

## Appendix: COCO Fine-grained Categories

Anonymous CVPR submission

Paper ID 207

The COCO [1] dataset does not have bounding box annotations associated with specific phrases or entities in the caption. We use category level detection annotations and create a category mapping list that maps the object categories like <Person> to a list of potential fine-grained labels like [“child”, “man”, “baker”,...]. We first use the Stanford lemmatization toolbox [2] to get the base form of the entity words in the caption. For each category class, we retrieve the top 200 similar words in the WordVec [3] space. We then manually verify each word in the list, resulting in 413 fine-grained classes. A complete list of the fine-grained class for each object category can be found in Table 1 and Table 2.

### References

- [1] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick. Microsoft coco: Common objects in context. In *ECCV*, 2014. 1
- [2] C. D. Manning, M. Surdeanu, J. Bauer, J. Finkel, S. J. Bethard, and D. McClosky. The Stanford CoreNLP natural language processing toolkit. In *ACL*, 2014. 1
- [3] T. Mikolov, K. Chen, G. Corrado, and J. Dean. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*, 2013. 1

Object category	Fine-grained class
<person>	person, girl, boy, man, woman, kid, child, chef, baker, people, adult, rider, children, baby, worker, passenger, sister, biker, policeman, officer, lady, cowboy, bride, groom, male, female, guy, traveler, mother, father, gentleman, pitcher, player, skier, snowboarder, skater, skateboarder, foreigner, caller, offender, coworker, trespasser, patient, politician, soldier, grandchild, serviceman, walker, drinker, doctor, bicyclist, thief, buyer, teenager, student, camper, driver, soldier, hunter, shopper, villager, cop
<bicycle>	bicycle, bike, unicycle, minibike, trike
<car>	car, automobile, van, minivan, sedan, suv, hatchback, cab, jeep, coupe, taxicab, limo, taxi
<motorcycle>	motorcycle, scooter, motor bike, motor cycle, motorbike, moped
<airplane>	airplane, jetliner, plane, air plane, monoplane, aircraft, jet, airbus, biplane, seaplane bus, minibus, trolley
<bus>	bus, minibus, schoolbus, trolley
<train>	train, locomotive, tramway, caboose
<truck>	truck, pickup, lorry, hauler, firetruck
<boat>	boat, ship, liner, sailboat, motorboat, dinghy, powerboat, speedboat, canoe, skiff, yacht, kayak, catamaran, pontoon, houseboat, vessel, rowboat, trawler, ferryboat, watercraft, tugboat, schooner, barge, ferry, sailboard, paddleboat, lifeboat, freighter, steamboat, riverboat, surfboard, battleship, steamship
<traffic light>	traffic light, street light, traffic signal, stop light, streetlight, stoplight
<fire hydrant>	fire hydrant, hydrant
<stop sign>	stop sign, street sign
<parking meter>	parking meter
<bench>	bench, pew
<cat>	cat, kitten, feline, tabby
<dog>	dog, puppy, beagle, pup, chihuahua, schnauzer, dachshund, rottweiler, canine, pitbull, collie, pug, terrier, poodle, labrador, doggie, doberman, mutt, doggy, spaniel, bulldog, sheepdog, weimaraner, corgi, cocker, greyhound, retriever, brindle, hound, whippet, husky
<horse>	horse, colt, pony, racehorse, stallion, equine, mare, foal, palomino, mustang, clydesdale, bronc, bronco
<sheep>	sheep, lamb, goat, ram, cattle, ewe
<cow>	cow, cattle, oxen, ox, calf, ewe, holstein, heifer, buffalo, bull, zebu, bison
<elephant>	elephant
<bear>	bear, panda
<zebra>	zebra
<giraffe>	giraffe
<backpack>	backpack, knapsack
<umbrella>	umbrella
<handbag>	handbag, handbag, wallet, purse, briefcase
<tie>	tie
<suitcase>	suitcase, suit case, luggage
<frisbee>	frisbee
<skis>	skis, ski
<snowboard>	snowboard
<sports ball>	sports ball, baseball, ball, football, soccer, basketball, softball, volleyball, pinball, fastball, racquetball
<kite>	kite
<baseball bat>	baseball bat
<baseball glove>	baseball glove
<skateboard>	skateboard
<surfboard>	surfboard, longboard, skimboard, shortboard, wakeboard
<tennis racket>	tennis racket
<bottle>	bottle

Table 1. COCO category mapping list for visual words.

216			270
217			271
218			272
219			273
220			274
221			275
222			276
223			277
224			278
225	<b>Object category</b>	<b>Fine-grained class</b>	279
226			280
227	<wine glass>	wine glass	281
228	<cup>	cup	282
229	<fork>	fork	283
230	<knife>	knife, pocketknife, knife	284
231	<spoon>	spoon	285
232	<banana>	banana	286
233	<apple>	apple	287
234	<sandwich>	sandwich, burger, sub, cheeseburger, hamburger	288
235	<orange>	orange, lemons	289
236	<broccoli>	broccoli	290
237	<carrot>	carrot	291
238	<hot dog>	hot dog	292
239	<pizza>	pizza	293
240	<donut>	donut, doughnut, bagel	294
241	<cake>	cake, cheesecake, cupcake, shortcake, coffeecake, pancake	295
242	<bird>	bird, ostrich, owl, seagull, goose, duck, parakeet, falcon, robin, pelican, waterfowl, heron, hummingbird, mallard, finch, pigeon, sparrow, seabird, osprey, blackbird, fowl, shorebird, woodpecker, egret, chickadee, quail, bluebird, kingfisher, buzzard, willet, gull, swan, bluejay, flamingo, cormorant, parrot, loon, gosling, waterbird, pheasant, rooster, sandpiper, crow, raven, turkey, oriole, cowbird, warbler, magpie, peacock, cockatiel, lorikeet, puffin, vulture, condor, macaw, peafowl, cockatoo, songbird	296
243	<chair>	chair, seat, recliner, stool	297
244	<couch>	couch, sofa, recliner, futon, loveseat, settee, chesterfield	298
245	<potted plant>	potted plant, houseplant	299
246	<bed>	bed	300
247	<dining table>	dining table, table	301
248	<toilet>	toilet, urinal, commode, lavatory, potty	302
249	<tv>	tv, monitor, televison, television	303
250	<laptop>	laptop, computer, notebook, netbook, lenovo, macbook	304
251	<mouse>	mouse	305
252	<remote>	remote	306
253	<keyboard>	keyboard	307
254	<cell phone>	cell phone, mobile phone, phone, cellphone, telephone, phon, smartphone, iPhone	308
255	<sink>	sink	309
256	<refrigerator>	refrigerator, fridge, refrigerator, fridge, freezer, refridgerator, frig	310
257	<book>	book	311
258	<clock>	clock	312
259	<vase>	vase	313
260	<scissors>	scissors	314
261	<teddy bear>	teddy bear, teddybear	315
262	<hair drier>	hair drier, hairdryer	316
263	<toothbrush>	toothbrush	317
264			318
265			319
266			320
267			321
268			322
269			323

Table 2. COCO category mapping list for visual words (continued).