

Supplemental Material: Diagnosing Errors in Video Relation Detectors

Shuo Chen
 Pascal Mettes
 Cees G.M. Snoek
 {s.chen3,p.s.m.mettes,cgmsnoek}@uva.nl

VIS Lab,
 University of Amsterdam,
 The Netherlands

Super Category	Category	Classes	Instances
Subjects & Objects			
Animal	turtle, antelope, lion, cattle, red_panda, horse, monkey, fox, elephant, bird, sheep, giant_panda, squirrel, bear, tiger, snake, rabbit, whale, dog, domestic_cat, lizard, hamster, zebra	23	39097 (63.57%)
Furniture	sofa	1	356 (0.58%)
Person	person	1	8536 (13.88%)
Sports	ball, frisbee, skateboard	3	519 (0.84%)
Vehicle	bicycle, motorcycle, airplane, watercraft, bus, train, car	7	12996 (21.13%)
Predicates			
Action	bite, chase, drive	14	2956 (9.61%)

Spatial	walk_above, stand_behind, next_to, fly_toward, sit_right, jump_with, walk_behind, creep_above, stand_front, run_front, run_left, jump_next_to, right, creep_right, walk_left, fly_left, swim_beneath, swim_behind, creep_left, creep_away, creep_next_to, lie_left, creep_behind, walk_right, stand_inside, stand_left, jump_above, move_past, run_past, walk_toward, left, creep_toward, jump_toward, walk_next_to, sit_inside, stand_right, run_next_to, lie_behind, fly_right, lie_beneath, sit_left, past, run_away, stop_above, move_with, move_right, lie_above, stop_with, jump_left, stop_right, front, jump_beneath, walk_past, sit_behind, move_above, lie_next_to, walk_beneath, walk_with, move_beneath, run_above, run_with, toward, run_beneath, stop_behind, jump_behind, move_left, walk_front, move_toward, move_behind, above, move_away, swim_left, stand_with, stop_left, stand_beneath, be- neath, stand_next_to, swim_front, creep_beneath, lie_front, move_front, fly_above, sit_beneath, sit_front, jump_away, stop_beneath, sit_above, run_behind, fly_front, creep_front, faster, stop_next_to, away, lie_with, run_toward, lie_right, lie_inside, stop_front, run_right, taller, stand_above, swim_with, jump_past, fly_away, creep_past, walk_away, behind, move_next_to, jump_front, swim_next_to, jump_right, swim_right, sit_next_to, fly_with, larger, fly_behind, fly_next_to, fly_past	118	27796 (90.39%)
---------	---	-----	----------------

Table 1: Subject/object and relation categories in Imagenet-VidVRD dataset, organized by super categories. Note the bias towards animals and spatial relations.

Super Category	Category	Classes	Instances
Subjects & Objects			
Accessory	handbag, backpack, suitcase	3	5948 (1.00%)
Animal	leopard, snake, chicken, hamster/rat, stingray, antelope, turtle, panda, tiger, sheep/goat, crocodile, pig, fish, cat, dog, lion, bird, elephant, duck, camel, kangaroo, crab, cattle/cow, penguin, horse, squirrel, bear, rabbit	28	51009 (8.58%)
Appliance	refrigerator, sink, oven, microwave, electric_fan	5	2034 (0.34%)
Electronic	camera, cellphone, screen/monitor, laptop	4	16005 (2.69%)
Food	fruits, bread, cake, vegetables, dish	5	8094 (1.36%)
Furniture	table, toilet, stool, chair, sofa	5	41089 (6.91%)
Indoor	toy	1	30034 (5.05%)
Kitchen	bottle, cup	2	21330 (3.59%)
Other	piano, baby_walker, baby_seat, faucet, guitar	5	15789 (2.65%)
Outdoor	bench, stop_sign, traffic_light	3	1744 (0.29%)
Person	adult, baby, child	3	368549 (61.97%)
Sports	bat, snowboard, surfboard, skateboard, racket, ball/sports_ball, ski, frisbee	8	16697 (2.81%)
Vehicle	scooter, watercraft, bus/truck, bicycle, aircraft, car, motorcycle, train	8	16382 (2.75%)
Predicates			
Action	kiss, bite, push, point_to, wave_hand_to, drive, carry, open, watch, throw, clean, feed, wave, shake_hand_with, play(instrument), get_off, hug, touch, hold, pat, press, chase, close, release, grab, lift, smell, hold_hand_of, knock, lick, cut, kick, pull, get_on, lean_on, hit, speak_to, ride, shout_at, squeeze, caress, use	42	69066 (23.23%)
Spatial	above, next_to, beneath, in_front_of, away, behind, inside, towards	8	228286 (76.77%)

Table 2: Subject/object and relation categories in VidOR, organized by their super categories. Note the bias towards persons and spatial relations.