

ECAI23 Supplementary Materials

1 Additional Analyses on the Proposed Framework

1.1 Additional Case Study

As shown in Figure 1, we included additional examples of masked video clips in in-distribution (ID) and out-of-distribution (OOD) for a better understanding of our attention-based masking framework. Again, we can see that our frame selection strategy selected the frames with low motion variances, e.g., *push up* action in Figure 1 (a) and *casting fishing line* in Figure 1 (b). In addition, patch masking reasonably separates the area corresponding to the foreground from the frame and erases the background and the other patches less related to the action.

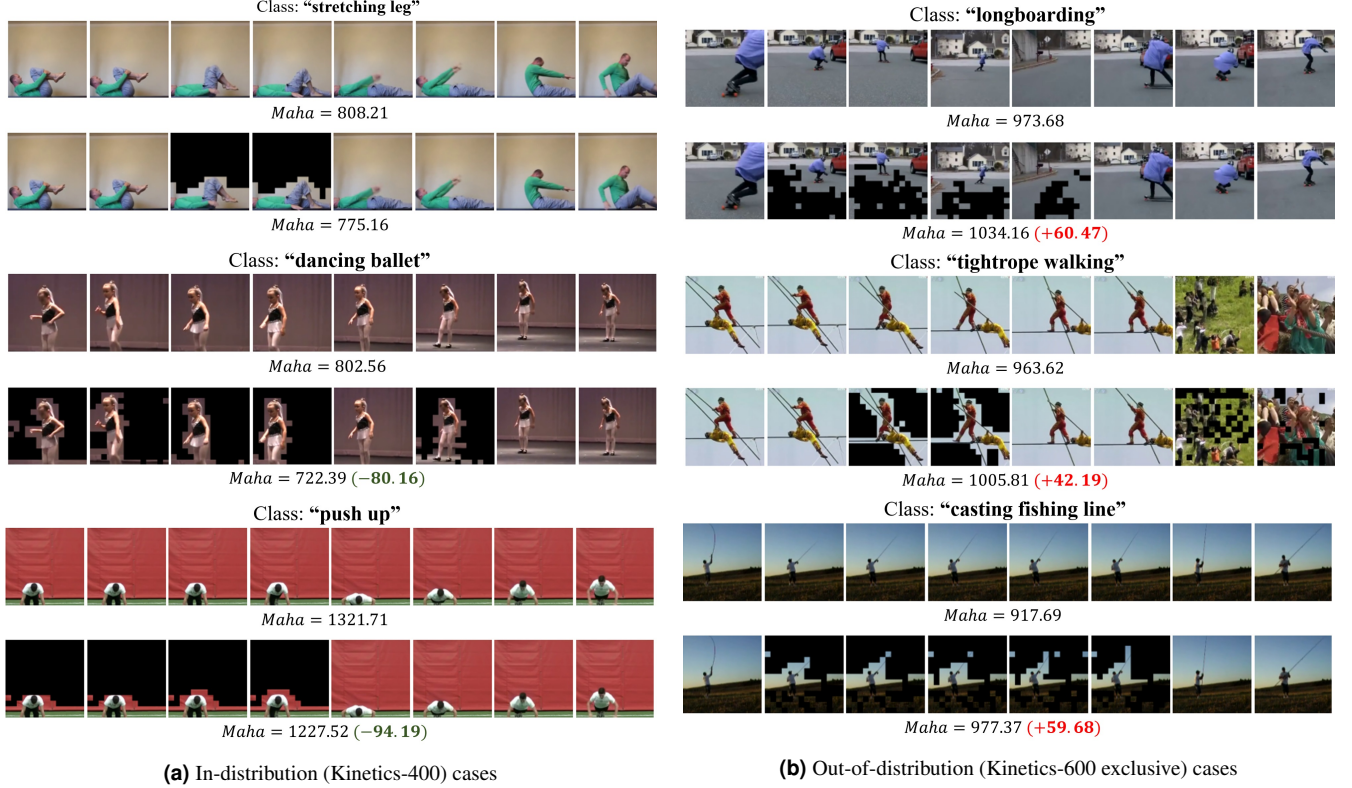
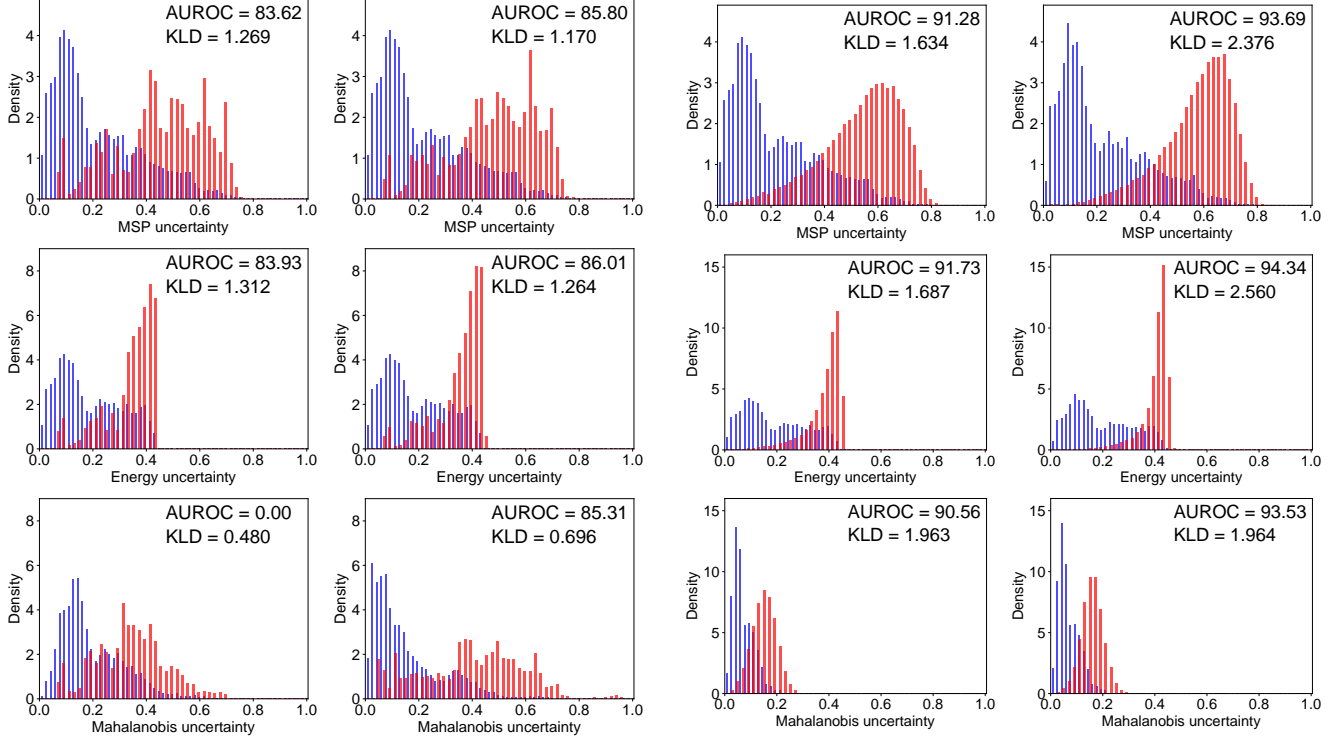


Figure 1. More examples of masked video clips. We include additional examples of ID (Kinetics-400) and OOD (Kinetics-600 exclusive) with the differences in video clips and Mahalanobis distance before and after applying our framework. $lt_threshold$ parameters γ_l and γ_s are set to 0.01 and 0.05, respectively. Again, the value in parentheses means the difference, **red** and **green** indicate negative and positive differences, respectively.

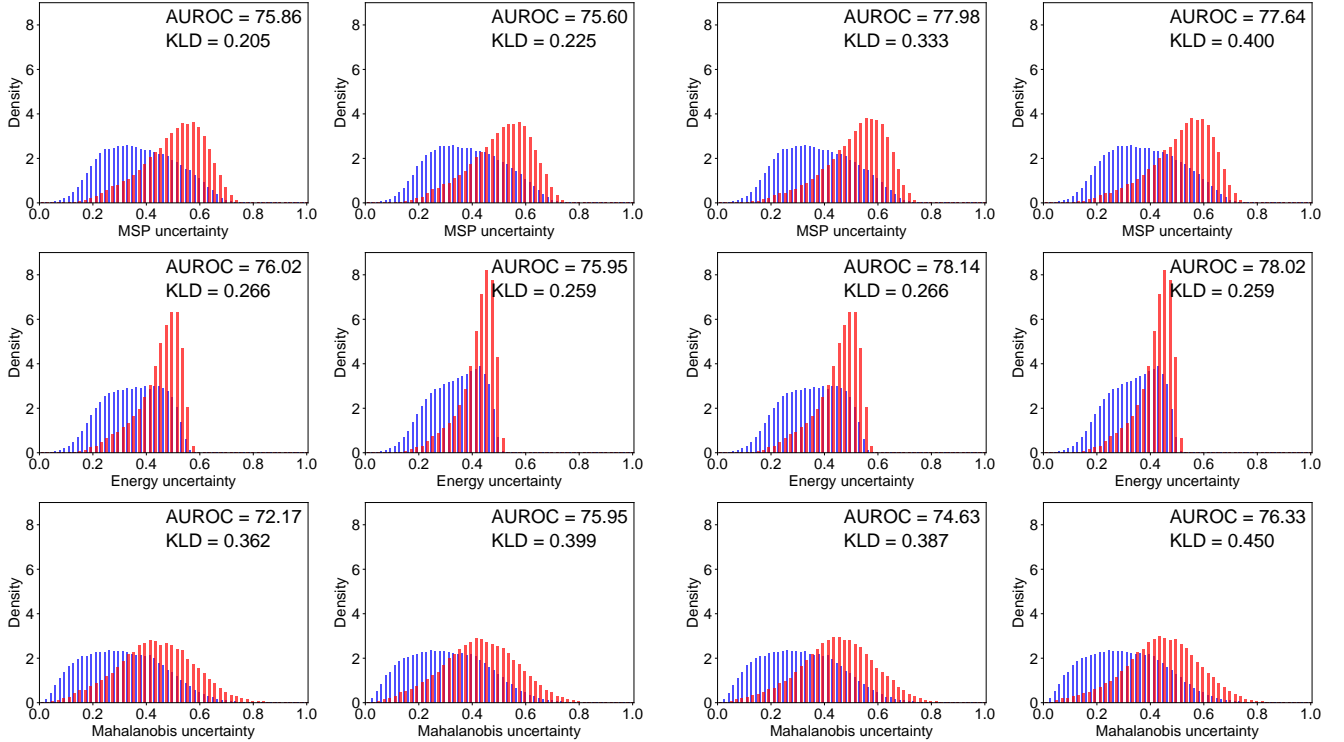
1.2 Histogram Statistics

Our framework for out-of-distribution (OOD) action detection presents a simple yet powerful solution to address the static bias problem in human action recognition (HAR). To evaluate the effectiveness of our framework, we conducted experiments on four OOD action detection tasks: UCF-101 vs. HMDB-51, UCF-101 vs. MiT-v2, Kinetics-400 vs. Kinetics-600 exclusive, and Kinetics-400 vs. Kinetics-700 exclusive, as shown in Figure 2. Before and after applying our framework, we examined the histogram statistics of both in-distribution (ID) and OOD samples. The results showed significant differences in the histogram distributions between the two categories. However, after applying our framework, we observed a notable increase in the discrepancy of histogram statistics, indicated by the increased AUROC and KL-divergence scores. These findings underscore the effectiveness of our proposed framework in accurately identifying OOD data, which holds promising implications for enhancing the robustness and reliability of AI models in various action detection tasks.



(a) UCF-101 (in) vs. HMDB-51 (out) histograms

(b) UCF-101 (in) vs. MiT-v2 (out) histograms



(c) Kinetics-400 (in) vs. Kinetics-600 exclusive (out) histograms

(d) Kinetics-400 (in) vs. Kinetics-700 exclusive (out) histograms

Figure 2. Histogram Statistics. We analyze histogram statistics before and after applying our framework to four OOD action detection tasks. Note that for each OOD action detection task, figures in the left and right columns denote histogram statistics of ID (blue) and OOD (red) samples before and after applying our framework. Here, we observe a significant increase in the discrepancy between ID and OOD data. This highlights the efficacy of our approach in accurately detecting OOD samples.

1.3 Failure Cases

In this section, we provide illustrative examples of classes that fall within the groups where our method was found to be ineffective. In these cases, the videos primarily rely on static biased cues, such as water in the "surfing water" class and snow in the "snowboarding" class. Interestingly, we observed that applying attention-based masking had a negative impact on the OOD detection performance. This led to an increase in the Mahalanobis distance, indicating a potential vulnerability of our method in scenarios where static cues play a significant role. These findings highlight the need for further exploration and refinement of our framework to address such specific challenges.

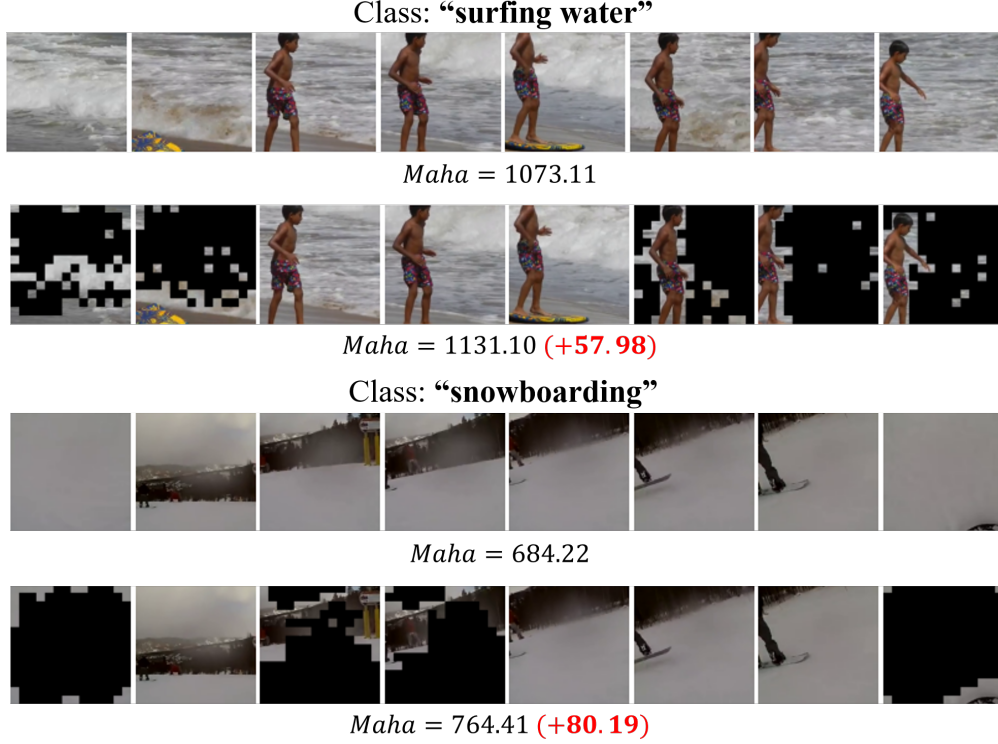


Figure 3. In-distribution (Kinetics-400) cases

Figure 4. Examples on worst groups. We present examples of classes included in the groups where our method was ineffective. In most cases, the videos are highly dependent on the static biased cues (e.g., water in the *surfing water* and snow in the *snowboarding* class). Applying attention-based masking adversely affects OOD detection, which increases Mahalanobis distance in this case.

2 Experimental Details

2.1 Details on the Groupwise Analysis

For the top/bottom five groups for which our framework was effective, we listed the Kinetics-400 classes included in each action group in Table 1. Note that this taxonomy is obtained from the original Kinetics-400 paper. We can see that the top five groups mainly consist of actions less dependent on the static biased cues (e.g., *stretching leg* in *body motions* action group and *golf driving* in *golf* group). In this case, removing static biased cues with attention-based masking helped the model to focus on the salient parts of the action and gave benefits in terms of OOD detection performance. However, the bottom five groups usually consist of actions highly related to the static bias. For example, as shown in Figure 4, *swimming* and *snow ice* groups are dependent on the background (e.g., water and snow) and *athletics – throwing+launching* group is dependent on the static object appearing in the scene (e.g., hammer and discus).

Action Groups	Classes
golf	golf putting (930) golf driving (684) golf chipping (545)
makeup	filling eyebrows (892) dying hair (885) doing nails (758) getting a tattoo (584) applying cream (318)
body motions	air drumming (977) headbanging (935) pumping fist (849) shaking head (726) finger snapping (674) stretching leg (660) stretching arm (557) headbutting (479) bending back (472) baby waking up (451) clapping (334) cracking neck (290) drumming fingers (257) applauding (256) swinging legs (255)
gymnastics	gymnastics tumbling (987) somersaulting (825) bouncing on trampoline (529) cartwheeling (455) vault (410)
athletics – jumping	pole vault (820) high jump (798) long jump (678) triple jump (628) hurdling (467) parkour (348)
paper	reading book (971) bookbinding (758) folding paper (718) opening present (696) unboxing (694) folding napkins (693) wrapping present (672) writing (564) counting money (503) ripping paper (447) reading newspaper (266) shredding paper (249)
swimming	swimming backstroke (912) swimming breaststroke (680) swimming butterfly stroke (521)
snow ice	snowkiting (990) tobogganing (989) ice skating (971) skiing (not slalom or crosscountry) (971) biking through snow (895) ski jumping (874) snowboarding (787) playing ice hockey (750) shoveling snow (719) ice climbing (691) sled dog racing (623) making snowman (594) snowmobiling (453) bobsledding (400) ice fishing (377) skiing slalom (352) hockey stop (318) skiing rosscountry (318)
cloths	folding napkins (693) tying knot (not on a tie) (675) folding clothes (523) making bed (522) bandaging (413) ironing (383) doing laundry (307) tying bow tie (217)
athletics – throwing+launching	hammer throw (992) archery (978) throwing discus (943) catching or throwing frisbee (891) shot put (824) javelin throw (745) throwing axe (662) throwing ball (479) disc golfing (415)

Table 1. Classes in each selected action group. We list the classes of each selected action group, along with the number of training videos. The classes were sorted in descending order based on the number of training videos it has, and a maximum of 5 classes were included per group.

2.2 Details on Kinetics-Exclusive Datasets

We followed the mainstream OOD detection tasks that focused on the semantic shift of the dataset and removed the classes with the same labels of Kinetics-400 from Kinetics-600 and Kinetics-700. The list of classes in Kinetics-600 exclusive and Kinetics-700 exclusive are shown in Table 2 and Table 3. We referred to the list of newly added classes mentioned in the Kinetics-600 and Kinetics-700 paper, and the resulting number of classes in Kinetics-600 exclusive and Kinetics-700 exclusive are 213 and 309, respectively.

1	acting in play	55	embroidering	109	pillow fight	163	skipping stone
2	adjusting glasses	56	falling off bike	110	pinching	164	sleeping
3	alligator wrestling	57	falling off chair	111	pirouetting	165	smashing
4	archaeological excavation	58	fencing (sport)	112	planing wood	166	smelling feet
5	arguing	59	fidgeting	113	playing beer pong	167	smoking pipe
6	assembling bicycle	60	fixing bicycle	114	playing darts	168	spelunking
7	attending conference	61	flint knapping	115	playing dominoes	169	square dancing
8	backflip (human)	62	fly tying	116	playing field hockey	170	standing on hands
9	base jumping	63	geocaching	117	playing gong	171	staring
10	bathing dog	64	getting a piercing	118	playing hand clapping games	172	steer roping
11	battle rope training	65	gold panning	119	playing laser tag	173	sucking lolly
12	blowdrying hair	66	gospel singing in church	120	playing lute	174	swimming front crawl
13	blowing bubble gum	67	hand washing clothes	121	playing maracas	175	swinging baseball bat
14	bodysurfing	68	head stand	122	playing marbles	176	sword swallowing
15	bottling	69	historical reenactment	123	playing netball	177	tackling
16	bouncing on bouncy castle	70	home roasting coffee	124	playing ocarina	178	tagging graffiti
17	breaking boards	71	huddling	125	playing pan pipes	179	talking on cell phone
18	breathing fire	72	ice swimming	126	playing pinball	180	tasting wine
19	building lego	73	installing carpet	127	playing ping pong	181	threading needle
20	building sandcastle	74	ironing hair	128	playing polo	182	throwing ball
21	bull fighting	75	jaywalking	129	playing rubiks cube	183	throwing knife
22	bulldozing	76	jumping bicycle	130	playing scrabble	184	throwing snowballs
23	burping	77	jumping jacks	131	playing with trains	185	throwing tantrum
24	calculating	78	karaoke	132	poking bellybutton	186	throwing water balloon
25	calligraphy	79	land sailing	133	polishing metal	187	tie dying
26	capsizing	80	lawn mower racing	134	popping balloons	188	tightrope walking
27	card stacking	81	laying concrete	135	pouring beer	189	tiptoeing
28	card throwing	82	laying stone	136	pushing wheelbarrow	190	trimming shrubs
29	carving ice	83	laying tiles	137	putting in contact lenses	191	twiddling fingers
30	casting fishing line	84	leatherworking	138	putting on eyeliner	192	tying necktie
31	changing gear in car	85	licking	139	putting on foundation	193	tying shoe laces
32	chewing gum	86	lifting hat	140	putting on lipstick	194	using a microscope
33	chiseling stone	87	lighting fire	141	putting on mascara	195	using a paint roller
34	chiseling wood	88	lock picking	142	putting on sari	196	using a power drill
35	chopping meat	89	longboarding	143	putting on shoes	197	using a sledge hammer
36	chopping vegetables	90	looking at phone	144	raising eyebrows	198	using a wrench
37	coloring in	91	luge	145	repairing puncture	199	using atm
38	combing hair	92	making balloon shapes	146	riding snow blower	200	using bagging machine
39	contorting	93	making bubbles	147	roasting marshmallows	201	using circular saw
40	cooking scallops	94	making cheese	148	roasting pig	202	using inhaler
41	cosplaying	95	making horseshoes	149	rolling pastry	203	using puppets
42	cracking back	96	making paper aeroplanes	150	rope pushdown	204	visiting the zoo
43	cracking knuckles	97	marriage proposal	151	sausage making	205	wading through mud
44	crossing eyes	98	massaging neck	152	sawing wood	206	wading through water
45	cumbia	99	moon walking	153	scrapbooking	207	walking through snow
46	curling (sport)	100	mosh pit dancing	154	scrubbing face	208	watching tv
47	cutting apple	101	mountain climber (exercise)	155	separating eggs	209	waving hand
48	cutting orange	102	mushroom foraging	156	sewing	210	weaving fabric
49	delivering mail	103	needle felting	157	shaping bread dough	211	winking
50	directing traffic	104	opening door	158	shining flashlight	212	wood burning (art)
51	docking boat	105	opening refrigerator	159	shopping	213	yarn spinning
52	doing jigsaw puzzle	106	packing	160	shucking oysters		
53	drooling	107	photobombing	161	shuffling feet		
54	dumpster diving	108	photocopying	162	sipping cup		

Table 2. Classes in Kinetics-600 exclusive

1	acting in play	79	fencing (sport)	157	pinching	235	shucking oysters
2	adjusting glasses	80	fidgeting	158	pirouetting	236	shuffling feet
3	alligator wrestling	81	filling cake	159	planing wood	237	sieving
4	archaeological excavation	82	fixing bicycle	160	playing american football	238	silent disco
5	arguing	83	flint knapping	161	playing beer pong	239	sipping cup
6	arresting	84	flipping bottle	162	playing billiards	240	ski ballet
7	assembling bicycle	85	fly tying	163	playing blackjack	241	skipping stone
8	attending conference	86	geocaching	164	playing checkers	242	sleeping
9	backflip (human)	87	getting a piercing	165	playing darts	243	slicing onion
10	bathing dog	88	gold panning	166	playing dominoes	244	smashing
11	battle rope training	89	gospel singing in church	167	playing field hockey	245	smelling feet
12	being excited	90	grooming cat	168	playing gong	246	smoking pipe
13	being in zero gravity	91	hand washing clothes	169	playing hand clapping games	247	spelunking
14	blending fruit	92	head stand	170	playing laser tag	248	spinning plates
15	blowdrying hair	93	helmet diving	171	playing lute	249	splashing water
16	blowing bubble gum	94	herding cattle	172	playing mahjong	250	square dancing
17	bodysurfing	95	high fiving	173	playing maracas	251	squeezing orange
18	bottling	96	historical reenactment	174	playing marbles	252	stacking cups
19	bouncing ball (not juggling)	97	home roasting coffee	175	playing netball	253	stacking dice
20	bouncing on bouncy castle	98	huddling	176	playing nose flute	254	standing on hands
21	breaking boards	99	hugging (not baby)	177	playing oboe	255	staring
22	breaking glass	100	hugging baby	178	playing ocarina	256	steer roping
23	breathing fire	101	ice swimming	179	playing pan pipes	257	steering car
24	brushing floor	102	installing carpet	180	playing piccolo	258	sucking lolly
25	building lego	103	ironing hair	181	playing pinball	259	surveying
26	building sandcastle	104	jaywalking	182	playing ping pong	260	swimming front crawl
27	bulldozing	105	jumping bicycle	183	playing polo	261	swimming with dolphins
28	burping	106	jumping jacks	184	playing road hockey	262	swimming with sharks
29	calculating	107	jumping sofa	185	playing rounders	263	swinging baseball bat
30	calligraphy	108	karaoke	186	playing rubiks cube	264	sword swallowing
31	capsizing	109	land sailing	187	playing scrabble	265	tackling
32	card stacking	110	lawn mower racing	188	playing shuffleboard	266	tagging graffiti
33	card throwing	111	laying concrete	189	playing slot machine	267	taking photo
34	carrying weight	112	laying decking	190	playing with trains	268	talking on cell phone
35	carving ice	113	laying stone	191	poaching eggs	269	tasting wine
36	carving marble	114	laying tiles	192	poking bellybutton	270	threading needle
37	carving wood with a knife	115	leatherworking	193	polishing furniture	271	throwing ball
38	casting fishing line	116	letting go of balloon	194	polishing metal	272	throwing knife
39	changing gear in car	117	licking	195	popping balloons	273	throwing snowballs
40	chasing	118	lifting hat	196	pouring beer	274	throwing tantrum
41	checking watch	119	lighting candle	197	pouring milk	275	throwing water balloon
42	chewing gum	120	lighting fire	198	pouring wine	276	tie dying
43	chiseling stone	121	listening with headphones	199	preparing salad	277	tightrope walking
44	chiseling wood	122	lock picking	200	pretending to be a statue	278	tiptoeing
45	chopping meat	123	longboarding	201	pulling espresso shot	279	treating wood
46	clam digging	124	looking at phone	202	pulling rope (game)	280	trimming shrubs
47	closing door	125	looking in mirror	203	pushing wheelbarrow	281	twiddling fingers
48	coloring in	126	luge	204	putting in contact lenses	282	tying shoe laces
49	combing hair	127	making balloon shapes	205	putting on eyeliner	283	uncorking champagne
50	contorting	128	making bubbles	206	putting on foundation	284	using a microscope
51	cooking scallops	129	making cheese	207	putting on lipstick	285	using a paint roller
52	cosplaying	130	making horseshoes	208	putting on mascara	286	using a power drill
53	coughing	131	making latte art	209	putting on sari	287	using a sledge hammer
54	cracking back	132	making paper aeroplanes	210	putting on shoes	288	using a wrench
55	cracking knuckles	133	making slime	211	putting wallpaper on wall	289	using atm
56	crocheting	134	marriage proposal	212	raising eyebrows	290	using bagging machine
57	crossing eyes	135	massaging neck	213	repairing puncture	291	using circular saw
58	cumbia	136	metal detecting	214	riding snow blower	292	using inhaler
59	curling (sport)	137	milking goat	215	roasting marshmallows	293	using megaphone
60	curling eyelashes	138	mixing colours	216	roasting pig	294	using puppets
61	cutting apple	139	moon walking	217	rolling eyes	295	vacuuming car
62	cutting cake	140	mosh pit dancing	218	rolling pastry	296	visiting the zoo
63	cutting orange	141	mountain climber (exercise)	219	rope pushdown	297	wading through mud
64	dealing cards	142	moving baby	220	saluting	298	wading through water
65	decoupage	143	moving child	221	sanding wood	299	waking up
66	delivering mail	144	mushroom foraging	222	sausage making	300	walking on stilts
67	directing traffic	145	needle felting	223	sawing wood	301	walking through snow
68	docking boat	146	opening coconuts	224	scrapbooking	302	walking with crutches
69	doing jigsaw puzzle	147	opening door	225	scrubbing face	303	watching tv
70	doing sudoku	148	opening refrigerator	226	seasoning food	304	waving hand
71	drooling	149	opening wine bottle	227	separating eggs	305	waxing armpits
72	dumpster diving	150	packing	228	sewing	306	weaving fabric
73	dyeing eyebrows	151	passing soccer ball	229	shaping bread dough	307	winking
74	eating nachos	152	peeling banana	230	shining flashlight	308	wood burning (art)
75	embroidering	153	petting horse	231	shoot dance	309	yarn spinning
76	entering church	154	photobombing	232	shooting off fireworks		
77	falling off bike	155	photocopying	233	shopping		
78	falling off chair	156	pillow fight	234	shouting		

Table 3. Classes in Kinetics-700 exclusive