A. Prompt Design and Queried Attributes

A.1. GPT3

Inspired by recent work on querying LLMs [47, 35], we start with the following prompt and a demonstration to query GPT3:

```
Q: What are useful visual features to
distinguish a lemur in a photo?
A: There are several useful visual
features to tell there is a lemur in
a photo:
- four-limbed primate
- black, grey, white, brown, or
red-brown
- wet and hairless nose with curved
nostrils
- long tail
- large eyes
- furry bodies
- clawed hands and feet
Q: What are useful visual features
to distinguish class_name in a photo?
A: There are several useful visual
features to distinguish class_name in a
photo:
```

To elicit knowledge within a certain domain, we also test the following prompt to specify the domain given a task:

`Q: What are useful visual features to distinguish` *class_name* `from other` *domain_name* `in a photo? A: There are several useful visual features to distinguish` *class_name* `from other` *domain_name* `in a photo:`

Here *class_name* is the name of each class in the datasets. For instance, in CIFAR-10, *class_name* is from {`airplane, automobile, bird, cat, deer, dog, frog, horse, ship, truck`}. *domain_name* is the domain of the datasets. We set *domain_name* to be `birds, objects, objects, flowers, foods, dogs and cats, cars, animals` for datasets CUB, CIFAR-10, CIFAR-100, Flowwer, Food, Oxford-pets, Stanford-cars, Imagenet-Animals, respectively.

A.2. ChatGPT

Query ChatGPT for CUB As CUB is in a specific domain of bird species, we use ChatGPT to design structured and compositional attributes as in [44]. Specifically, we first query ChatGPT with the following prompts to obtain the possible names describing *body parts* for the birds:

`What are the possible body parts to visually distinguish birds in the photo?`

We obtain the following attributes for *body parts*:

$\mathcal{BP}$ ={`wings, beak, feet, tail, head, breast, abdomen, leg, feathers`}.

Then we query for possible colors:

`What are the possible colors that are possible to appear on a bird?`

which results in a set of colors:

$\mathcal{C}$ ={`red, orange, yellow, green, blue, purple, brown, black, white, gray`}.

Then we query the shapes of each possible body part, take `wings` for example:

`What are the possible shapes for bird wings?`

Finally, with the colors $\mathcal{C}$ and the possible shapes for each body part shown in Table 7, we build 440 attributes for CUB with examples shown below:

```
red swallowtail or fork-tailed wings.
red round wings.
green webbed legs.
orange round wings.
```

Query ChatGPT for CIFAR-100 We utilize the batch prompting method described in Section 2 to query ChatGPT for the attributes of CIFAR-100.

We use the following prompt to query ChatGPT:

`Q: Here are five` *superclass_name*`:` {*class_name_1,...,class_name_N*}`. What are the useful visual features for distinguishing them in a photo? Please list every attribute in bullet points.`

Here *superclass_name* is the name of each superclass in the datasets. For instance, in CIFAR-100, beaver, dolphin, otter, seal, whale belongs to the superclass aquatic mammals.

A.3. Comparing Different Attribute Concept Pools

With the above prompting templates, we explore the effects of different concept pools. Comparing with the concept pool constructed from GPT-3 prompts with corresponding class names for each dataset, we add the following two pools to discuss the effects: (1) the concepts from Imagenet queried from GPT-3, which would be larger and also noisier; (2) The concepts from ChatGPT. Note that we conduct the ablation study on two dataset CUB and CIFAR-100, since their classes are generally covered by the classes from Imagenet. For CUB, we manually designed the attributes in the pool and for CIFAR-100, the attributes are

Body parts	Possible shapes
wings	`Swallowtail or fork-tailed wings, Round wings, Long, narrow wings, Short, broad wings, Elliptical wings`
beak	`Conical beaks, Hooked beaks, Probe-like beaks, Wide, flat beaks, Short, stubby beaks, Long, thin beaks`
feet	`Webbed feet, Talons, Perching feet, Scaling feet, Running feet`
tail	`Fan-shaped tails, Square-shaped tails, Rounded tails, Forked tails, Tails with streamers`
head	`Conical heads, Round heads, Elongated heads, Wide heads, Stout heads, Narrow heads`
breast	`Flat breasts, Round breasts, Bulky breasts, Slender breasts`
leg	`Long and slender legs, Short and thick legs, Webbed legs, Talons legs, Perching legs`
abdomen	`Round and plump abdomen, Slim and streamlined abdomen, Long and thin abdomen, Puffed out abdomen`
feathers	`Long, narrow feathers, Short, broad feathers, Round feathers, Streamer-like feathers`

Table 7: Possible shapes for each body part of birds.

queried in a hierarchical way (see Appendix A for the details). The results are shown in Table 8.

Overall, our learning-to-search method is robust to different attribute pools, and we do not observe significant performance change using GPT-3 or ChatGPT. Though our human-designed compositional attributes with ChatGPT on CUB is worse than pure LLM-generated attributes.

Datasets	CUB			CIFAR-100		
K	8	16	32	8	16	32
GPT-3	**31.67**	48.55	60.27	**34.77**	**52.24**	66.30
GPT-3-Imagenet	30.81	**49.29**	**60.41**	33.80	51.01	65.61
ChatGPT	21.66	40.28	47.46	33.79	51.26	**67.06**

Table 8: Comparison *w.r.t.* different Concept Pools.

A.4. Robustness Check

To confirm the effectiveness of our prompts and the robustness of GPT-3 prompts, we conduct the experiments with the concepts queried from GPT3 using different prompts. We design semantically instructive and misleading prompts as shown in Table 9. Overall, we observe that other instructive prompts perform similar to ours, while misleading prompts could hurt the performance drastically.

B. Implementation Details

B.1. Linear Probing

After we obtain attribute embeddings $\mathbf{T}^*$ from Eq.(6), we then calculate the semantic vector $\mathbf{A}^*$ of each image I with the D-dimensional image embedding $\mathbf{V} = \Theta_V(I) \in \mathbb{R}^D$:

$$s_j = \cos(\mathbf{V}, \mathbf{T}_j^*), j = 1, \cdots, K, \tag{8}$$

$$\mathbf{A}^* = (s_1, \cdots, s_K)^\top. \tag{9}$$

Then we calculate the score vectors of all the images in the training and testing dataset. We then use linear probing to evaluate the performance. Since we use a task-guided searching during the first stage to find K attribute embeddings, we can readily use the classification head in the first stage (i.e., a linear model $f_\theta \in \mathbb{R}^K \rightarrow \mathbb{R}^{K_C}$ with one fully connected layer) for our second stage with lightweight finetuning instead of training from scratch. where K_C is the number of classes. Then, we train f_θ with a cross-entropy loss:

$$\mathcal{L} = -\frac{1}{M}\sum_{i=1}^{M}\sum_{c=1}^{K_C} y_{i,c} \log p'_{i,c}, \tag{10}$$

the same as Eq.(2), M is the number of images in a mini-batch. $y_{i,c}$ is the binary indicator of i-th image the mini-batch belonging to class c, and $p_{i,c}$ is the predicted probability of the i-th image belonging to class c. Then $p'_{i,c}, c \in \{1, \cdots, K_C\}$ is calculated as:

$$[p'_{i,1}, \cdots, p'_{i,K_C}]^\top = \mathrm{Softmax}(f_\theta(\mathbf{A}_i^*)) \tag{11}$$

where $\mathbf{A}_i^*$ is the semantic vector of the i-th image in the mini-batch. Then we will use f_θ to classify the images in the test set to yield the performances.

Category	Prompts	Acc
Instructive	`What are the useful visual features to distinguish` *class_name*?	31.67
	`What are the helpful visual features to distinguish` *class_name*?	**32.71**
	`What are the distinctive visual features to distinguish` *class_name*?	30.38
Misleading	`What are the useless visual features to distinguish` *class_name*?	19.64
	`Give me some random visual features in a photo to distinguish` *domain_name*:	5.85

Table 9: Robustness study against different prompts on CUB, with $K = 8$

C. Additional Experiments

C.1. Comparison with Zero-Shot Classifications

Datasets	CIFAR-100	Stanford-cars
CLIP-ZS w/ class names	54.49	57.87
CLIP-ZS w/ attributes	30.07	5.42
CLIP-Train Visual	79.30	79.95
Ours (K=512)	75.41	74.67
Datasets	Flower	Imagenet-Animals
CLIP-ZS w/ class names	60.19	59.44
CLIP-ZS w/ attributes	9.80	9.13
CLIP-Train Visual	92.35	75.31
Ours (K=512)	90.29	75.60

Table 10: Comparison with zero-shot classification methods.

We deliver more results in Table 10. We use *A photo of* as the prompt for all methods. Zero-shot (CLIP-ZS) is worse than supervised training. Note that CLIP-ZS with class names may not be a fair comparison, **as our goal is to classify images with attributes instead of class names**, thereby gaining a level of interpretability and fine-grained understanding of visual recognition. If we use only attributes for CLIP-ZS, the performance drastically decreases.

C.2. Human Evaluation

To further evaluate the quality of our learned attributes, we conduct a pairwise human evaluation on Amazon Mechanical Turk. Specifically, we compare our attributes with uniformly sampled attributes from the GPT-3 generated attributes, and ask human to decide which set of attributes are better. Since datasets with hundreds of classes are hard to reason and compare, we evaluate our results on CIFAR-10. We sample 100 sets of 4 attributes from the attribute pool, and create 100 pairs of each random set of 4 attributes with our learned 4 attributes. Each pair was assigned to 5 workers to eliminate human variance. For each attribute pair, workers are presented with sampled images from CIFAR-10. We instruct the workers to consider which set of attributes are more useful to classfy the 10 classes. As shown in Table 11, even though in most cases, the attributes look similar to human, workers still favor *Our method* over *Uniform sampling*, which is consistent with the classification accuracy.

Choice (%)	Ours	Uniform	Tie
Score	31.6	19.0	49.4

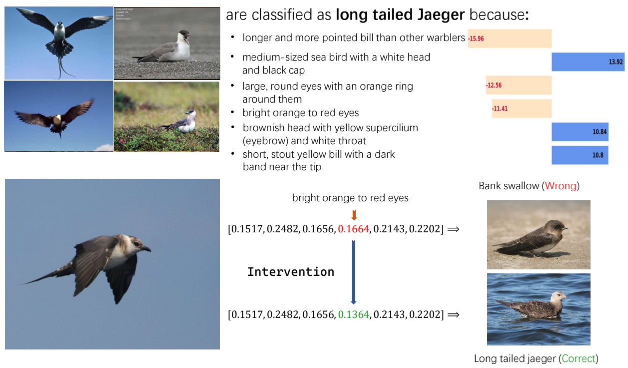
Table 11: Human evaluation results on CIFAR-10. Human are asked to vote which attributes are better, where *tie* means the two sets looks the same to annotators

Model Architectures	8	16	32
CLIP RN50	24.28	38.95	56.11
CLIP RN101	27.36	46.10	56.06
CLIP RN50x16	28.91	53.46	57.69
CLIP ViT-B/32	31.67	48.55	60.27
CLIP ViT-B/16	36.69	55.64	63.70
CLIP ViT-L/14	38.71	65.52	74.99
CLIP ViT-L/14@336px	40.95	66.58	76.04
Open-CLIP ViT-H-14 LAION-2B	49.84	73.28	82.60

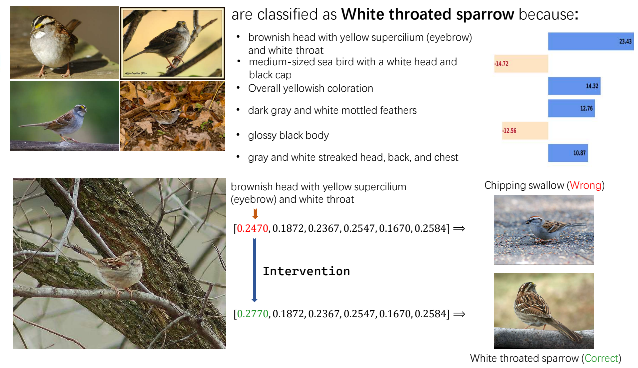
Table 12: Ablation study on different VLMs with bottleneck size K=8,16,32 on the CUB dataset.

D. More ablations

Better V&L models We evaluate different variants of CLIP style models, as shown in Table 12. Overall, our method is model-agnostic. It can be applied with any VLMs that compute image-text similarities. We also observe that in general, a stronger VLM will result in more accurate estimation of semantic vectors, hence improves classfication performance.



(a) Long Tailed Jaeger



(b) White Throated Sparrow

Figure 8: Examples on interpretability and interactivity. (1) The upper half of each figure show important attributes for two classes of birds. We choose 6 out of 32 attributes with highest importance scores, which are computed by multiplication between clip scores and weights in the linear probe, defined in Eq. (7). (2) The lower half of each figure demonstrates the intervention on the semantic vector (i.e., CLIP scores) to correct the prediction, we use δ=0.03 for all interventions on clip scores as an empirical value. The array of 6 scores are of the same order as the attributes.

Datasets	CUB			CIFAR-100		
# of Non-zeros	8	16	32	8	16	32
One-Hot	36.22	44.35	48.96	57.90	61.85	65.43
Only Top Scores	36.03	44.32	49.43	58.32	61.90	65.86

Table 13: Comparison between scores and one-hot

Effectiveness of semantic projection. Scores vs one-hot In this part, we consider the following baseline to show that the information within the similarity scores are useful: For an image I, after calculating all the similarity scores between every attribute $a_i \in \{a_1, \cdots, a_N\}$ and the image I to obtain the vector $\mathbf{A} \in \mathbb{R}^N$. Then we wipe out the information in the scores by setting top-K large scores in $\mathbf{A}$ as 1, and setting the left scores as 0, which will give us a binary vector $\mathbf{A}_{bin} \in \{0, 1\}^N$. Then we train and test the classification model on the corresponding binary vectors to compare with our methods. We conduct the ablation study on CUB as an example. We choose $K = 8$ and $K = 16$ for the comparisons.

From the results in Table 13, we observe the information from the similarity scores provides information for classification, while removing the information in the scores (converting top-K scores to 1) may lead to performance drop.

GPT-3 attributes vs. Random words We present more results on 8 datasets to verify that a large number of GPT-3 attributes behaves similar as random words. The observations are coherent on all eight datasets. When K is large, even if we randomly create K meaningless phrases from the entire vocabulary, we can still obtain competitive classification performance.

	Nonsense		GPT3	
K	4	512	4	512
CUB	2.42	64.79	**12.98**	67.64
CIFAR-10	31.26	92.81	**60.30**	93.67
CIFAR-100	10.17	77.40	**16.13**	77.55
Flower	3.33	90.20*	**28.92**	90.78*
Food	7.79	82.65*	**16.23**	82.50*
Oxford_pets	14.31	83.61*	**28.07**	86.01*
Stanford_cars	5.06	75.09	**13.41**	75.13
Imagenet_Animals	3.78	75.12	**8.81**	75.75

Table 14: GPT3 vs Nonsense. $*$ means the results are obtained when setting the number of attributes to the size of the concept pool from GPT3, which corresponding to the results of "full" in Figure 4.

E. Additional Case Study

E.1. Test-time Intervention

We provide more case studies in Figure 8 to show the interpretability and interactivity of our method.

E.2. Visualization of Discovered Attributes

We present our learned 32 attributes for each dataset (by setting $K = 32$) in Table 15 and 16. Similar to Figure 6, we can observe these attributes are distinctive within each domain, and provides fine-grained attributes to summarize a dataset. To some level, we can view these automatically learned attributes as a form of knowledge to help humans understand how visual recognition works.

Dataset	Learned 32 attributes for each dataset
CUB	(1) Brown, gray and white feathers on the upper parts of the body, with a rusty red or pinkish tinge to the head; (2) bright yellow and black coloring; (3) distinctive white throat; (4) Broad tail that is shorter than other pelican species ; (5) short legs for perching in trees ; (6) bright yellow throat, breast, and flanks with black bars ; (7) Brown and white mottled back ; (8) pinkish red breast patch with white edges ; (9) white throat patch bordered by black stripes ; (10) unique pattern of spots on lower throat and breast.; (11) Large feet for scratching in leaf litter; (12) brownish head with yellow supercilium (eyebrow) and white throat ; (13) red or orange coloration; (14) iridescent black body with blue and purple highlights ; (15) red, black and white feathers; (16) grayish brown body with darker wings and tail; (17) Heavy bill for crushing seeds ; (18) olive green back and wings; (19) white throat, belly and wing bars ; (20) Long, slender bill with yellow tip; (21) large, white bird; (22) grayish brown head and back; (23) red/orange coloration on the face during breeding season; (24) Gray head and yellow throat with white eye rings; (25) barred wings and tail feathers in black, white and grey patterning ; (26) distinctive white throat patch ; (27) long, heavy bill; (28) yellow head; (29) male mallards have a green head, yellow beak and white neck ring; (30) Broad white eyering ; (31) white throat and belly region; (32) bright orange and black plumage
CIFAR-10	(1) antlers (in males); (2) some birds have crests on their heads; (3) propellers or jet engines; (4) fur coat of varying colors and patterns; (5) a tail with a horizontal stabilizer; (6) portholes along the hull; (7) large body with a cab and a bed; (8) landing gear; (9) four wheels; (10) fuselage and other structural elements; (11) four-wheeled vehicle; (12) long head with a mane and tail; (13) has a mast with sails or flags; (14) paws with clawed toes; (15) feathers of various colors and patterns; (16) tail lights; (17) furry body; (18) a beak or bill for eating, preening, and other activities; (19) pointed bow and stern; (20) mane of hair along the neck and back; (21) masts, sails, and rigging; (22) two wings and two legs; (23) moist slimy skin; (24) windshield and side windows; (25) rudders at the stern for steering; (26) smokestacks or funnels on top of the ship; (27) a large deck or superstructure; (28) hooves on each foot; (29) typically has a steering wheel and pedals for driving; (30) control surfaces (flaps, ailerons, rudder); (31) grille or front fascia; (32) may have a cargo area in the back
CIFAR-100	(1) A pair of pedipalps near the mouth used for sensing, holding prey and mating; (2) multiple petals in shades of pink, red, yellow or white; (3) Headboard and footboard; (4) pouch on the abdomen of female kangaroos; (5) fruits are two winged samaras in clusters; (6) large courtyard area surrounded by buildings; (7) orange or yellow fur with black stripes; (8) buds clustered at branch tips in winter months; (9) large wheels and tires; (10) Mattress and bedding; (11) large, floppy ears; (12) silver scales with black and red spots on the sides; (13) long snout with sharp teeth; (14) smooth oval shaped sepals; (15) clustered, coconuts at the tip of each branch; (16) white foam from waves breaking against rocks and shorelines; (17) designs, colors, or patterns on the can; (18) shaggy fur; (19) tailgate at rear end; (20) long bushy tail usually with a tuft of hair at the end; (21) portcullis at entrance to gatehouse; (22) cabin or operator's seat in the middle of the vehicle; (23) drawbridge over a moat; (24) dialing pad with numbers 0–9; (25) catkins (flowers) in spring; (26) a large, tawnycolored body with a shaggy mane; (27) Ten walking legs and two large antennae; (28) Rim around the edge of the bowl; (29) armrests and backrests; (30) Long stem or pole extending from the shade to the base; (31) an ovary located at the base of the flower; (32) waxy texture of the petals and leaves
Flower	(1) bright purple petals that are fused together to form a thistle-like shape; (2) trumpet shaped flowers in shades of blue, purple, and white; (3) blue, purple, or white flowers with a thistle-like appearance; (4) an upright inflorescence (flower spike) bearing several clustered flowers on each branch ; (5) umbel of several small flowers on top of a single stem ; (6) Center of the spathe that looks like a tail or spadix; (7) the flower is a daisy-like plant with white petals and yellow center; (8) layered petals with a yellow center and pink edges ; (9) Large, bright pink to red flower ; (10) Six distinct petal segments surrounding an inner cup of short filaments and a trumpet center.; (11) trumpet shaped orange center with yellow stamens protruding from it ; (12) large, white petals with a yellow center; (13) fragrant single or double blooms in white, pink, or red; (14) brightly colored petals in shades of oranges, reds, and yellows; (15) large, yellow petals that form a daisy-like shape; (16) umbrella shaped clusters of white to pinkish flowers ; (17) tall, leafless stem; (18) hibiscus shaped leaves that are serrated around the edges; (19) pink to purple colored petals with red lips; (20) single stem with a rosette of leaves; (21) tall, slender stem with a single umbel of flowers; (22) large blue or purple flowers with five petals and a hooded center; (23) intricate patterns of blue, purple, pink and white lines on the petals ; (24) yellowish green sepals below the flower; (25) dark purple petals; (26) five petals arranged around a central column of white stamens and stigma ; (27) pink, white, or lavender flowers with five petals; (28) bright yellow flower head; (29) bright pink, red, or white petals with fringed edges; (30) bright red, orange, or yellow blooms; (31) bright red and yellow petals; (32) deep red, orange or yellow petals

Table 15: Learned 32 attributes on CUB, CIFAR-10, CIFAR-100 and Flower.

Dataset	Learned 32 attributes for each dataset
FOOD	(1) sliced strawberries arranged over the cream/whipped topping; (2) A white or wheat bun with a golden brown exterior ; (3) Long, thin rice noodles; (4) lattice pattern on the top layer made by weaving strips of pastry dough; (5) slices of apples arranged in a spiral pattern ; (6) large pieces of clams visible in the chowder; (7) Gyoza is typically shaped like a half-moon or dumpling and can have either open or closed tops. ; (8) The broth will usually have either a sour or spicy taste depending on the type of soup ; (9) red sauce layered between the noodles and cheese; (10) shredded carrots embedded within the cake ; (11) dollop of sour cream or guacamole ; (12) thin layers of phyllo dough; (13) tender squid rings inside ; (14) distinctive pattern of takoyaki sauce on top; (15) moist and dark brown cake with visible cocoa powder; (16) Served on top of a bed of shredded daikon radish or grated daikon ; (17) lobster chunks mixed with mayonnaise and spices; (18) two layers of toast with lettuce, bacon, and tomatoes in between; (19) mashed avocado texture ; (20) cooked shrimp in a variety of colors (pink, orange, etc.); (21) steaming bowl of soup with steam rising up ; (22) melted cheese over the chips; (23) The presence of Mandarin pancakes, cucumber slices and spring onion used in traditional preparation methods; (24) fried or served cold with dipping sauce; (25) a custard base topped with a layer of hardened caramelized sugar; (26) a dollop of gochujang (red pepper) paste ; (27) butter or oil is used to toast the bread on both sides ; (28) chunks of vegetables, tofu, and seaweed floating in it; (29) gooey mixture of sugar, butter and cinnamon visible between the layers of apple slices; (30) olive oil and soy sauce dressing ; (31) toppings such as egg, vegetables, seaweed, and pork slices ; (32) toppings such as jalapenos, tomatoes, onions and/or peppers
Oxford-pets	(1) Long legs and neck; (2) large upright ears; (3) "Ragdoll" appearance with a long body and short legs; (4) Soft wiry coat in black or brindle colors; (5) Pointed ears; (6) Shade of red or wheaten color; (7) dark brown or black coat with white markings; (8) Markings resembling a leopard or tiger in various colors (brown, black, white, orange); (9) mediumsized dog; (10) greyish blue fur with silver tips; (11) short, almost hairless body with wrinkles; (12) Round eyes in shades of blue or green; (13) White and grey fur; (14) A tail that curls over its back; (15) Ears that are small and rounded at the tips; (16) loose skin on the face and neck that can create wrinkles; (17) triangular ears; (18) white blaze on face and chest; (19) droopy ears that hang close to the head; (20) wide eyes with prominent wrinkles around them; (21) thick, white double coat; (22) foxy head and face with a curled tail; (23) Curly tail that curls over the back; (24) black face mask on white fur background; (25) long, silky coat in white or white and black colors; (26) thick mane around neck and chest; (27) distinctive wrinkles on the face; (28) short coat of glossy black fur; (29) Short, glossy coat of black and silver; (30) double coat of fur that is typically fawn, black or silver; (31) black and tan coloring; (32) Visible spots on the body
Stanford-cars	(1) large grille with a classic Bentley badge; (2) Signature wheel arches; (3) large tailgate spoiler on the liftgate; (4) unique wheels with five spokes and silver finish; (5) Front bumper has a skid plate design; (6) flared wheel arches that give the car an aggressive look; (7) Ron Fellows Edition badge on the rear of the car; (8) The distinct hexagonal grille with the Volvo emblem at the center; (9) Distinctive red and black racing stripes with Abarth logo; (10) "4Runner" badge on the rear liftgate; (11) Hatchback style trunk/boot area; (12) High performance tires with "Type R" on the sidewall; (13) horizontal three bar tail lamps with the running Mustang logo at its center.; (14) distinctive grille with a mesh pattern and Spyker logo; (15) Interior: Leather wrapped steering wheel with audio controls; (16) kidney grille with large blue and white BMW logo; (17) black power convertible top; (18) Red badge with "Integra Type R" logo or lettering on the hood and trunk lid; (19) Chrome grille with the Chevrolet logo; (20) the Fiat logo on the front grille and rear of car; (21) Quattro badge on the rear right side of the car; (22) gloss black paint job with distinct yellow detailing; (23) LED tail lights with a unique curved design to give it a modern look.; (24) Chrome grille with the Chrysler emblem in the center; (25) chrome grille with the Chevrolet logo; (26) wide grille with a large chrome Bentley badge in the center; (27) two door hardtop convertible body style; (28) distinctive side windows with curved lines and signature Maybach logo; (29) tailgate spoiler on the rear hatchback door; (30) Chrome grille outlining the Honda logo; (31) Black brake calipers with Corvette lettering; (32) Twodoor, fourseater convertible hardtop
Imagenet-Animals	(1) smaller and lighter than other Welsh corgis; (2) from Airedale, England; (3) male rams have large, thick horns, while female rams have smaller, thinner horns; (4) dense, flat coat; (5) the Maltese has a reputation for being lively, playful and affectionate; (6) coat is predominantly black and tan; (7) Gordon setters are typically black with tan markings; (8) Saint Bernards are large dogs; (9) Kerry blue terriers are from Ireland; (10) the breed name (elkhound); (11) English setters are bred in England; (12) shaggy, matted coat; (13) all black coat; (14) spotted or striped fur; (15) dark brown or black coat with a distinctive "water spaniel" curl; (16) glossy black feathers with a green or blue sheen; (17) pink, orange, or yellow stripes on the shell; (18) black and white, blue and white, or wheaten (red) coloration; (19) wrinkles on the face and head; (20) coat is wheaten in color (ranging from pale cream to rich gold); (21) big ears; (22) male finches have a bright red breast; (23) creamy white or wheaten-colored coat; (24) white throat; (25) large, broad carapace; (26) dark plumage with black and iridescent blue feathers; (27) long, black antennae; (28) large, black and white dolphin; (29) longer legs than other hound breeds; (30) fawn or brindle coloration; (31) long, wirehaired coat; (32) fawn to mahogany coat

Table 16: Learned 32 attributes on Food, Oxford-pets, Stanford-cars and Imagenet-Animals.