Supplemental Material:
From Large Scale Image Categorization to Entry-Level Categories

Vicente Ordonez[1], Jia Deng[2], Yejin Choi[3], Alexander C. Berg[1], Tamara L. Berg[1]
[1]University of North Carolina at Chapel Hill, [2]Stanford University, [3]Stony Brook University
[vicente,aberg,tlberg]@cs.unc.edu, jdeng@stanford.edu, ychoi@cs.stonybrook.edu

1. Translation Mappings

Figure 1 extends Figure 3 in the main paper. It shows more examples of concept mappings using our Language-only translation method described in section 3.1 and our Visually-informed Translation method described in section 3.2.

	Input Concept	Ngram-translation	SVM-translation	Human - translation
1	eastern kingbird	bird	bird	bird
2	cactus wren	bird	bird	bird
3	buzzard, Buteo buteo	hawk	bird	hawk
4	whinchat, Saxicola rubetra	chat	bird	bird
6	Weimaraner	dog	dog	dog
7	Gordon setter	dog	dog	dog
8	numbat, banded anteater, anteater	anteater	cat	anteater
9	rhea, Rhea americana	bird	grass	ostrich
10	Africanized bee, killer bee, Apis mellifera	bee	flower	bee
11	conger, conger eel	eel	water	fish
12	merino, merino sheep	sheep	dog	sheep
13	Europ. black grouse, heathfowl, Lyrurus tetrix	bird	duck	bird
14	yellowbelly marmot, rockchuck, Marm. flaviventris	marmot	rock	squirrel
15	snorkeling, snorkel diving	swimming	water	snorkel
16	American crow, Corvus brachyrhyncos	crow	bird	bird
17	common nutcracker, Nucifraga caryocatactes	bird	bird	bird
18	giant salamander, Megalobatrachus maximus	salamander	rock	lizard
19	carrier pigeon	homer	bird	bird
20	rhinoceros beetle	beetle	bird	bug
21	bottom, freighter, merchantman, merchant ship	bottom	ship	ship
22	bulletproof vest	protection	shirt	vest
23	chain wrench	tool	bead	chain
24	chateau	home	castle	castle
25	polonaise	dress	dress	dress
26	bicorn, bicorne	hat	dress	hat
27	jeroboam, double-magnum	bottle	bottle	wine
28	shoe shop, shoe-shop, shoe store	store	market	shoe
29	field speedwell, Veronica agrestis	flower	flower	flower
30	tobacco mildew, Peronospora hyoscyami	mildew	flower	leaf

Figure 1. Translations from ImageNet leaf node synset categories to entry level categories using our automatic approaches from the main paper sections 3.1 (left) and 3.2 (center) and crowd-sourced human annotations from section 2 (right).

2. Supervised Learning of Mappings

Figures 2, 3 extend Figure 5 in our paper. They show more examples of mappings between the fine grained level categories of [1] and base level categories.

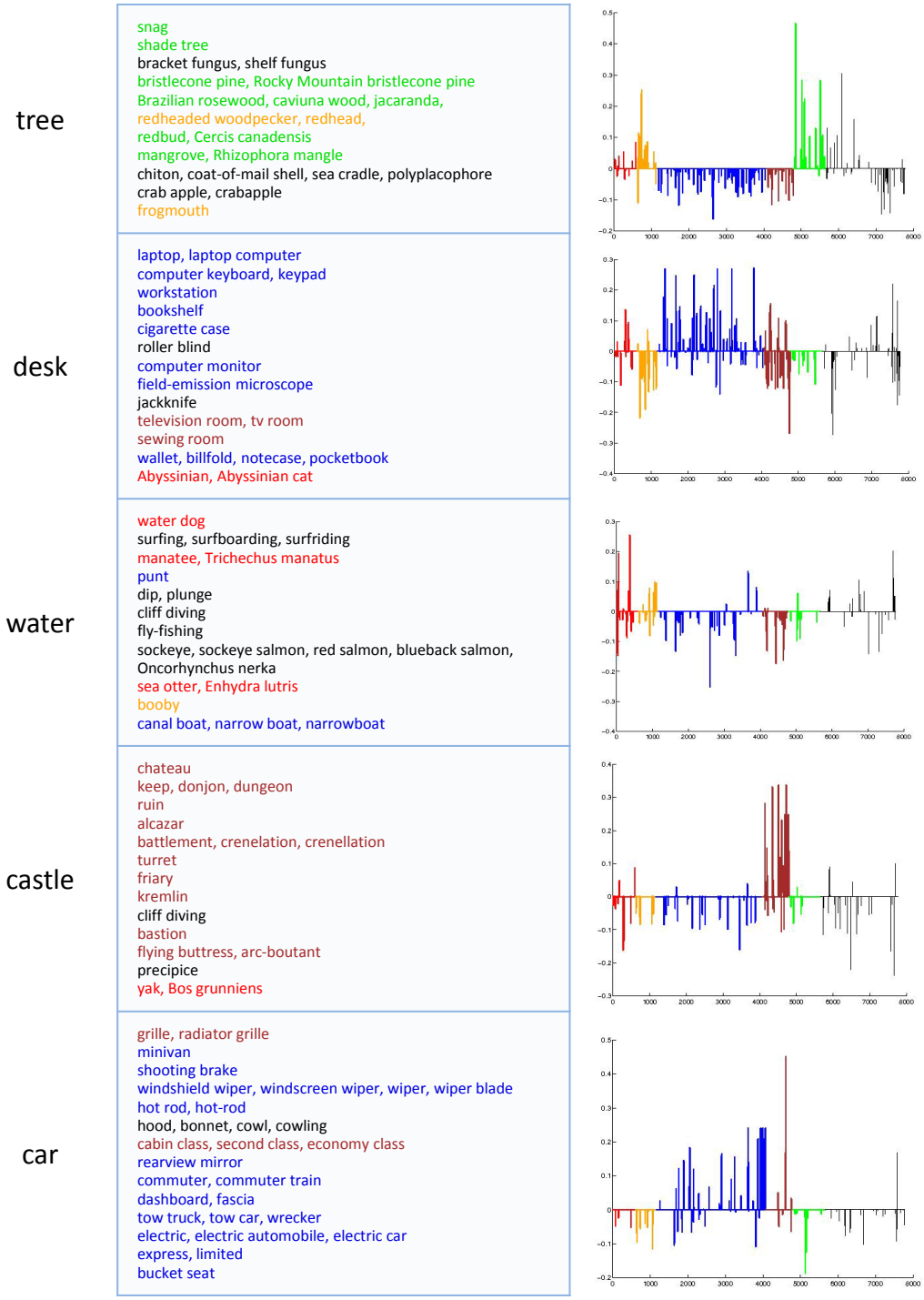


Figure 2. Entry-level categories with their corresponding top weighted leaf node features after training an SVM on our noisy data and a visualization of weights grouped by an arbitrary categorization of leaf nodes. vegetation(green), birds(orange), instruments(blue), structures(brown), mammals(red), others(black).

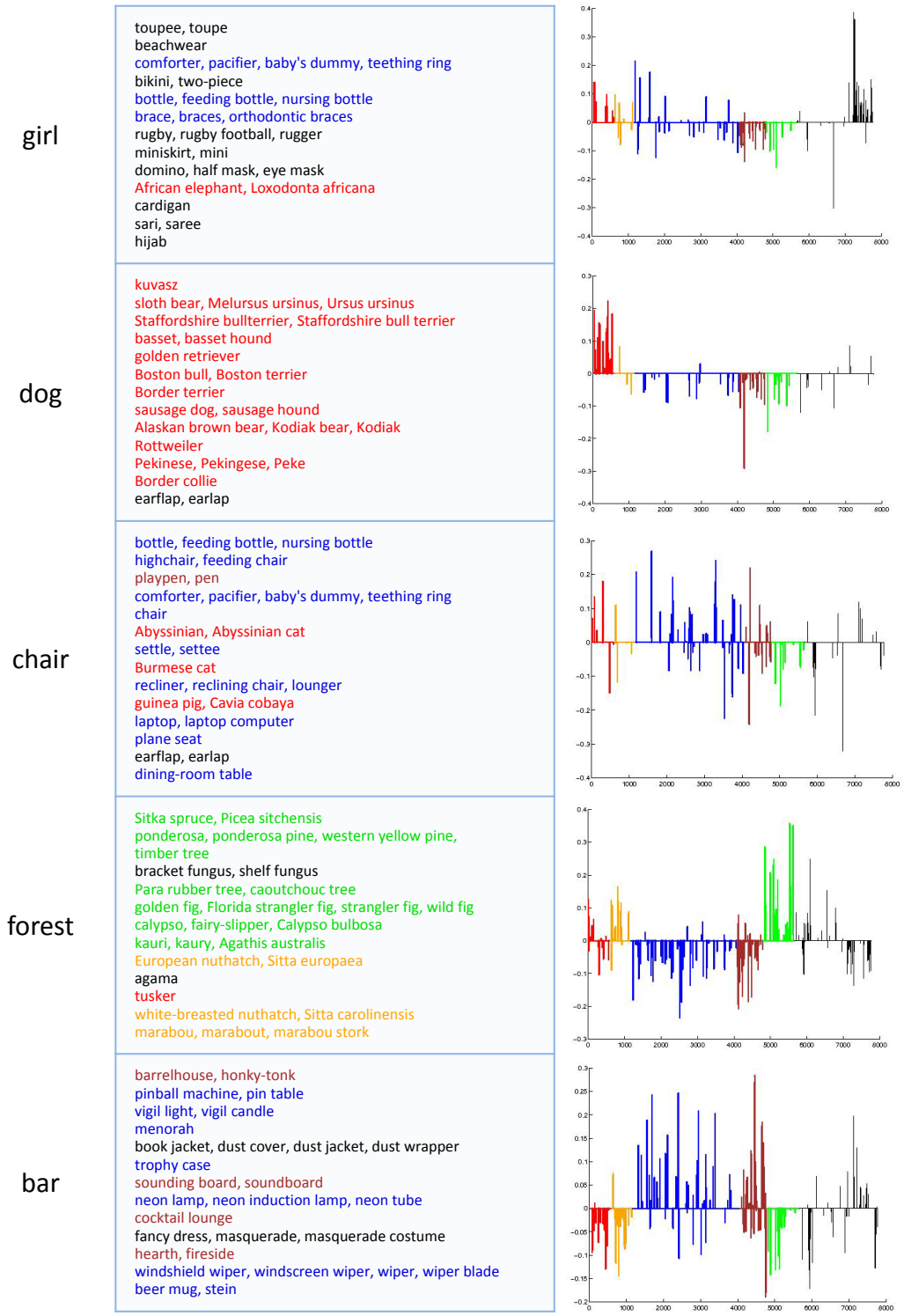


Figure 3. (Continuation of Figure 5.) Entry-level categories with their corresponding top weighted leaf node features after training an SVM on our noisy data and a visualization of weights grouped by an arbitrary categorization of leaf nodes. vegetation(green), birds(orange), instruments(blue), structures(brown), mammals(red), others(black).

3. Entry-Level Category Prediction Results

We show more qualitative results for predicting entry-level categories. Figures 4 and 5 show additional results for **Dataset A** and Figures 6 and 7 show additional results for **Dataset B**. All these figures extend Figure 8 in the main paper.

	Images	Labels	Flat Classifier	Hedging [6]	Ngram-based	SVM-based	Joint
Results in the top 25%		building bush, field fountain grass, home house, window manor, sky tree, yard white house	farmhouse stately ranch courthouse **manor**	**house** **home** **building** housing residence	**building** **home** **house** structure housing	neighborhood street **tree** **house** bridge	**building** **house** **home** structure **tree**
		bush driveway field, flower grass road, rock street, tree	umbrella flamboyant titus grape gleditsium	woody **tree** plant vascular flowering	**tree** plant oak structure framework	**grass** **field** **road** mountain forest	**tree** plant **grass** **field** **road**
		creek, day, water lake, nature landscape, sky mountain, park outside, reflection river, rock	catchment riverside caldera parrotfish wing	formation catchment depression side bank	formation tree structure catchment side	**river** **lake** **mountain** **water** sand	formation tree **river** **lake** **water**
		blue dress bush, dress girl, child grass, plant sky, tree	Hyla large wind Honduras Salix	woody **tree** **plant** vascular conifer	**tree** **plant** material flower wear	**dress** **girl** field beach boy	**dress** **girl** field **tree** beach
		front yard grass, window house, lawn potted plant sidewalk stair, tree	camper stoop chicken dacha detach	camper trailer stoop porch structure	structure trailer porch stoop camper	neighborhood **house** **window** bedroom door	neighborhood **house** building **window** bedroom
		duck duckling fin, fowl goose gosling lake, outdoor pond, water	Canada whistle **gosling** large Hyla	**goose** Canada aquatic anseriform waterfowl	**goose** bird tree Canada waterfowl	beach **water** grass sand field	beach **water** **duck** grass sand
Results in the bottom 25%		airport bus depot state tile tourist train station wall	box stilt balk zip webbing	structure box stilt office material	structure material tree device cover	reflection glass bathroom door floor	reflection building glass bathroom door
		basket, broom child, man dustpan food garbage rake, stick	pant king rubber electrical macrame	cover implement pant leg king	tree cover material device good	bridge river neighborhood road car	bridge river neighborhood road cross

Figure 4. Example translations on **Dataset A**.1^{st} col shows images. 2^{nd} col shows MTurk associated nouns. These represent the ground truth annotations (entry-level categories) we would like to predict (colored in blue). 3^{rd} col shows predicted nouns using a standard multiclass flat-classifier. 4^{th} col shows nouns predicted by the method of [2]. 5^{th} col shows our n-gram based method predictions. 6^{th} col shows our SVM mapping predictions and finally the 7^{th} column shows the labels predicted by our joint model. Matches are colored in green. Tables 1, 2 in the main paper show the measured improvements in recall and precision.

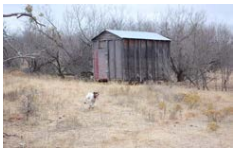
	Images	Labels	Flat Classifier	Hedging [6]	Ngram-based	SVM-based	Joint
Results in the top 25%		barn, brown building, cabin dirt, dog farm, field grass, meadow shack, shed tree, turkey, animal	privy coop gnu snowbank dacha	privy outbuilding **building** structure ungulate	**building** structure **tree** privy outbuilding	fence wooden **barn** **tree** door	**tree** **building** structure house **barn**
		beach, boat building cloud, daylight dock house, house boat ripple, river sky, tower, water	shipping catamaran defense Hyla ketch	vessel transport craft shipping vehicle	vessel tree vehicle shipping craft	**sky** **cloud** field **beach** view	**sky** **cloud** **boat** field **beach**
		flower foliage forest petal plant, shrub	large Hyla doe wind Honduras	woody tree **plant** vascular flowering	tree **flower** **plant** bird oak	**forest** tree garden **plant** pine tree	tree **forest** **plant** **flower** garden
		building city, coast grass, island ocean, park structure tower, tree, water	oil ski acropoli condo castle	**structure** **building** place establishment woody	**structure** **building** **tree** home hotel	neighborhood view **building** apartment hill	**building** neighborhood **structure** **tree** view
		child, dress eye, nose, girl glass, hair lady, woman lip, lip eyebrow mirror reflection, wall	papillon tape seat large turtle	papillon toy toy dog domestic	tree dog cover toy wear	**girl** **dress** face **mirror** bathroom	**dress** **girl** face **mirror** white
		boat, hill lake, oar, paint ripple river sand, sea ship, water	trawler race marina lifeboat cruiser	vessel craft transport vehicle **boat**	vessel vehicle **boat** craft tree	**boat** fishing boat beach **water** floor	**boat** fishing beach **water** view
Results in the bottom 25%		baseball bicycle, bike book pole prospects road	large crossbow lawn Hyla cannon	instrument arm cover device large	device cover equipment wear good	street car box sign dog	street box sign dog mirror
		forest house, hut lady, porch raise-floor stair, tree, tribe	log fixer-upper rest hip woodsh	building structure home housing **house**	structure building **tree** home **house**	bridge fence office building boat bike	building bridge fence office boat

Figure 5. Example translations on **Dataset A**.1^{st} col shows images. 2^{nd} col shows MTurk associated nouns. These represent the ground truth annotations (entry-level categories) we would like to predict (colored in blue). 3^{rd} col shows predicted nouns using a standard multiclass flat-classifier. 4^{th} col shows nouns predicted by the method of [2]. 5^{th} col shows our n-gram based method predictions. 6^{th} col shows our SVM mapping predictions and finally the 7^{th} column shows the labels predicted by our joint model. Matches are colored in green. Tables 1, 2 in the main paper show the measured improvements in recall and precision.

	Images	Labels	Flat Classifier	Hedging [6]	Ngram-based	SVM-based	Joint
Results in the top 25%		daisy flower, grass leaf, petal plant sky, soil stem sunflower	helianthus cosmos dry goshawk large	**sunflower** helianthus seed flowering **flower**	**flower** tree **sunflower** **plant** bird	**sunflower** **flower** **daisy** **sky** color	**flower** **sunflower** **daisy** **sky** **plant**
		building sky station subway train wire	customhouse free hip city ministry	**building** structure customhouse government free	structure **building** home tree government	sign **station** bus **train** market	**building** sign **station** bus **train**
		beak, bird feather, ripple lake, neck pond, pool swan, water	hooded cygnet whooper drake bottlenose	aquatic anseriform waterfowl **swan** duck	**bird** duck **swan** tree material	duck **water** **lake** beach sand	duck **swan** **water** **lake** boat
		bucket, cabinet counter, dish, floor, kitchen microwave oven, range rug, shelf, sink stove, tank, teapot trash, trashcan	**microwave** console gas freezer electric	**microwave** **kitchen** home durable appliance	**microwave** **kitchen** appliance good home	**counter** **stove** **sink** **cabinet** room	**microwave** room console **counter** **stove**
		building, car, city light, light post office, cone, sign pavement, road sidewalk, window structure, uptown van, vehicle, street	limited Hyla Segway wagon lumber	transport wheel **vehicle** **structure** container	**structure** tree equipment **vehicle** container	**street** tent bus **sign** office building	**street** tent bus **building** **sign**
		farm, fence field horse, mule kite, dirt people tree, zoo	gelding yearling shire yearling draft	**horse** equine perissodactyl ungulate male	**horse** **tree** equine male gelding	**horse** pasture **field** cow **fence**	**horse** pasture **field** cow **fence**
Results in the bottom 25%		fence, junk sign stop sign street sign trash can tree	feeder Hyla cleaner box large	woody **tree** structure plant vascular	**tree** structure building plant area	logo street neighborhood building office building	logo street neighborhood building office
		circle earring hook jewel jewelry make up stone	clasp fob enamel chain gold	clasp fix constraint device chain	clasp fix constraint device chain	bead pearl bracelet silver sterling	clasp fix constraint device bead

Figure 6. Example translations on **Dataset B**. 1^{st} col shows images. 2^{nd} col shows MTurk associated nouns. These represent the ground truth annotations (entry-level categories) we would like to predict (colored in blue). 3^{rd} col shows predicted nouns using a standard multiclass flat-classifier. 4^{th} col shows nouns predicted by the method of [2]. 5^{th} col shows our n-gram based method predictions. 6^{th} col shows our SVM mapping predictions and finally the 7^{th} column shows the labels predicted by our joint model. Matches are colored in green. Tables 1, 2 in the main paper show the measured improvements in recall and precision.

	Images	Labels	Flat Classifier	Hedging [6]	Ngram-based	SVM-based	Joint
Results in the top 25%		conifer, grass horse, animal lawn, livestock mammal pasture, roan tree, white	yearling rhodesian **lawn** **roan** vizslum	**horse** equine perissodactyl ungulate placental	**horse** dog equine **tree** mount	**horse** **pasture** field cow dog	**horse** **pasture** **tree** field cow
		cloud hawaius palm, palm tree sky, sun, sunset tree, leaf	date backlighting caryota key Hyla	**palm** woody **tree** plant vascular	**tree** **palm** plant oak equipment	**sunset** **palm tree** sunflower **sky** **sun**	**palm** **sunset** sunflower **sky** **sun**
		animal beak, bird, duck feather lake, water mallard, wildlife	**mallard** drake quack-quack wild aythya	**mallard** **duck** anseriform waterfowl drake	**duck** **mallard** waterfowl drake **bird**	**duck** **water** sand **lake** boat	**duck** **mallard** waterfowl drake **bird**
		boat, ship bridge, vacation building, city father, fishing harbor, pole water, cloud sky, skyline	dredger shipping trawler bascule cantilever	dredger vessel lighter craft transport	vessel lighter vehicle **boat** craft	neighborhood **ship** **bridge** **boat** river	**boat** **ship** neighborhood **bridge** river
		building church door historic, bell house, minaret pretty, tower	belfry **church** clock **minaret** large	belfry **tower** **church** structure room	structure **tower** area room belfry	neighborhood clock tower **door** **tower** **church**	clock **building** neighborhood **door** **church**
		army truck army vehicle car, jeep detachable trailer drive, highway road, spare tire	**jeep** garbage personnel half 4wd	**jeep** self-propelled wheel motor **car**	**car** **jeep** motor vehicle container	logo truck **car** bus market	**car** **jeep** motor vehicle container
Results in the bottom 25%		grass, awning, people bicycle, biker, biking spectator, helmet competitor, athlete crowd, dirt, tree event, outdoor, man race, garbage can mud, rain, sweat tent, tent pole	cowboy broodmare large Hyla gray	woody **tree** horse equine perissodactyl	**tree** wear horse good cover	market vegetable festival shirt street	market vegetable festival shirt street
		change dispenser equipment machine, public vending machine	gas readout fire generator Hyla	gas readout pump electronic mechanical	pump device gas **equipment** readout	logo sign bead desk bedroom	sign logo bead desk bedroom

Figure 7. Example translations on **Dataset B**. 1^{st} col shows images. 2^{nd} col shows MTurk associated nouns. These represent the ground truth annotations (entry-level categories) we would like to predict (colored in blue). 3^{rd} col shows predicted nouns using a standard multiclass flat-classifier. 4^{th} col shows nouns predicted by the method of [2]. 5^{th} col shows our n-gram based method predictions. 6^{th} col shows our SVM mapping predictions and finally the 7^{th} column shows the labels predicted by our joint model. Matches are colored in green. Tables 1, 2 in the main paper show the measured improvements in recall and precision.

References

[1] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. In *CVPR09*, 2009. 2

[2] J. Deng, J. Krause, A. C. Berg, and L. Fei-Fei. Hedging your bets: Optimizing accuracy-specificity trade-offs in large scale visual recognition. In *CVPR*, 2012. 4, 5, 6, 7