Roll to Handstand
63	MFE	6	MFE6X09	Handstand Single Leg Bend Downward Jump Turn 45° to Stand
53	MFE	6	MFE6X10	Side Flip with 90° Inside Turn
54	MFE	6	MFE6X11	Vertical Jump and Land
55	MFE	6	MFE6X12	Backward Flip Jump and Land to Stand
64	MHB	3	MHB3X00	Overhand Dead Hang Swing
65	MHB	4	MHB4X00	Forward 4 Steps and Lift Hips
66	MHB	4	MHB4X01	Forward 4 Steps and Turn 180°
67	MHB	4	MHB4X02	Walk forward; swing back to standing position
68	MHB	4	MHB4X03	Kick feet to ground
69	MHB	4	MHB4X04	Backward swing dismount
70	MHB	5	MHB5T00	Jump up to support
71	MHB	5	MHB5T01	Support backward swing
72	MHB	5	MHB5T02	Single leg swing to straddle support
73	MHB	5	MHB5T03	Back leg swings to the side
74	MHB	5	MHB5X00	Tuck and flip up to support
75	MHB	5	MHB5X01	Swing backward to shoulder level
76	MHB	5	MHB5X02	Slow descent from front flip to dead hang
77	MHB	5	MHB5X03	Hang swing; back swing; swinging in dead hang
78	MHB	5	MHB5X04	Forward and backward swing to dismount
79	MHB	6	MHB6T00	Flip up to support
80	MHB	6	MHB6T01	Single leg swing to straddle support
81	MHB	6	MHB6T02	Backward swing single leg hang
82	MHB	6	MHB6T03	Turn 180° to support
83	MHB	6	MHB6T04	Backward swing dismount
84	MHB	6	MHB6X00	Overhand dead hang swing; lift hips after swinging
85	MHB	6	MHB6X01	Forward swing; bend and straighten legs to lift up
86	MHB	6	MHB6X02	Backward swing to support; swing back to dead hang
87	MHB	6	MHB6X03	Arc swing from backward to forward swing
88	MHB	6	MHB6X04	Swinging in dead hang
89	MHB	6	MHB6X05	Backward swing dismount
90	MPB	3	MPB3X00	Jump up to support
91	MPB	4	MPB4T00	Jump up to support
92	MPB	4	MPB4T01	Sequentially move support forward
93	MPB	4	MPB4T02	Tucked L-Sit on Parallel Bars
94	MPB	4	MPB4T03	Jump off bars
95	MPB	4	MPB4X00	Jump up to support
96	MPB	4	MPB4X01	Small swing below bar level
97	MPB	4	MPB4X02	Forward swing to one bar side sit
98	MPB	4	MPB4X03	Turn 90° out; jump off in straddle position
99	MPB	5	MPB5T00	Jump up to straddle sit
100	MPB	5	MPB5T01	Straddle sit forward to straddle sit
101	MPB	5	MPB5T02	Straddle sit forward
102	MPB	5	MPB5T03	Forward swing to side sit
103	MPB	5	MPB5T04	Forward swing to right angle between two bars
104	MPB	5	MPB5X00	Swinging in dead hang
105	MPB	5	MPB5X01	Forward swing; bend and straighten legs to lift up to straddle support
106	MPB	5	MPB5X02	Slowly rise to shoulder level with legs split
107	MPB	5	MPB5X03	Swinging in support; back and forth
108	MPB	5	MPB5X04	Forward and backward swing; jump off in single arm support (60°)
109	MPB	6	MPB6T00	Jump up to right angle support (2")
110	MPB	6	MPB6T01	Forward lean to straddle sit
111	MPB	6	MPB6T02	Straddle sit forward
112	MPB	6	MPB6T03	Backward swing 180° to straddle sit
113	MPB	6	MPB6T04	Spring bars forward swing
114	MPB	6	MPB6T05	Backward and forward swing to right (left) single bar support (60°) dismount
115	MPB	6	MPB6X00	Inside grip dead hang; lift legs forward in arc swing to backward swing
116	MPB	6	MPB6X01	Swinging in dead hang
117	MPB	6	MPB6X02	Forward swing
118	MPB	6	MPB6X03	Slowly rise to shoulder level with legs split; support with legs together
119	MPB	6	MPB6X04	Swinging in support; back and forth
120	MPB	6	MPB6X05	Forward and backward swing to right (left) side; jump off in single bar support (60°)

121	MVT	3	MVT3X00	Run-up onto board and jump onto high mat with straight body
122	MVT	4	MVT4T00	Goat Split Leap
123	MVT	4	MVT4X00	Run-up onto board and jump onto high mat with tucked body
124	MVT	5	MVT5T00	Goat Bent Leg Leap
125	MVT	5	MVT5X00	Jump onto forward roll
126	MVT	6	MVT6T00	Horizontal Box Split Jump
127	MVT	6	MVT6X00	Run-up onto board and jump into handstand on mat
128	WBB	3	WBB3X00	Jump onto beam with feet together and stand upright
129	WBB	3	WBB3X01	Take one step forward
130	WBB	3	WBB3X02	Rise up on toes once
131	WBB	3	WBB3X03	Straight body jump
132	WBB	3	WBB3X04	Step forward to beam and stand with feet together
133	WBB	3	WBB3X05	Straight body jump down
134	WBB	4	WBB4T00	Swing back to squat support and stand upright
135	WBB	4	WBB4T01	Take 4 steps forward
136	WBB	4	WBB4T02	Turn 180° with feet together
137	WBB	4	WBB4T03	Take 4 steps backward
138	WBB	4	WBB4T04	Vertical jump
139	WBB	4	WBB4T05	Turn 90° with feet together
140	WBB	4	WBB4T06	Jump off from horizontal beam
141	WBB	4	WBB4X00	Swing back to squat
142	WBB	4	WBB4X01	Take one step forward
143	WBB	4	WBB4X02	Turn 180° with feet together
144	WBB	4	WBB4X03	Take one step backward
145	WBB	4	WBB4X04	Straight body jump
146	WBB	4	WBB4X05	Straight body jump down
147	WBB	5	WBB5T00	Jump onto beam sideways; turn 180° into straddle support
148	WBB	5	WBB5T01	Kneeling balance (2")
149	WBB	5	WBB5T02	Take one step forward into single-leg stand
150	WBB	5	WBB5T03	Single-leg jump and land with both heels raised and stand upright
151	WBB	5	WBB5T04	Take one step forward; squat and turn 180° into upright position
152	WBB	5	WBB5T05	Leg kicks during movement (front; side; back; left and right leg each once)
153	WBB	5	WBB5T06	Turn 90° with feet together
154	WBB	5	WBB5T07	Jump off from horizontal beam with split legs
155	WBB	5	WBB5X00	Jump onto beam and hold straddle support (hold for 2")
156	WBB	5	WBB5X01	Leg kicks during movement (front; right leg)
158	WBB	5	WBB5X02	Leg kicks during movement (front; left leg)
159	WBB	5	WBB5X03	Leg kicks during movement (side; right leg)
160	WBB	5	WBB5X04	Leg kicks during movement (side; left leg)
161	WBB	5	WBB5X05	Leg kicks during movement (back; right leg)
162	WBB	5	WBB5X06	Leg kicks during movement (back; left leg)
163	WBB	5	WBB5X07	Straight body jump; switch legs mid-air and jump
164	WBB	5	WBB5X08	Turn 180° with feet together
165	WBB	5	WBB5X09	Swallow balance (2")
157	WBB	5	WBB5X10	Jump off from beam with feet together and straight body forward
166	WBB	6	WBB6X00	Jump onto beam and hold straddle support (2")
167	WBB	6	WBB6X01	Leg kicks during movement (front; right leg)
174	WBB	6	WBB6X02	Leg kicks during movement (front; left leg)
175	WBB	6	WBB6X03	Leg kicks during movement (side; right leg)
176	WBB	6	WBB6X04	Leg kicks during movement (side; left leg)
177	WBB	6	WBB6X05	Leg kicks during movement (back; right leg)
178	WBB	6	WBB6X06	Leg kicks during movement (back; left leg)
179	WBB	6	WBB6X07	Vertical split jump
180	WBB	6	WBB6X08	Straight body jump; switch legs mid-air and jump
181	WBB	6	WBB6X09	Wolf jump
168	WBB	6	WBB6X10	Single-leg turn 180°
169	WBB	6	WBB6X11	Forward somersault
170	WBB	6	WBB6X12	Sideways somersault; turn 90° into lunge position
171	WBB	6	WBB6X13	Leg extension balance (2")
172	WBB	6	WBB6X14	Body wave
173	WBB	6	WBB6X15	Jump off from beam with feet together and straight body backward
182	WFE	3	WFE3X00	Roll Forward in Tuck Position
183	WFE	3	WFE3X01	Backward Roll in Tuck Position
186	WFE	3	WFE3X02	Shoulder Elbow Stand (2")
187	WFE	3	WFE3X03	Straddle Jump
188	WFE	3	WFE3X04	Front Split (Left Leg in Front)
189	WFE	3	WFE3X05	Front Split (Right Leg in Front)
190	WFE	3	WFE3X06	Front Kick (Left Leg) during movement
191	WFE	3	WFE3X07	Sideways Kick (Right Leg) during movement
192	WFE	3	WFE3X08	Back Kick (Left Leg) during movement
193	WFE	3	WFE3X09	Front Kick (Right Leg) during movement
184	WFE	3	WFE3X10	Sideways Kick (Left Leg) during movement
185	WFE	3	WFE3X11	Back Kick (Right Leg) during movement
194	WFE	4	WFE4T00	Forward Roll in Tuck Position
195	WFE	4	WFE4T01	Swallow Balance
196	WFE	4	WFE4T02	Leap with Extended Support
197	WFE	4	WFE4T03	Straddle Jump
198	WFE	4	WFE4X00	Forward Roll in Pike Position to Straight Leg Stand
199	WFE	4	WFE4X01	Forward Roll

200	WFE	4	WFE4X02	Backward Roll
201	WFE	4	WFE4X03	Side Aerial
202	WFE	4	WFE4X04	180° Straight Jump Turn
203	WFE	4	WFE4X05	Swallow Balance (2")
204	WFE	4	WFE4X06	Body Wave
205	WFE	5	WFE5T00	Front Roll with Split Leg Stand
206	WFE	5	WFE5T01	Front Roll to Sit in Right Angle
207	WFE	5	WFE5T02	Back Shoulder Elbow Stand
208	WFE	5	WFE5T03	One Arm Back Roll
209	WFE	5	WFE5T04	One Knee to Kneel Balance
210	WFE	5	WFE5T05	Kneeling Push-Up to Sit to Kneel Jump
211	WFE	5	WFE5X00	Forward Roll to Sit in Right Angle
212	WFE	5	WFE5X01	Backward Roll
213	WFE	5	WFE5X02	Side Aerial
214	WFE	5	WFE5X03	Back Layout
215	WFE	5	WFE5X04	Front Split
216	WFE	5	WFE5X05	Side Split
217	WFE	5	WFE5X06	Straddle Jump
218	WFE	5	WFE5X07	Western Jump
219	WFE	5	WFE5X08	360° Straight Jump Turn
220	WFE	6	WFE6T00	Sideways Aerial with 90° Turn
221	WFE	6	WFE6T01	Front Roll with Split Leg Stand
222	WFE	6	WFE6T02	Front Roll to Stand
223	WFE	6	WFE6T03	Swallow Balance
224	WFE	6	WFE6T04	Back Leg Extended into Front Split
225	WFE	6	WFE6T05	Front Leg Extended into Right Angle Sit
226	WFE	6	WFE6T06	Backward Roll to Stand
227	WFE	6	WFE6X00	Three-Step Run-Up into Double Leg Front Handspring
228	WFE	6	WFE6X01	Three-Step Run-Up into Take-Off Step Jointed with Straight Jump
229	WFE	6	WFE6X02	Front Layout into Back Layout
230	WFE	6	WFE6X03	360° Turn on One Foot
231	WFE	6	WFE6X04	Parallel Feet Jump
232	WFE	6	WFE6X05	Long Stride Jump
233	WFE	6	WFE6X06	Wolf Jump
234	WFE	6	WFE6X07	Split Jump with Straight Body
235	WFE	6	WFE6X08	Backward Roll to Handstand
236	WUB	3	WUB3X00	Tuck In 90° during Roll
237	WUB	3	WUB3X01	Change Hands Left to Right
238	WUB	3	WUB3X02	Overhand Grip Hang (10")
239	WUB	3	WUB3X03	Drop Down to Stand
240	WUB	4	WUB4X00	Jump Up to Support (3")
241	WUB	4	WUB4X01	Consecutive Support and Back Swing (Below Horizontal) 2 times
242	WUB	4	WUB4X02	Forward Roll to Tuck Hang (2")
243	WUB	5	WUB5X00	Flick Flack to Handstand
244	WUB	5	WUB5X01	Back Swing after Support
245	WUB	5	WUB5X02	Abdominal Roll
246	WUB	5	WUB5X03	Back Swing to Straight Body Jump Down
247	WUB	6	WUB6X00	Long Swing Bend Up
248	WUB	6	WUB6X01	Back Swing after Support (Above Horizontal)
249	WUB	6	WUB6X02	Abdominal Roll to Support
250	WUB	6	WUB6X03	Legs Together Straight Body Swing to Handstand
251	WUB	6	WUB6X04	Pike Support to Back Swing Bend Down
252	WVT	3	WVT3X00	Run-Up on Springboard to Jump on High Mat with Straight Body
253	WVT	4	WVT4T00	Goat Split Leg Vault
254	WVT	4	WVT4X00	Run-Up on Springboard to Jump on High Mat in Tuck Position
255	WVT	5	WVT5T00	Goat Bend Leg Vault
256	WVT	5	WVT5X00	Run-Up on Springboard to Jump on High Mat with Forward Roll
257	WVT	6	WVT6T00	Horizontal box Vault
258	WVT	6	WVT6X00	TRun-Up on Springboard to Jump on High Mat with Handstand

139

8

References

[1] Yang Bai, Desen Zhou, Songyang Zhang, Jian Wang, Errui Ding, Yu Guan, Yang Long, and Jingdong Wang. Action Quality Assessment with Temporal Parsing Transformer. In *Computer Vision – ECCV 2022*, volume 13664, pages 422–438. Springer Nature Switzerland, 2022. doi: 10.1007/978-3-031-19772-7_25.

[2] Joao Carreira and Andrew Zisserman. Quo Vadis, Action Recognition? A New Model and the Kinetics Dataset. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4724–4733. IEEE, 2017. doi: 10.1109/CVPR.2017.502.

[3] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *CoRR*, abs/1412.6980, 2014. URL https://api.semanticscholar.org/CorpusID: 6628106.

[4] Paritosh Parmar and Brendan Morris. Action Quality Assessment Across Multiple Actions. In *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 1468–1476, 2019. doi: 10.1109/WACV.2019.00161.

[5] Paritosh Parmar and Brendan Tran Morris. What and How Well You Performed? A Multitask Learning Approach to Action Quality Assessment. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 304–313, 2019.

[6] Amir Shahroudy, Jun Liu, Tian-Tsong Ng, and Gang Wang. Ntu rgb+d: A large scale dataset for 3d human activity analysis. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1010–1019, 2016.

[7] Yansong Tang, Zanlin Ni, Jiahuan Zhou, Danyang Zhang, Jiwen Lu, Ying Wu, and Jie Zhou. Uncertainty-Aware Score Distribution Learning for Action Quality Assessment. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 9836–9845. IEEE, 2020. doi: 10.1109/CVPR42600.2020.00986.

[8] Jinglin Xu, Yongming Rao, Xumin Yu, Guangyi Chen, Jie Zhou, and Jiwen Lu. FineDiving: A Fine-grained Dataset for Procedure-aware Action Quality Assessment. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2939–2948. IEEE, 2022. doi: 10.1109/CVPR52688.2022.00296.

[9] Sijie Yan, Yuanjun Xiong, and Dahua Lin. Spatial temporal graph convolutional networks for skeleton-based action recognition. In *AAAI*, 2018.

[10] Xumin Yu, Yongming Rao, Wenliang Zhao, Jiwen Lu, and Jie Zhou. Group-aware Contrastive Regression for Action Quality Assessment. In *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 7899–7908. IEEE, 2021. doi: 10.1109/ICCV48922.2021.00782.