

SUPPLEMENTARY MATERIALS

Note: We provide detailed figures for multi-aspect logit distillation in Section A, implementation and dataset details in Section B, additional ablation study results in Section C, further details on the visualization of the logit distribution in Section D, and computational cost analyses in Section E, which were not included in the main paper due to space limitations.

A DETAILS OF MULTI-ASPECT LOGIT DISTILLATION

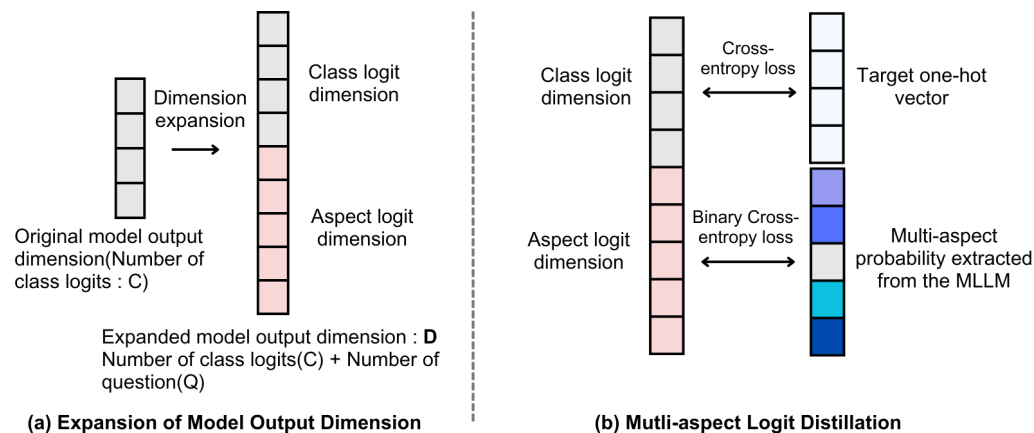


Figure 1: **Multi-aspect knowledge distillation.** To distill knowledge about multi-aspect questions into the model, we simply expand the dimension of model output. Also, we consider the expanded dimensions as the class logit dimension and the aspect logit dimension. We apply cross-entropy loss to the class logit dimension and binary cross-entropy loss to the aspect logit dimension.

B IMPLEMENTATION DETAILS

B.1 DATASET DETAILS

StanfordCars Krause et al. (2013). StanfordCars contains 16,185 images of 196 classes of cars. The data is split into 8,144 training images and 8,041 testing images, where each class has been split roughly in a 50-50 split. Classes are typically at the level of Make, Model, Year, ex. 2012 Tesla Model S or 2012 BMW M3 coupe.

OxfordPets Parkhi et al. (2012). OxfordPets comprises 7,384 images of 37 distinct cat and dog breeds, with around 200 images per class. It is divided into 3,690 images for training and 3,694 images for testing. The dataset features significant variations in scale, pose, lighting, and others.

Describable Textures Dataset (DTD) Cimpoi et al. (2014). DTD consists of 47 texture classes and a total of 5,640 images. It is divided into 3,760 images for training and 1,880 for testing, with each class containing 120 images. The image sizes range from 300x300 to 640x640 pixels, and each image contains at least 90% of the surface area representing the category’s attribute.

102Flowers Nilsback & Zisserman (2008). 102Flowers is designed for image classification, featuring 102 different flower classes. It is divided into 6,552 training images and 1,637 testing images. Each class includes between 40 and 258 images, with significant variations in scale, pose, and lighting conditions across the images.

CUB200 Wah et al. (2011). CUB200 is one of the most commonly used datasets for fine-grained visual categorization tasks. It comprises 11,788 images across 200 bird subcategories, with 5,994 images for training and 5,794 for testing. Each image has detailed annotations, including 1 subcategory label, 15 part locations, 312 binary attributes, and 1 bounding box.

054 **FGVC-Aircraft Maji et al. (2013).** FGVC-Aircraft consists of 9,967 aircraft images, with around
055 100 images corresponding to each of the 100 different model variants, the majority being airplanes.
056 The dataset is divided into 6,667 images for training and 3,300 for testing. Each image includes
057 annotations with a tight bounding box and a hierarchical label for the airplane model. The aircraft
058 models are arranged in a four-level hierarchical structure.

059 **Caltech101 Fei-Fei et al. (2004).** Caltech101 includes images from 101 object categories, along
060 with a background category consisting of images unrelated to those 101 categories. To focus purely
061 on class classification, we exclude the background category. The dataset is divided into 4,310 images
062 for training and 4,367 images for testing. Each category contains between 40 and 800 images, with
063 most classes having approximately 50 images. The image resolution is roughly 300×200 pixels.

064 **Mini-ImageNet Ravi & Larochelle (2016).** Mini-ImageNet is a reduced version of the larger
065 ImageNet Deng et al. (2009) dataset, specifically designed for few-shot learning tasks. It consists
066 of 50,000 training images and 10,000 testing images distributed across 100 classes. Additionally, to
067 use a higher resolution, we utilize the dataset from Ravi & Larochelle (2016).

068 **Microsoft Common Objects in Context (MS-COCO) Lin et al. (2014).** MS-COCO is a large-
069 scale object detection, segmentation, key-point detection, and captioning dataset. The dataset con-
070 sists of 328K images. We use the MS-COCO dataset’s 2017 version, which consists of a train-
071 ing/validation split of 118K/5K images.

073 B.2 TRAINING DETAILS

074
075 For the image classification experiments, we employed baseline models such as ResNet18,
076 ResNet34, MobileNet-V1, and EfficientNet-b0 across various fine-grained datasets, including Stan-
077 fordCars, OxfordPets, DTD, 102Flowers, CUB200, and FGVC-Aircraft, as well as coarse-grained
078 datasets such as Caltech101 and Mini-ImageNet.

079 **Data preprocessing.** Input images were normalized using the channel-wise mean (0.485, 0.456,
080 0.406) and standard deviation (0.229, 0.224, 0.225) for RGB channels. For training, we applied a
081 series of transformations: RandomResizedCrop with a target size of 224, followed by RandomHor-
082 izontalFlip, conversion to tensor using ToTensor, and normalization.

083 **Hyperparameters for fine-grained datasets.** The models were trained for 240 epochs with a batch
084 size of 16. The initial learning rate was set to 0.01 and decreased by a factor of 10 at the 150th,
085 180th, and 210th epochs. We use the SGD optimizer with a momentum of 0.9 for all experiments,
086 and weight decay is set to 5e-4.

087 **Hyperparameters for Caltech101 dataset.** The models were trained for 240 epochs with a batch
088 size of 16. The initial learning rate was set to 0.01 and decreased by a factor of 10 at the 150th,
089 180th, and 210th epochs. We use the SGD optimizer with a momentum of 0.9 for all experiments,
090 and weight decay is set to 5e-4.

091 **Hyperparameters for Mini-ImageNet dataset.** The models were trained for 100 epochs with a
092 batch size of 64. The initial learning rate was set to 0.2 and decreased by a factor of 10 at the 30th,
093 60th, and 90th epochs. We use the SGD optimizer with a momentum of 0.9 for all experiments, and
094 weight decay is set to 5e-4.

096 B.3 MULTI-ASPECT QUESTION SAMPLES

097
098 Table 8, 9, 10, 11, 12, 13, 14, and 15 present the multi-aspect questions generated by GPT-4o for
099 the StanfordCars, OxfordPets, DTD, 102Flowers, CUB200, FGVC-Aircraft, Caltech101, and Mini-
100 ImageNet datasets, respectively. Meanwhile, Table 16, 17, and 18 show the multi-aspect questions
101 generated by GPT-3.5-turbo for the StanfordCars, OxfordPets, and Caltech101 datasets, respec-
102 tively.
103

104 B.4 DETAILS OF LOGIT DISTILLATION WITH MLLM

105 We also include how MLLM distills logit to student model, as described in Section 4.4 and 5.1 of
106 the main paper.
107

In traditional knowledge distillation, the teacher model typically outputs soft targets as a probability distribution over classes. The student model is then trained based on the KL divergence loss between the soft targets and the student’s predicted target logits, as well as the cross-entropy loss with the actual hard targets. To enable MLLM to perform logit distillation, we make the following assumption:

‘Could the logits generated for both the predicted class index token and the remaining class index token logits during zero-shot classification be considered soft targets?’

In this assumption, since MLLM receives information about the range of possible answers through prompts, it restricts the range of tokens generated. Given a first logit vector $\mathbf{z} \in \mathbb{R}^V$, where V is the vocabulary size of tokenizer, the logits corresponding to the numerical tokens are indexed by the set $\mathcal{N} \subseteq \{1, 2, \dots, C\}$. The softmax function applied to the logits of the numerical tokens within a specified range $[1, C]$ is given by:

$$P(t \mid 1 \leq t \leq C) = \frac{\exp(z_t)}{\sum_{n \in \mathcal{N}, 1 \leq n \leq C} \exp(z_n)} \quad \text{for } t \in \mathcal{N}, 1 \leq t \leq C$$

where z_t is the logit corresponding to token t and the sum in the denominator is computed over all numerical tokens n in the range $[1, C]$. This can be interpreted as MLLM producing a probability distribution for classifying specific classes, allowing it to generate soft targets as a teacher model. These generated soft targets can be used for training in the same way as in traditional knowledge distillation, as they remain unchanged while the student model is being trained.

C ADDITIONAL ABLATION STUDY RESULTS

In this section, we provide the additional ablation results for the OxfordPets fine-grained dataset and Caltech101 coarse-grained dataset. The additional ablation study results are presented in Table 1 and 2, showing the outcomes for the OxfordPets and Caltech101 datasets, respectively. Table 3 displays the results of the extension to logit distillation on the OxfordPets dataset.

Table 1: **Additional Ablation study on OxfordPets.** Each table reports the accuracy(%) on OxfordPets. Res18 for ResNet18, Res34 for ResNet34, Mb-N1 for MobileNetV1 and EffiNet for EfficientNet-b0. Rand for our method with random logits instead of multi-aspect logits. KL for our method with KL-Divergence loss on multi-aspect logit. α for the weighting factor of multi-aspect logit loss. We run each experiment 3 times and report the average results.

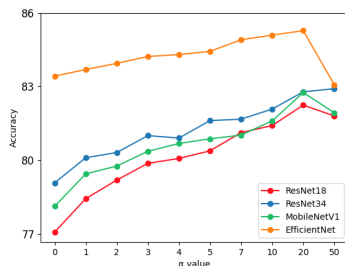
(a) Effect of the loss function

	Res18	Res34	Mb-N1	EffiNet
KL	75.96	79.52	79.71	83.92
Ours	82.24	82.78	82.75	85.27

(b) Effect of the multi-aspect logit

	Res18	Res34	Mb-N1	EffiNet
Rand	78.64	79.17	77.70	83.23
Ours	82.24	82.78	82.75	85.27

(c) Weights to the multi-aspect loss



(d) Effect of LLM and MLLM

	Res18	Res34	Mb-N1	EffiNet
Base	77.07	79.07	78.12	83.42
Ours(L: GPT-3.5)	82.72	83.08	82.66	85.22
Ours(M: LLaVA)	82.94	83.04	82.94	85.19
Ours	82.24	82.78	82.75	85.27

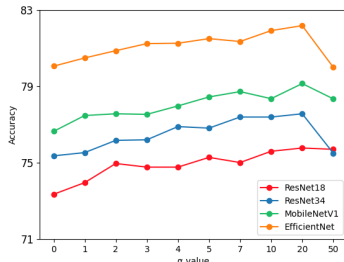
D ADDITIONAL DETAILS OF VISUALIZATION (HISTOGRAM, T-SNE AND ERROR BAR)

In this section, we provide additional visualizations for our analysis, including error bars for the experimental results on each dataset, visualization of the average logit distribution, visualization of

Table 2: **Additional Ablation study on Caltech101.** Each table reports the accuracy(%) on Caltech101. Res18 for ResNet18, Res34 for ResNet34, Mb-N1 for MobileNetV1 and EffiNet for EfficientNet-b0. Rand for our method with random logits instead of multi-aspect logits. KL for our method with KL-Divergence loss on multi-aspect logit. α for the weighting factor of multi-aspect logit loss. We run each experiment 3 times and report the average results.

(a) Effect of the loss function					(b) Effect of the multi-aspect logit				
	Res18	Res34	Mb-N1	EffiNet		Res18	Res34	Mb-N1	EffiNet
KL	75.10	74.74	77.67	80.73	Rand	73.90	74.94	76.64	79.80
Ours	75.77	77.56	79.14	82.17	Ours	75.77	77.56	79.14	82.17

(c) Weights to the multi-aspect loss



(d) Effect of LLM and MLLM

	Res18	Res34	Mb-N1	EffiNet
Base	73.35	75.36	76.64	80.05
Ours(L: GPT-3.5)	76.16	76.78	79.13	81.95
Ours(M: LLaVA)	76.02	77.46	78.81	81.65
Ours	75.77	77.56	79.14	82.17

Table 3: **Extension to logit distillation on OxfordPets.** We can simply extend our method to logit distillation. We run each experiment three times and report the average results.

Dataset	Teacher	ResNet34(79.07)	EfficientNet-b0(83.42)
	Student	ResNet18(77.07)	MobileNetV1(78.12)
OxfordPets	KD	79.01	80.90
	Ours + KD	82.68	83.13

t-SNE embedding, and visualizations of the probability values between the MLLM and the classification model for the multi-aspect questions.

Error bar. To help in evaluating the quality of the experiments, we include error bars representing the standard error for the conducted experiments. The error bars for the StanfordCars, OxfordPets, DTD, 102Flowers, CUB200, FGVC-Aircraft, Caltech101, and Mini-ImageNet datasets are presented in Figure 2. We run each experiment 3 times and report the average results.

Visualization of the average logit distribution. We provide the average logit distribution for all aspects of the datasets. The visualizations of the average logit distribution graphs for the StanfordCars, OxfordPets, DTD, 102Flowers, CUB200, FGVC-Aircraft, Caltech101, and Mini-ImageNet datasets are shown in Figure 3, 4, 5, 6, 7, 8, 9, and 10, respectively.

Visualization of t-SNE embeddings. We use t-SNE to reduce the dimensionality of the predicted aspect logit probabilities from our ResNet18 and the MLLM’s aspect logit probabilities to better visualize the results. For each dataset, we display the train and test results across 50 aspects. Yellow points represent a higher probability of ‘Yes’ (closer to 1), while purple points represent a higher probability of ‘No’ (closer to 0). The ground-truth and predicted result t-SNE embedding visualizations for the training data from the StanfordCars, OxfordPets, DTD, 102Flowers, CUB200, FGVC-Aircraft, Caltech101, and Mini-ImageNet datasets are shown in Figure 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, and 26.

E DETAILS OF COMPUTATIONAL COSTS

We calculate the computational cost on StanfordCars and OxfordPets using a single NVIDIA RTX 3090 GPU. Training refers to the average of three experiments and represents the number of seconds

216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269

taken per epoch. Inference indicates the number of seconds required to process all test sets. Table 4 shows that even with the extension of our method, there is no significant difference in time.

StanfordCars						
Aspect	ResNet18			ResNet34		
	Training	Inference	FLOPs	Training	Inference	FLOPs
0	25.418	20.5869	58.1873G	60.416	51.6867	117.4380G
10	25.916	21.0294	58.1875G	60.639	51.8848	117.4382G
20	26.411	21.4692	58.1877G	60.861	52.0821	117.4384G
30	26.910	21.9126	58.1878G	61.233	52.4127	117.4386G
50	27.903	22.7951	58.1882G	61.529	52.6757	117.4389G
OxfordPets						
Aspect	ResNet18			ResNet34		
	Training	Inference	FLOPs	Training	Inference	FLOPs
0	11.522	9.3810	24.0335G	16.169	12.5678	48.5060G
10	11.772	9.6062	24.0337G	16.520	12.8841	48.5062G
20	12.022	9.8315	24.0338G	16.871	13.2003	48.5064G
30	12.311	9.0919	24.0340G	17.277	13.5661	48.5066G
50	12.825	10.5550	24.0344G	17.998	14.2157	48.5069G

Table 4: The computational cost.

Table 5: **Dataset class indices.** We provide the class indices for DTD, 102Flowers, and FGVC-Aircraft, which have 47, 102, and 100 classes, respectively.

Index	DTD Class	102Flowers Class	FGVC-Aircraft Class
0	banded	alpine_sea_holly	707-320
1	blotchy	anthurium	727-200
2	braided	artichoke	737-200
3	bubbly	azalea	737-300
4	bumpy	ball_moss	737-400
5	chequered	balloon_flower	737-500
6	cobwebbed	barbeton_daisy	737-600
7	cracked	bearded_iris	737-700
8	crosshatched	bee_balm	737-800
9	crystalline	bird_of_paradise	737-900
10	dotted	bishop_of_llandaff	747-100
11	fibrous	black-eyed_susan	747-200
12	flecked	blackberry_lily	747-300
13	freckled	blanket_flower	747-400
14	frilly	bolero_deep_blue	757-200
15	gauzy	bougainvillea	757-300
16	grid	bromelia	767-200
17	grooved	buttercup	767-300
18	honeycombed	californian_poppy	767-400
19	interlaced	camellia	777-200
20	knitted	canna_lily	777-300
21	lacelike	canterbury_bells	A300B4
22	lined	cape_flower	A310
23	marbled	carnation	A318
24	matted	cautleya_spicata	A319
25	meshed	clematis	A320
26	paisley	coltsfoot	A321
27	perforated	columbine	A330-200
28	pitted	common_dandelion	A330-300
29	pleated	corn_poppy	A340-200
30	polka-dotted	cyclamen	A340-300
31	porous	daffodil	A340-500

	Index	DTD Class	102Flowers Class	FGVC-Aircraft Class
270				
271	32	potholed	desert-rose	A340-600
272	33	scaly	english_marigold	A380
273	34	smearred	fire_lily	ATR-42
274	35	spiralled	foxglove	ATR-72
275	36	sprinkled	frangipani	An-12
276	37	stained	fritillary	BAE-125
277	38	stratified	garden_phlox	BAE_146-200
278	39	striped	gaura	BAE_146-300
279	40	studded	gazania	Beechcraft_1900
280	41	swirly	geranium	Boeing_717
281	42	veined	giant_white_arum_lily	C-130
282	43	waffled	globe-flower	C-47
283	44	woven	globe_thistle	CRJ-200
284	45	wrinkled	grape_hyacinth	CRJ-700
285	46	zigzagged	great_masterwort	CRJ-900
286	47		hard-leaved_pocket_orchid	Cessna_172
287	48		hibiscus	Cessna_208
288	49		hippeastrum	Cessna_525
289	50		japanese_anemone	Cessna_560
290	51		king_protea	Challenger_600
291	52		lenten_rose	DC-10
292	53		lotus_lotus	DC-3
293	54		love_in_the_mist	DC-6
294	55		magnolia	DC-8
295	56		mallow	DC-9-30
296	57		marigold	DH-82
297	58		mexican_aster	DHC-1
298	59		mexican_petunia	DHC-6
299	60		monkshood	DHC-8-100
300	61		moon_orchid	DHC-8-300
301	62		morning_glory	DR-400
302	63		orange_dahlia	Dornier_328
303	64		osteospermum	E-170
304	65		oxeye_daisy	E-190
305	66		passion_flower	E-195
306	67		pelargonium	EMB-120
307	68		peruvian_lily	ERJ_135
308	69		petunia	ERJ_145
309	70		pincushion_flower	Embraer_Legacy_600
310	71		pink-yellow_dahlia	Eurofighter_Typhoon
311	72		pink_primrose	F-16A
312	73		poinsettia	FA-18
313	74		primula	Falcon_2000
314	75		prince_of_wales_feathers	Falcon_900
315	76		purple_coneflower	Fokker_100
316	77		red_ginger	Fokker_50
317	78		rose	Fokker_70
318	79		ruby-lipped_cattleya	Global_Express
319	80		siam_tulip	Gulfstream_IV
320	81		silverbush	Gulfstream_V
321	82		snapdragon	Hawk_T1
322	83		spear_thistle	Il-76
323	84		spring_crocus	L-1011
	85		stemless_gentian	MD-11
	86		sunflower	MD-80
	87		sweet_pea	MD-87
	88		sweet_william	MD-90
	89		sword_lily	Metroliner

324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377

Index	DTD Class	102Flowers Class	FGVC-Aircraft Class
90		thorn_apple	Model_B200
91		tiger_lily	PA-28
92		toad_lily	SR-20
93		tree_mallow	Saab_2000
94		tree_poppy	Saab_340
95		trumpet_creeper	Spitfire
96		wallflower	Tornado
97		water_lily	Tu-134
98		watercress	Tu-154
99		wild_pansy	Yak-42
100		windflower	
101		yellow_iris	

Table 6: **Dataset class indices.** We provide the class indices for StanfordCars, and CUB200, which have 196, and 200 classes, respectively.

Index	StanfordCars Class	CUB200 Class
0	AM General Hummer SUV 2000	Acadian_Flycatcher
1	Acura Integra Type R 2001	American_Crow
2	Acura RL Sedan 2012	American_Goldfinch
3	Acura TL Sedan 2012	American_Pipit
4	Acura TL Type-S 2008	American_Redstart
5	Acura TSX Sedan 2012	American_Three_toed_Woodpecker
6	Acura ZDX Hatchback 2012	Anna_Hummingbird
7	Aston Martin V8 Vantage Convertible 2012	Artic_Tern
8	Aston Martin V8 Vantage Coupe 2012	Baird_Sparrow
9	Aston Martin Virage Convertible 2012	Baltimore_Oriole
10	Aston Martin Virage Coupe 2012	Bank_Swallow
11	Audi 100 Sedan 1994	Barn_Swallow
12	Audi 100 Wagon 1994	Bay_breasted_Warbler
13	Audi A5 Coupe 2012	Belted_Kingfisher
14	Audi R8 Coupe 2012	Bewick_Wren
15	Audi RS 4 Convertible 2008	Black_Tern
16	Audi S4 Sedan 2007	Black_and_white_Warbler
17	Audi S4 Sedan 2012	Black_billed_Cuckoo
18	Audi S5 Convertible 2012	Black_capped_Vireo
19	Audi S5 Coupe 2012	Black_footed_Albatross
20	Audi S6 Sedan 2011	Black_throated_Blue_Warbler
21	Audi TT Hatchback 2011	Black_throated_Sparrow
22	Audi TT RS Coupe 2012	Blue_Grosbeak
23	Audi TTS Coupe 2012	Blue_Jay
24	Audi V8 Sedan 1994	Blue_headed_Vireo
25	BMW 1 Series Convertible 2012	Blue_winged_Warbler
26	BMW 1 Series Coupe 2012	Boat_tailed_Grackle
27	BMW 3 Series Sedan 2012	Bobolink
28	BMW 3 Series Wagon 2012	Bohemian_Waxwing
29	BMW 6 Series Convertible 2007	Brandt_Cormorant
30	BMW ActiveHybrid 5 Sedan 2012	Brewer_Blackbird
31	BMW M3 Coupe 2012	Brewer_Sparrow
32	BMW M5 Sedan 2010	Bronzed_Cowbird
33	BMW M6 Convertible 2010	Brown_Creeper
34	BMW X3 SUV 2012	Brown_Pelican
35	BMW X5 SUV 2007	Brown_Thrasher
36	BMW X6 SUV 2012	Cactus_Wren
37	BMW Z4 Convertible 2012	California_Gull
38	Bentley Arnage Sedan 2009	Canada_Warbler
39	Bentley Continental Flying Spur Sedan 2007	Cape_Glossy_Starling

	Index	StanfordCars Class	CUB200 Class
378			
379	40	Bentley Continental GT Coupe 2007	Cape_May_Warbler
380	41	Bentley Continental GT Coupe 2012	Cardinal
381	42	Bentley Continental Supersports Conv. Convertible 2012	Carolina_Wren
382	43	Bentley Mulsanne Sedan 2011	Caspian_Tern
383	44	Bugatti Veyron 16.4 Convertible 2009	Cedar_Waxwing
384	45	Bugatti Veyron 16.4 Coupe 2009	Cerulean_Warbler
385	46	Buick Enclave SUV 2012	Chestnut_sided_Warbler
386	47	Buick Rainier SUV 2007	Chipping_Sparrow
387	48	Buick Regal GS 2012	Chuck_will_Widow
388	49	Buick Verano Sedan 2012	Clark_Nutcracker
389	50	Cadillac CTS-V Sedan 2012	Clay_colored_Sparrow
390	51	Cadillac Escalade EXT Crew Cab 2007	Cliff_Swallow
391	52	Cadillac SRX SUV 2012	Common_Raven
392	53	Chevrolet Avalanche Crew Cab 2012	Common_Tern
393	54	Chevrolet Camaro Convertible 2012	Common_Yellowthroat
394	55	Chevrolet Cobalt SS 2010	Crested_Auklet
395	56	Chevrolet Corvette Convertible 2012	Dark_eyed_Junco
396	57	Chevrolet Corvette Ron Fellows Edition Z06 2007	Downy_Woodpecker
397	58	Chevrolet Corvette ZR1 2012	Eared_Grebe
398	59	Chevrolet Express Cargo Van 2007	Eastern_Towhee
399	60	Chevrolet Express Van 2007	Elegant_Tern
400	61	Chevrolet HHR SS 2010	European_Goldfinch
401	62	Chevrolet Impala Sedan 2007	Evening_Grosbeak
402	63	Chevrolet Malibu Hybrid Sedan 2010	Field_Sparrow
403	64	Chevrolet Malibu Sedan 2007	Fish_Crow
404	65	Chevrolet Monte Carlo Coupe 2007	Florida_Jay
405	66	Chevrolet Silverado 1500 Classic Extended Cab 2007	Forsters_Tern
406	67	Chevrolet Silverado 1500 Extended Cab 2012	Fox_Sparrow
407	68	Chevrolet Silverado 1500 Hybrid Crew Cab 2012	Frigatebird
408	69	Chevrolet Silverado 1500 Regular Cab 2012	Gadwall
409	70	Chevrolet Silverado 2500HD Regular Cab 2012	Geococcyx
410	71	Chevrolet Sonic Sedan 2012	Glaucous_winged_Gull
411	72	Chevrolet Tahoe Hybrid SUV 2012	Golden_winged_Warbler
412	73	Chevrolet TrailBlazer SS 2009	Grasshopper_Sparrow
413	74	Chevrolet Traverse SUV 2012	Gray_Catbird
414	75	Chrysler 300 SRT-8 2010	Gray_Kingbird
415	76	Chrysler Aspen SUV 2009	Gray_crowned_Rosy_Finch
416	77	Chrysler Crossfire Convertible 2008	Great_Crested_Flycatcher
417	78	Chrysler PT Cruiser Convertible 2008	Great_Grey_Shrike
418	79	Chrysler Sebring Convertible 2010	Green_Jay
419	80	Chrysler Town and Country Minivan 2012	Green_Kingfisher
420	81	Daewoo Nubira Wagon 2002	Green_Violetear
421	82	Dodge Caliber Wagon 2007	Green_tailed_Towhee
422	83	Dodge Caliber Wagon 2012	Groove_billed_Ani
423	84	Dodge Caravan Minivan 1997	Harris_Sparrow
424	85	Dodge Challenger SRT8 2011	Heermann_Gull
425	86	Dodge Charger SRT-8 2009	Henslow_Sparrow
426	87	Dodge Charger Sedan 2012	Herring_Gull
427	88	Dodge Dakota Club Cab 2007	Hooded_Merganser
428	89	Dodge Dakota Crew Cab 2010	Hooded_Oriole
429	90	Dodge Durango SUV 2007	Hooded_Warbler
430	91	Dodge Durango SUV 2012	Horned_Grebe
431	92	Dodge Journey SUV 2012	Horned_Lark
	93	Dodge Magnum Wagon 2008	Horned_Puffin
	94	Dodge Ram Pickup 3500 Crew Cab 2010	House_Sparrow
	95	Dodge Ram Pickup 3500 Quad Cab 2009	House_Wren
	96	Dodge Sprinter Cargo Van 2009	Indigo_Bunting
	97	Eagle Talon Hatchback 1998	Ivory_Gull

	Index	StanfordCars Class	CUB200 Class
432			
433	98	FIAT 500 Abarth 2012	Kentucky_Warbler
434	99	FIAT 500 Convertible 2012	Laysan_Albatross
435	100	Ferrari 458 Italia Convertible 2012	Lazuli_Bunting
436	101	Ferrari 458 Italia Coupe 2012	Le_Conte_Sparrow
437	102	Ferrari California Convertible 2012	Least_Auklet
438	103	Ferrari FF Coupe 2012	Least_Flycatcher
439	104	Fisker Karma Sedan 2012	Least_Tern
440	105	Ford E-Series Wagon Van 2012	Lincoln_Sparrow
441	106	Ford Edge SUV 2012	Loggerhead_Shrike
442	107	Ford Expedition EL SUV 2009	Long_tailed_Jaeger
443	108	Ford F-150 Regular Cab 2007	Louisiana_Waterthrush
444	109	Ford F-150 Regular Cab 2012	Magnolia_Warbler
444	110	Ford F-450 Super Duty Crew Cab 2012	Mallard
445	111	Ford Fiesta Sedan 2012	Mangrove_Cuckoo
446	112	Ford Focus Sedan 2007	Marsh_Wren
447	113	Ford Freestar Minivan 2007	Mockingbird
448	114	Ford GT Coupe 2006	Mourning_Warbler
449	115	Ford Mustang Convertible 2007	Myrtle_Warbler
450	116	Ford Ranger SuperCab 2011	Nashville_Warbler
451	117	GMC Acadia SUV 2012	Nelson_Sharp_tailed_Sparrow
452	118	GMC Canyon Extended Cab 2012	Nighthawk
453	119	GMC Savana Van 2012	Northern_Flicker
454	120	GMC Terrain SUV 2012	Northern_Fulmar
454	121	GMC Yukon Hybrid SUV 2012	Northern_Waterthrush
455	122	Geo Metro Convertible 1993	Olive_sided_Flycatcher
456	123	HUMMER H2 SUT Crew Cab 2009	Orange_crowned_Warbler
457	124	HUMMER H3T Crew Cab 2010	Orchard_Oriole
458	125	Honda Accord Coupe 2012	Ovenbird
459	126	Honda Accord Sedan 2012	Pacific_Loon
460	127	Honda Odyssey Minivan 2007	Painted_Bunting
461	128	Honda Odyssey Minivan 2012	Palm_Warbler
462	129	Hyundai Accent Sedan 2012	Parakeet_Auklet
463	130	Hyundai Azera Sedan 2012	Pelagic_Cormorant
464	131	Hyundai Elantra Sedan 2007	Philadelphia_Vireo
464	132	Hyundai Elantra Touring Hatchback 2012	Pied_Kingfisher
465	133	Hyundai Genesis Sedan 2012	Pied_billed_Grebe
466	134	Hyundai Santa Fe SUV 2012	Pigeon_Guillemot
467	135	Hyundai Sonata Hybrid Sedan 2012	Pileated_Woodpecker
468	136	Hyundai Sonata Sedan 2012	Pine_Grosbeak
469	137	Hyundai Tucson SUV 2012	Pine_Warbler
470	138	Hyundai Veloster Hatchback 2012	Pomarine_Jaeger
471	139	Hyundai Veracruz SUV 2012	Prairie_Warbler
472	140	Infiniti G Coupe IPL 2012	Prothonotary_Warbler
473	141	Infiniti QX56 SUV 2011	Purple_Finch
474	142	Isuzu Ascender SUV 2008	Red_bellied_Woodpecker
475	143	Jaguar XK XKR 2012	Red_breasted_Merganser
475	144	Jeep Compass SUV 2012	Red_cockaded_Woodpecker
476	145	Jeep Grand Cherokee SUV 2012	Red_eyed_Vireo
477	146	Jeep Liberty SUV 2012	Red_faced_Cormorant
478	147	Jeep Patriot SUV 2012	Red_headed_Woodpecker
479	148	Jeep Wrangler SUV 2012	Red_legged_Kittiwake
480	149	Lamborghini Aventador Coupe 2012	Red_winged_Blackbird
481	150	Lamborghini Diablo Coupe 2001	Rhinoceros_Auklet
482	151	Lamborghini Gallardo LP 570-4 Superleggera 2012	Ring_billed_Gull
483	152	Lamborghini Reventon Coupe 2008	Ringed_Kingfisher
484	153	Land Rover LR2 SUV 2012	Rock_Wren
484	154	Land Rover Range Rover SUV 2012	Rose_breasted_Grosbeak
485	155	Lincoln Town Car Sedan 2011	Ruby_throated_Hummingbird

Index	StanfordCars Class	CUB200 Class
486		
487	156 MINI Cooper Roadster Convertible 2012	Rufous_Hummingbird
488	157 Maybach Landulet Convertible 2012	Rusty_Blackbird
489	158 Mazda Tribute SUV 2011	Sage_Thrasher
490	159 McLaren MP4-12C Coupe 2012	Savannah_Sparrow
491	160 Mercedes-Benz 300-Class Convertible 1993	Sayornis
492	161 Mercedes-Benz C-Class Sedan 2012	Scarlet_Tanager
493	162 Mercedes-Benz E-Class Sedan 2012	Scissor_tailed_Flycatcher
494	163 Mercedes-Benz S-Class Sedan 2012	Scott_Oriole
495	164 Mercedes-Benz SL-Class Coupe 2009	Seaside_Sparrow
496	165 Mercedes-Benz Sprinter Van 2012	Shiny_Cowbird
497	166 Mitsubishi Lancer Sedan 2012	Slaty_backed_Gull
498	167 Nissan 240SX Coupe 1998	Song_Sparrow
499	168 Nissan Juke Hatchback 2012	Sooty_Albatross
500	169 Nissan Leaf Hatchback 2012	Spotted_Catbird
501	170 Nissan NV Passenger Van 2012	Summer_Tanager
502	171 Plymouth Neon Coupe 1999	Swainson_Warbler
503	172 Porsche Panamera Sedan 2012	Tennessee_Warbler
504	173 Ram C-V Cargo Van Minivan 2012	Tree_Sparrow
505	174 Rolls-Royce Ghost Sedan 2012	Tree_Swallow
506	175 Rolls-Royce Phantom Drophead Coupe Convertible 2012	Tropical_Kingbird
507	176 Rolls-Royce Phantom Sedan 2012	Vermilion_Flycatcher
508	177 Scion xD Hatchback 2012	Vesper_Sparrow
509	178 Spyker C8 Convertible 2009	Warbling_Vireo
510	179 Spyker C8 Coupe 2009	Western_Grebe
511	180 Suzuki Aerio Sedan 2007	Western_Gull
512	181 Suzuki Kizashi Sedan 2012	Western_Meadowlark
513	182 Suzuki SX4 Hatchback 2012	Western_Wood_Pewee
514	183 Suzuki SX4 Sedan 2012	Whip_poor_Will
515	184 Tesla Model S Sedan 2012	White_Pelican
516	185 Toyota 4Runner SUV 2012	White_breasted_Kingfisher
517	186 Toyota Camry Sedan 2012	White_breasted_Nuthatch
518	187 Toyota Corolla Sedan 2012	White_crowned_Sparrow
519	188 Toyota Sequoia SUV 2012	White_eyed_Vireo
520	189 Volkswagen Beetle Hatchback 2012	White_necked_Raven
521	190 Volkswagen Golf Hatchback 1991	White_throated_Sparrow
522	191 Volkswagen Golf Hatchback 2012	Wilson_Warbler
523	192 Volvo 240 Sedan 1993	Winter_Wren
524	193 Volvo C30 Hatchback 2012	Worm_eating_Warbler
525	194 Volvo XC90 SUV 2007	Yellow_Warbler
526	195 smart fortwo Convertible 2012	Yellow_bellied_Flycatcher
527	196	Yellow_billed_Cuckoo
528	197	Yellow_breasted_Chat
529	198	Yellow_headed_Blackbird
530	199	Yellow_throated_Vireo

Table 7: **Dataset class indices.** We provide the class indices for OxfordPets, Caltech101, and Mini-ImageNet, which have 37, 101, and 100 classes, respectively.

Index	OxfordPets Class	Caltech101 Class	Mini-ImageNet Class
532			
533	0 Abyssinian	Faces	African_hunting_dog
534	1 Bengal	Faces_easy	Arctic_fox
535	2 Birman	Leopards	French_bulldog
536	3 Bombay	Motorbikes	Gordon_setter
537	4 British_Shorthair	accordion	Ibizan_hound
538	5 Egyptian_Mau	airplanes	Newfoundland
539	6 Maine_Coon	anchor	Saluki
	7 Persian	ant	Tibetan_mastiff

	Index	OxfordPets Class	Caltech101 Class	Mini-ImageNet Class
540				
541	8	Ragdoll	barrel	Walker_hound
542	9	Russian_Blue	bass	aircraft_carrier
543	10	Siamese	beaver	ant
544	11	Sphynx	binocular	ashcan
545	12	american_bulldog	bonsai	barrel
546	13	american_pit_bull_terrier	brain	beer_bottle
547	14	basset_hound	brontosaurus	black-footed_ferret
548	15	beagle	buddha	bolete
549	16	boxer	butterfly	bookshop
549	17	chihuahua	camera	boxer
550	18	english_cocker_spaniel	cannon	cannon
551	19	english_setter	car_side	carousel
552	20	german_shorthaired	ceiling_fan	carton
553	21	great_pyrenees	cellphone	catamaran
554	22	havanese	chair	chime
555	23	japanese_chin	chandelier	cliff
556	24	keeshond	cougar_body	clog
557	25	leonberger	cougar_face	cocktail_shaker
558	26	miniature_pinscher	crab	combination_lock
559	27	newfoundland	crayfish	consomme
560	28	pomeranian	crocodile	coral_reef
560	29	pug	crocodile_head	crate
561	30	saint_bernard	cup	cuirass
562	31	samoyed	dalmatian	dalmatian
563	32	scottish_terrier	dollar_bill	dishrag
564	33	shiba_inu	dolphin	dome
565	34	staffordshire_bull_terrier	dragonfly	dugong
566	35	wheaten_terrier	electric_guitar	electric_guitar
567	36	yorkshire_terrier	elephant	file
568	37		emu	fire_screen
569	38		euphonium	frying_pan
570	39		ewer	garbage_truck
571	40		ferry	golden_retriever
572	41		flamingo	goose
572	42		flamingo_head	green_mamba
573	43		garfield	hair_slide
574	44		gerenuk	harvestman
575	45		gramophone	holster
576	46		grand_piano	horizontal_bar
577	47		hawksbill	hotdog
578	48		headphone	hourglass
579	49		hedgehog	house_finch
580	50		helicopter	iPod
581	51		ibis	jellyfish
582	52		inline_skate	king_crab
583	53		joshua_tree	komondor
584	54		kangaroo	ladybug
585	55		ketch	lion
585	56		lamp	lipstick
586	57		laptop	malamute
587	58		llama	meerkat
588	59		lobster	miniature_poodle
589	60		lotus	miniskirt
590	61		mandolin	missile
591	62		mayfly	mixing_bowl
592	63		menorah	nematode
592	64		metronome	oboe
593	65		minaret	orange

594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647

Index	OxfordPets Class	Caltech101 Class	Mini-ImageNet Class
66		nautilus	organ
67		octopus	parallel_bars
68		okapi	pencil_box
69		pagoda	photocopier
70		panda	poncho
71		pigeon	prayer_rug
72		pizza	reel
73		platypus	rhinoceros_beetle
74		pyramid	robin
75		revolver	rock_beauty
76		rhino	school_bus
77		rooster	scoreboard
78		saxophone	slot
79		schooner	snorkel
80		scissors	solar_dish
81		scorpion	spider_web
82		sea_horse	spike
83		snoopy	stage
84		soccer_ball	street_sign
85		stapler	tank
86		starfish	theater_curtain
87		stegosaurus	three-toed_sloth
88		stop_sign	tile_roof
89		strawberry	tobacco_shop
90		sunflower	toucan
91		tick	triceratops
92		trilobite	trifle
93		umbrella	unicycle
94		watch	upright_piano
95		water_lilly	vase
96		wheelchair	white_wolf
97		wild_cat	wok
98		windsor_chair	worm_fence
99		wrench	yawl
100		yin_yang	

648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701

REFERENCES

- Mircea Cimpoi, Subhransu Maji, Iasonas Kokkinos, Sammy Mohamed, and Andrea Vedaldi. Describing textures in the wild. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 3606–3613, 2014.
- Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*, pp. 248–255. Ieee, 2009.
- Li Fei-Fei, Rob Fergus, and Pietro Perona. Learning generative visual models from few training examples: An incremental bayesian approach tested on 101 object categories. In *2004 conference on computer vision and pattern recognition workshop*, pp. 178–178. IEEE, 2004.
- Jonathan Krause, Michael Stark, Jia Deng, and Li Fei-Fei. 3d object representations for fine-grained categorization. In *Proceedings of the IEEE international conference on computer vision workshops*, pp. 554–561, 2013.
- Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick. Microsoft coco: Common objects in context. In *Computer Vision—ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13*, pp. 740–755. Springer, 2014.
- Subhransu Maji, Esa Rahtu, Juho Kannala, Matthew Blaschko, and Andrea Vedaldi. Fine-grained visual classification of aircraft. *arXiv preprint arXiv:1306.5151*, 2013.
- Maria-Elena Nilsback and Andrew Zisserman. Automated flower classification over a large number of classes. In *2008 Sixth Indian conference on computer vision, graphics & image processing*, pp. 722–729. IEEE, 2008.
- Omkar M Parkhi, Andrea Vedaldi, Andrew Zisserman, and CV Jawahar. Cats and dogs. In *2012 IEEE conference on computer vision and pattern recognition*, pp. 3498–3505. IEEE, 2012.
- Sachin Ravi and Hugo Larochelle. Optimization as a model for few-shot learning. In *International conference on learning representations*, 2016.
- Catherine Wah, Steve Branson, Peter Welinder, Pietro Perona, and Serge Belongie. The caltech-ucsd birds-200-2011 dataset. 2011.

702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755

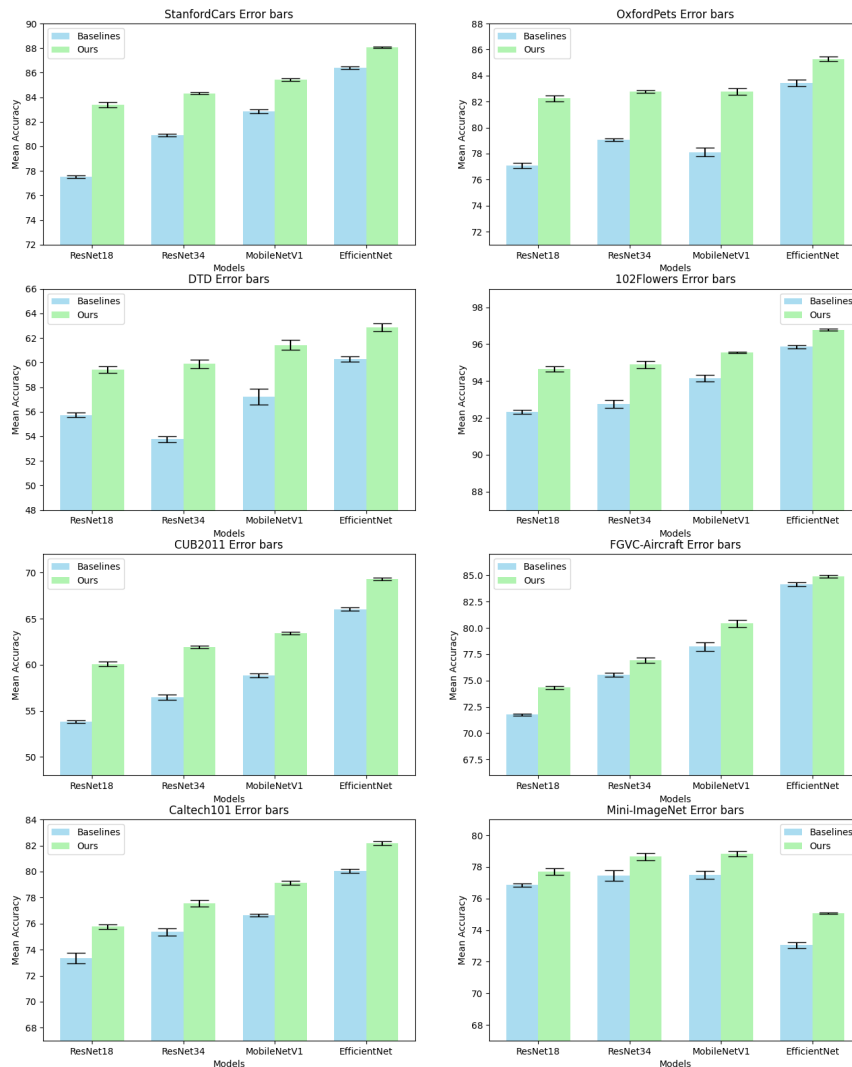


Figure 2: **Error bars of SOTA experimental results on fine-grained and coarse-grained datasets.** We run each experiment three times and report the average results. This results represent the standard deviation of the performance across multiple runs for both fine-grained and coarse-grained datasets, reflecting the variability and stability of the SOTA experiment results.

756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809

1. Does the car belong to the high-end luxury category (like Bugatti, Bentley, etc.)?
2. Is the car's make year post-2010?
3. Is the car equipped with a V8 engine?
4. Is the car's model a hatchback?
5. Is the car model a SUV?
6. Is the car a diesel-powered model?
7. Is the car a model of Chevrolet brand?
8. Does the car have a convertible roof?
9. Is the car a sports coupe model?
10. Does the car belong to the sedan category?
11. Does the picture depict a sports version of a typical car model (like Audi RS, Aston Martin V8 Vantage etc.)?
12. Is the car a hybrid vehicle?
13. Is the car from the minivan category?
14. Does the car have a noticeable rear spoiler?
15. Is the car model from the smaller Compact Class?
16. Is the make of the car BMW?
17. Is the car part of the Ford family?
18. Have the car images been taken after 2007?
19. Is the car a part of the Italian luxury car brands (like Ferrari, Lamborghini)?
20. Was the car model made in the V12 engine series?
21. Does the car have scissor doors?
22. Does the car have distinctive gull-wing doors?
23. Does the car have a rear engine layout?
24. Is the car an off-road vehicle or designed for rugged terrain usage?
25. Does the car feature a dual exhaust system?
26. Is the car a roadster model?
27. Is the car equipped with side skirts?
28. Is the car a station wagon?
29. Does the car feature a distinctive front grille with vertical slats?
30. Is the car from the Japanese automaker, Honda?
31. Is the car a part of the electric car category (like Tesla Model S, Chevrolet Bolt, etc.)?
32. Does the car have a long wheelbase version?
33. Does the car have a soft-top roof?
34. Is the car a coupe with two doors?
35. Is the car from a Korean manufacturer (like Hyundai, Kia)?
36. Does the car have a distinctive round headlight design?
37. Does the car belong to the pickup truck category?
38. Does the car have a distinctive boxy shape?
39. Is the car a 4-door model?
40. Is the car equipped with a turbocharger?
41. Is the car a plug-in hybrid?
42. Does the car feature a panoramic sunroof?
43. Is the car a muscle car (like Dodge Challenger, Chevrolet Camaro)?
44. Does the car have a noticeable hood scoop?
45. Does the car have all-wheel drive (AWD)?
46. Is the car a vintage model made before 2000?
47. Does the car have a prominent air intake on the front bumper?
48. Does the car have a distinctive rear diffuser?
49. Is the car from an American manufacturer?
50. Is the car a convertible with a hardtop?

Table 8: **Multi-aspect questions generated by GPT-4o for the StanfordCars dataset.**

810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863

1. Does the animal have a flat face?
2. Does the animal display a prominent ruff around the neck?
3. Are the ears of the breed long and floppy?
4. Is the animal's fur long?
5. Does the animal have a robust and muscular build?
6. Does the breed have a compact and muscular build?
7. Does the animal have long drooping ears?
8. Does the animal have distinctive facial markings?
9. Does the animal have striking blue eyes?
10. Does the animal have a brachycephalic (shortened head) skull?
11. Does the animal have a double coat?
12. Is the animal's body unusually slender and tall?
13. Is the breed's coat spotted or dappled?
14. Is the tail of the animal bushy or feathered?
15. Does the animal have webbed feet?
16. Does the animal have a short, stubby nose?
17. Does the animal have floppy ears?
18. Is the animal's coat curly or wavy?
19. Is the fur of the animal curly or wavy?
20. Does the animal have hairless skin?
21. Does the breed have a plumed tail?
22. Does the animal have a long, flowing coat?
23. Is the animal small-sized, typically less than 10 pounds?
24. Does the breed have a square-shaped body?
25. Is the animal typically solid-colored?
26. Is the animal predominantly white in color?
27. Is the breed characterized by a high-set tail?
28. Does the breed have a short snout?
29. Is the breed's tail bushy or fluffy?
30. Does the animal have a pronounced underbite?
31. Does the animal have an unusually squarish or boxy muzzle?
32. Is the fur patterned with spots or stripes?
33. Does the breed have a pointed muzzle?
34. Does the animal have a characteristically flat or pushed-in face with large, round eyes?
35. Does the animal have large, round eyes?
36. Is the animal's coat silky to the touch?
37. Does the breed have a long and slender tail?
38. Is the breed's coat rough or wiry?
39. Is the animal known for having a sleek and shiny coat?
40. Is the breed known for having a slender body?
41. Is the breed known for its distinctive coloration or pattern?
42. Does the breed have a broad chest?
43. Does the breed have a distinctive ruff or collar of fur around the neck?
44. Is the animal predominantly black in color?
45. Does the breed have large, bat-like ears?
46. Is the breed's coat short and dense?
47. Does the breed have a docked or naturally short tail?
48. Does the breed have small, pointed ears?
49. Is the animal known for having a lion-like appearance?
50. Is the animal's coat thick and woolly?

Table 9: Multi-aspect questions generated by GPT-4o for the OxfordPets dataset.

864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917

1. Are there noticeable cracks or fissures?
2. Does the texture have a scaly or reptilian appearance?
3. Does the texture feature honeycomb-like hexagonal shapes?
4. Does the texture look like a net or web?
5. Are there regular, grid-like patterns?
6. Are there noticeable swirls or spiral patterns?
7. Is the texture characterized by a dotted or spotted pattern?
8. Are there visible grooves or indentations?
9. Are there waffle-like grid patterns on the texture?
10. Does the texture have a marbled appearance with blended colors?
11. Is the texture perforated or has holes?
12. Is the pattern composed of crisscrossing lines?
13. Are there distinct, irregular blotches?
14. Does the texture have a checkered or chequered pattern?
15. Are there fibrous or thread-like elements visible?
16. Does the texture have a veined appearance?
17. Does the texture have a crystalline or gem-like appearance?
18. Are there raised, bumpy areas on the texture?
19. Does the texture appear braided with intertwining strands?
20. Is the texture wrinkled or creased?
21. Is the texture characterized by fine, lace-like details?
22. Does the texture have a smeared or smudged appearance?
23. Is the texture smeared with streaks or smears?
24. Are there visible stains or discolorations on the texture?
25. Does the texture feature pleated or folded sections?
26. Is the pattern composed of zigzag lines?
27. Does the texture have a sprinkled or speckled look?
28. Are there any interwoven or braided elements in the texture?
29. Is the texture banded with stripes of varying widths?
30. Is the texture composed of overlapping or interlaced elements?
31. Is the texture flecked with small, random spots?
32. Are there frilly or ruffled edges in the texture?
33. Does the texture have a porous or sponge-like look?
34. Is the texture marked by potholes or deep indentations?
35. Is the texture covered with polka dots?
36. Are there visible knitted or crocheted patterns?
37. Does the texture have a paisley or teardrop-shaped pattern?
38. Are there pitted or dimpled areas on the texture?
39. Does the texture have a stratified or layered appearance?
40. Does the texture feature stratified layers or bands?
41. Are there visible bubbles or circular shapes?
42. Is the texture cobwebbed with thin, thread-like lines?
43. Does the texture resemble fabric or woven material?
44. Is the texture marked by crosshatched lines?
45. Does the texture have parallel lines?
46. Are there noticeable stained or dirty areas?
47. Does the texture have a woven or interlaced look?
48. Is the texture swirly with swirling patterns?
49. Are there noticeable wrinkles or creases?
50. Does the texture have a zigzagged pattern?

Table 10: Multi-aspect questions generated by GPT-4o for the DTD dataset.

918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971

1. Does the flower have multiple petals arranged in a symmetrical pattern?
2. Does the flower have heart-shaped petals?
3. Does the flower have a prominent central disk surrounded by petals?
4. Are there multiple small flowers arranged in a cluster?
5. Is the flower predominantly blue or purple?
6. Does the flower exhibit a gradient of colors?
7. Does the flower have a spiky or thistle-like appearance?
8. Does the flower have a large, singular bloom?
9. Is the primary colors of the flower yellow?
10. Are the petals long and narrow, resembling a lily?
11. Does the flower have a tubular shape?
12. Is the flower predominantly red?
13. Are the petals arranged in layers or rows?
14. Are the petals overlapping?
15. Are the petals fringed or ruffled?
16. Does the flower grow in a cluster on a single stem?
17. Are the petals shaped like a star or have pointed tips?
18. Does the flower have a distinct, pronounced lip or 'tongue' petal?
19. Does the flower have a spurred petal or elongated appendage?
20. Are the petals arranged in a spiral pattern?
21. Does the flower have a strong fragrance?
22. Is the flower predominantly pink?
23. Does the flower have a bell or trumpet shape?
24. Does the flower have a daisy-like appearance?
25. Does the flower have a cup-shaped structure?
26. Does the flower have hairy or fuzzy petals?
27. Are the petals thin and delicate?
28. Are the petals bi-colored?
29. Are the petals flat and wide?
30. Does the flower have a central crown or corona?
31. Is the flower predominantly white?
32. Are the petals veined or patterned?
33. Are the petals rounded at the tips?
34. Does the flower have strap-like petals?
35. Are the petals twisted or curled?
36. Does the flower have a single petal?
37. Does the flower have a dome-shaped appearance?
38. Is the flower predominantly orange?
39. Does the flower have a flattened top?
40. Are the petals spoon-shaped?
41. Are the petals translucent or semi-transparent?
42. Does the flower have prominent stamens?
43. Is the flower predominantly green?
44. Does the flower have a geometric pattern on its petals?
45. Does the flower have a papery texture?
46. Are the petals serrated or jagged?
47. Are the petals clustered tightly together?
48. Is the flower predominantly violet?
49. Does the flower have drooping petals?
50. Are the petals reflexed or bent backward?

Table 11: Multi-aspect questions generated by GPT-4o for the 102Flowers dataset.

972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025

1. Does the bird have a curved beak?
2. Is the bird's beak long and pointed?
3. Is the bird predominantly blue?
4. Is the bird's primary habitat coastal areas?
5. Is the bird primarily found in water habitats?
6. Is the bird's beak hooked?
7. Is the bird's underside orange?
8. Does the bird have a long neck?
9. Is the bird's plumage mostly white?
10. Is the bird predominantly found in forests?
11. Does the bird have a thin, needle-like beak?
12. Does the bird have a crest on its head?
13. Does the bird have iridescent feathers?
14. Is the bird's beak short and thick?
15. Is the bird's beak conical?
16. Is the bird's plumage predominantly brown?
17. Does the bird have a fan-shaped tail?
18. Does the bird have a black and white striped pattern?
19. Does the bird have a red patch on its wings?
20. Is the bird's breast yellow?
21. Does the bird have a white eye stripe?
22. Does the bird have webbed feet?
23. Does the bird have a notched tail?
24. Is the bird's chest streaked?
25. Does the bird have a ring around its neck?
26. Does the bird have a black cap on its head?
27. Does the bird have a speckled breast?
28. Does the bird have long legs?
29. Is the bird's back green?
30. Does the bird have a black tail?
31. Does the bird have a mask-like pattern on its face?
32. Does the bird have a prominent eye ring?
33. Is the bird's tail short and square?
34. Does the bird have spots on its wings?
35. Is the bird's belly white?
36. Does the bird have a distinctive call that includes trills?
37. Does the bird have a yellow belly?
38. Is the bird's tail forked?
39. Is the bird's head and back grey?
40. Is the bird's wingspan larger than 12 inches?
41. Does the bird have a barred tail?
42. Is the bird's chest red?
43. Does the bird have a blue throat patch?
44. Does the bird have a bright orange beak?
45. Is the bird's head black?
46. Is the bird's beak straight?
47. Is the bird predominantly found in open grasslands?
48. Is the bird larger than a sparrow?
49. Does the bird have a yellow stripe on its wings?
50. Is the bird primarily insectivorous?

Table 12: Multi-aspect questions generated by GPT-4o for the CUB200 dataset.

1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079

1. Is this aircraft a turboprop model?
2. Does this aircraft have four engines?
3. Does this aircraft have a high-wing design?
4. Does this aircraft have two engines?
5. Is this a single-engine aircraft?
6. Is this aircraft used primarily for military purposes?
7. Is this aircraft a trijet (three engines)?
8. Does this aircraft feature a swept-wing design?
9. Is this aircraft a wide-body model?
10. Does the aircraft have propellers instead of jet engines?
11. Does the aircraft feature a T-tail design?
12. Does this aircraft have retractable landing gear?
13. Is this aircraft primarily used for cargo transportation?
14. Does this aircraft have an open cockpit?
15. Is the aircraft primarily used for commercial passenger flights?
16. Does this aircraft have a twin-jet engine configuration?
17. Is this aircraft primarily used for private or corporate purposes?
18. Does the aircraft have a delta wing configuration?
19. Does this aircraft have a single vertical stabilizer?
20. Is this aircraft a supersonic jet?
21. Is this aircraft used primarily for short regional flights?
22. Does this aircraft have an all-metal body?
23. Does this aircraft have winglets?
24. Does the aircraft have a long-range flight capacity?
25. Does this aircraft have a radial engine?
26. Does this aircraft have a tailwheel landing gear configuration?
27. Is this aircraft a high-performance jet?
28. Does this aircraft feature a pressurized cabin?
29. Does this aircraft have a twin-boom tail design?
30. Does this aircraft have a glass cockpit?
31. Does the aircraft feature swept-back wings?
32. Does this aircraft have a distinctive nose design?
33. Does this aircraft have a forward-swept wing design?
34. Does this aircraft have an all-composite body structure?
35. Does this aircraft have rear-mounted engines?
36. Does this aircraft have a tricycle landing gear configuration?
37. Does this aircraft have variable-sweep wings?
38. Does this aircraft have a high-wing design?
39. Does this aircraft have a distinctive humpback design?
40. Does this aircraft have a high bypass ratio engine?
41. Is this aircraft primarily designed for long-haul flights?
42. Is this aircraft often used for regional transportation?
43. Is this aircraft a narrow-body model?
44. Does the aircraft have turbojet engines?
45. Does this aircraft have twin tail fins?
46. Does this aircraft have a straight-wing design?
47. Is the aircraft designed for short takeoff and landing (STOL)?
48. Does the aircraft feature fixed landing gear?
49. Is this aircraft designed to operate from aircraft carriers?
50. Is the aircraft known for its high maneuverability?

Table 13: Multi-aspect questions generated by GPT-4o for the FGVC-Aircraft dataset.

1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133

1. Does the object have a recognizable face?
2. Does the object have wings?
3. Is the object known for its ability to fly?
4. Does the object have a screen and interface for digital interaction?
5. Does the object have limbs and a recognizable head/body structure?
6. Is the object known for its speed or ability to move quickly?
7. Does the object have wheels and an enclosed space for passengers?
8. Does the object have multiple legs?
9. Does the object have a prominent trunk or elongated nose?
10. Does the object have fur or hair?
11. Does the object have claws or pincers?
12. Is the object typically found in water or aquatic environments?
13. Is the object known for its ability to cut or pierce?
14. Does the object have a blade or sharp edge for cutting?
15. Does the object have feathers?
16. Is the object a type of vehicle used for transportation?
17. Is the object a type of instrument used to produce sound?
18. Does the object have keys or buttons for producing musical notes?
19. Does the object have a flat surface for placing items on?
20. Is the object used for capturing images or videos?
21. Is the object commonly associated with human activities or use?
22. Is the object typically found outdoors in a natural environment?
23. Does the object have a distinct shape?
24. Is the object commonly found in a garden or botanical setting?
25. Is the object primarily composed of organic materials?
26. Does the object have a shell or hard outer covering?
27. Is the object likely to be found in a domestic setting (home, kitchen)?
28. Is the object typically seen in a kitchen setting?
29. Does the object have any moving parts or mechanisms?
30. Does the object have a repeating pattern or design on its surface?
31. Is the object a type of plant?
32. Does the object have scales or a rough texture?
33. Does the object have an elongated neck?
34. Does the object have a circular or rounded shape?
35. Does the object have a cylindrical shape?
36. Does the object have a distinctive pattern on its surface, like spots or stripes?
37. Does the object have a clear, defined purpose or function?
38. Does this object occupy a large part of the image?
39. Is the object likely part of a larger system or assembly (e.g., part of a car)?
40. Is the object's primary purpose for entertainment or recreation?
41. Is the object typically used for writing or drawing?
42. Is the object often associated with religious or spiritual practices?
43. Is the object often used in sporting activities?
44. Is the object an example of marine life?
45. Is the object a piece of furniture used for seating?
46. Is the dominant color of the object a warm color (red, orange, yellow)?
47. Is the dominant color of the object a cool color (blue, skyblue, gray)?
48. Does the object appear to be handheld or designed for human interaction?
49. Is the object's shape primarily geometric (circles, squares, etc.)?
50. Is the object known for producing light?

Table 14: Multi-aspect questions generated by GPT-4o for the Caltech101 dataset.

1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187

1. Is the animal in the image a type of bird?
2. Is a creature known for its colorful appearance depicted in the image?
3. Does the image show a vehicle with wheels?
4. Does the image feature a musical instrument?
5. Is there something used for communication in the image?
6. Does the image contain a dog breed with long ears?
7. Is a creature with a distinct mane shown in the image?
8. Is the object in the image made of glass?
9. Is the object in the image typically found in a bathroom?
10. Is a weapon depicted in the image?
11. Is something used in artistic creation featured in the image?
12. Is a type of mushroom shown in the image?
13. Is something used for navigation present in the image?
14. Is something used for measurement present in the image?
15. Is something used for entertainment featured in the image?
16. Is a primarily nocturnal creature depicted in the image?
17. Is a mammal known for swimming depicted in the image?
18. Is there a clothing accessory in the image?
19. Can an amphibian be seen in the image?
20. Is the object shown typically found in a kitchen?
21. Does the image contain an animal with stripes?
22. Is the object depicted primarily used for transportation?
23. Does the image show an object used for cooling?
24. Is there an item associated with food preparation in the image?
25. Is a venomous creature shown in the image?
26. Can a type of beetle be seen in the image?
27. Is the object in the image typically found in water?
28. Is the depicted animal a type of cat?
29. Is there a vehicle without wheels in the image?
30. Is something typically used in a garden present in the image?
31. Can a large marine vessel be seen in the image?
32. Does the image feature an object used in sports?
33. Does the image contain a type of marine life?
34. Is a wild cat depicted in the image?
35. Is something used for timekeeping present in the image?
36. Is the depicted animal a type of amphibian?
37. Is something used for personal grooming in the image?
38. Is a herding dog breed shown in the image?
39. Can an insect with wings be seen in the image?
40. Is the animal in the image known for its speed?
41. Does the image contain a type of fruit?
42. Is the depicted animal a type of reptile?
43. Is something used in photography shown in the image?
44. Is something in the image commonly used for storage?
45. Does the image feature an object used in construction?
46. Is a rodent depicted in the image?
47. Is something typically found in an office present in the image?
48. Does the image contain an animal with a shell?
49. Is there protective gear depicted in the image?
50. Is a creature known for its strength depicted in the image?

Table 15: Multi-aspect questions generated by GPT-4o for the Mini-ImageNet dataset.

1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241

1. Is the car color black?
2. Is the car a convertible?
3. Is the car a sedan?
4. Is the car from the year 2012?
5. Is the car from the make Acura?
6. Is the car from the make Audi?
7. Is the car from the make BMW?
8. Is the car from the make Chevrolet?
9. Is the car from the make Dodge?
10. Is the car from the make Ferrari?
11. Is the car from the make Ford?
12. Is the car from the make Honda?
13. Is the car from the make Hyundai?
14. Is the car from the make Jeep?
15. Is the car from the make Lamborghini?
16. Is the car from the make Mercedes-Benz?
17. Is the car from the make Nissan?
18. Is the car from the make Porsche?
19. Is the car from the make Rolls-Royce?
20. Is the car from the make Toyota?
21. Is the car from the make Volkswagen?
22. Is the car a coupe?
23. Is the car an SUV?
24. Is the car a hatchback?
25. Is the car a wagon?
26. Is the car a hybrid?
27. Is the car a van?
28. Is the car a minivan?
29. Is the car a crew cab?
30. Is the car a regular cab?
31. Is the car a quad cab?
32. Is the car a club cab?
33. Is the car from the luxury category?
34. Is the car from the sports category?
35. Is the car from the economy category?
36. Is the car from the midsize category?
37. Is the car from the full-size category?
38. Is the car a high-performance model?
39. Is the car a low-performance model?
40. Is the car a high-end model?
41. Is the car a budget-friendly model?
42. Is the car a classic model?
43. Is the car a modern model?
44. Is the car a luxury sports car?
45. Is the car a sedan with a sunroof?
46. Is the car a coupe with a spoiler?
47. Is the car a convertible with a soft top?
48. Is the car a hatchback with a rear spoiler?
49. Is the car a wagon with roof racks?
50. Is the car a van with tinted windows?

Table 16: **Multi-aspect questions generated by GPT-3.5-turbo for the StanfordCars dataset.**

1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295

1. Does the breed have a short coat?
2. Does the breed have a long coat?
3. Are the ears of the breed floppy?
4. Are the ears of the breed erect?
5. Does the breed have a solid-colored coat?
6. Does the breed have a spotted coat pattern?
7. Is the breed known for its distinctive facial markings?
8. Is the breed large in size?
9. Is the breed small in size?
10. Does the breed have a curly tail?
11. Does the breed have a bushy tail?
12. Is the breed known for its playful nature?
13. Is the breed known for being affectionate?
14. Does the breed have a brachycephalic (short-nosed) face?
15. Is the coat of the breed fluffy?
16. Does the breed have a stocky build?
17. Is the breed known for its intelligence?
18. Is the breed known for its hunting abilities?
19. Is the breed known for its vocal nature?
20. Does the breed have a specific color pattern unique to its breed?
21. Does the breed have a distinct pattern on its face?
22. Does the breed have a breed-specific tail shape?
23. Is the breed known for its expressive eyes?
24. Does the breed have a muscular build?
25. Does the breed have a sleek and shiny coat?
26. Is the breed known for its agility?
27. Does the breed have a fluffy mane or collar?
28. Is the breed known for its endurance or stamina?
29. Does the breed have a double coat?
30. Is the breed known for its calm temperament?
31. Does the breed have a distinctive vocalization?
32. Is the breed known for its protective instincts?
33. Does the breed have prominent whiskers?
34. Does the breed have a distinctive head shape?
35. Is the breed known for its high energy levels?
36. Does the breed have a short, stubby nose?
37. Is the breed known for its unique tail carriage?
38. Does the breed have a sleek and elegant posture?
39. Is the breed known for its friendly disposition?
40. Does the breed have a sleek and slender build?
41. Is the breed known for its independent nature?
42. Does the breed have a luxurious coat texture?
43. Is the breed known for its social nature?
44. Does the breed have a silky or velvety coat?
45. Is the breed known for its athletic abilities?
46. Does the breed have distinctive facial hair or markings?
47. Is the breed known for its guarding instincts?
48. Does the breed have a thick, protective coat?
49. Is the breed known for its clownish behavior?
50. Does the breed have a unique tail length relative to its body size?

Table 17: Multi-aspect questions generated by GPT-3.5-turbo for the OxfordPets dataset.

1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349

1. Does the object have wheels?
2. Is the object a type of musical instrument?
3. Does the object have wings?
4. Is the object commonly found in water?
5. Does the object have fur?
6. Is the object commonly used for transportation?
7. Does the object have a shell?
8. Is the object typically found in a household setting?
9. Is the object typically found in nature?
10. Does the object have a long neck?
11. Is the object typically found in an office environment?
12. Does the object have scales?
13. Is the object a type of bird?
14. Does the object have antennas?
15. Does the object have claws?
16. Is the object commonly used for entertainment?
17. Does the object have a tail?
18. Is the object a type of plant?
19. Does the object have a sharp beak?
20. Is the object typically used for sports?
21. Does the object have multiple legs?
22. Is the object typically found in a museum?
23. Does the object have a curved shape?
24. Is the object typically used for cooking?
25. Does the object have a distinctive color pattern?
26. Is the object a type of reptile?
27. Does the object have a smooth texture?
28. Is the object typically found in the sky?
29. Does the object have horns?
30. Is the object typically used for communication?
31. Does the object have a protective shell?
32. Is the object typically found in tropical regions?
33. Does the object have a distinctive smell?
34. Is the object commonly found in urban environments?
35. Does the object have a long tail?
36. Is the object typically used for relaxation?
37. Does the object have a shiny surface?
38. Is the object typically found in cold climates?
39. Does the object have a round shape?
40. Is the object commonly associated with music?
41. Does the object have stripes?
42. Is the object typically found near water?
43. Does the object have a unique pattern?
44. Is the object typically found in forests?
45. Does the object have large ears?
46. Is the object typically found in a desert environment?
47. Does the object have a broad head?
48. Is the object commonly found on farms?
49. Does the object have a pointy nose?
50. Is the object typically used for navigation?

Table 18: Multi-aspect questions generated by GPT-3.5-turbo for the Caltech101 dataset.

1350
 1351
 1352
 1353
 1354
 1355
 1356
 1357
 1358
 1359
 1360
 1361
 1362
 1363
 1364
 1365
 1366
 1367
 1368
 1369
 1370
 1371
 1372
 1373
 1374
 1375
 1376
 1377
 1378
 1379
 1380
 1381
 1382
 1383
 1384
 1385
 1386
 1387
 1388
 1389
 1390
 1391
 1392
 1393
 1394
 1395
 1396
 1397
 1398
 1399
 1400
 1401
 1402
 1403

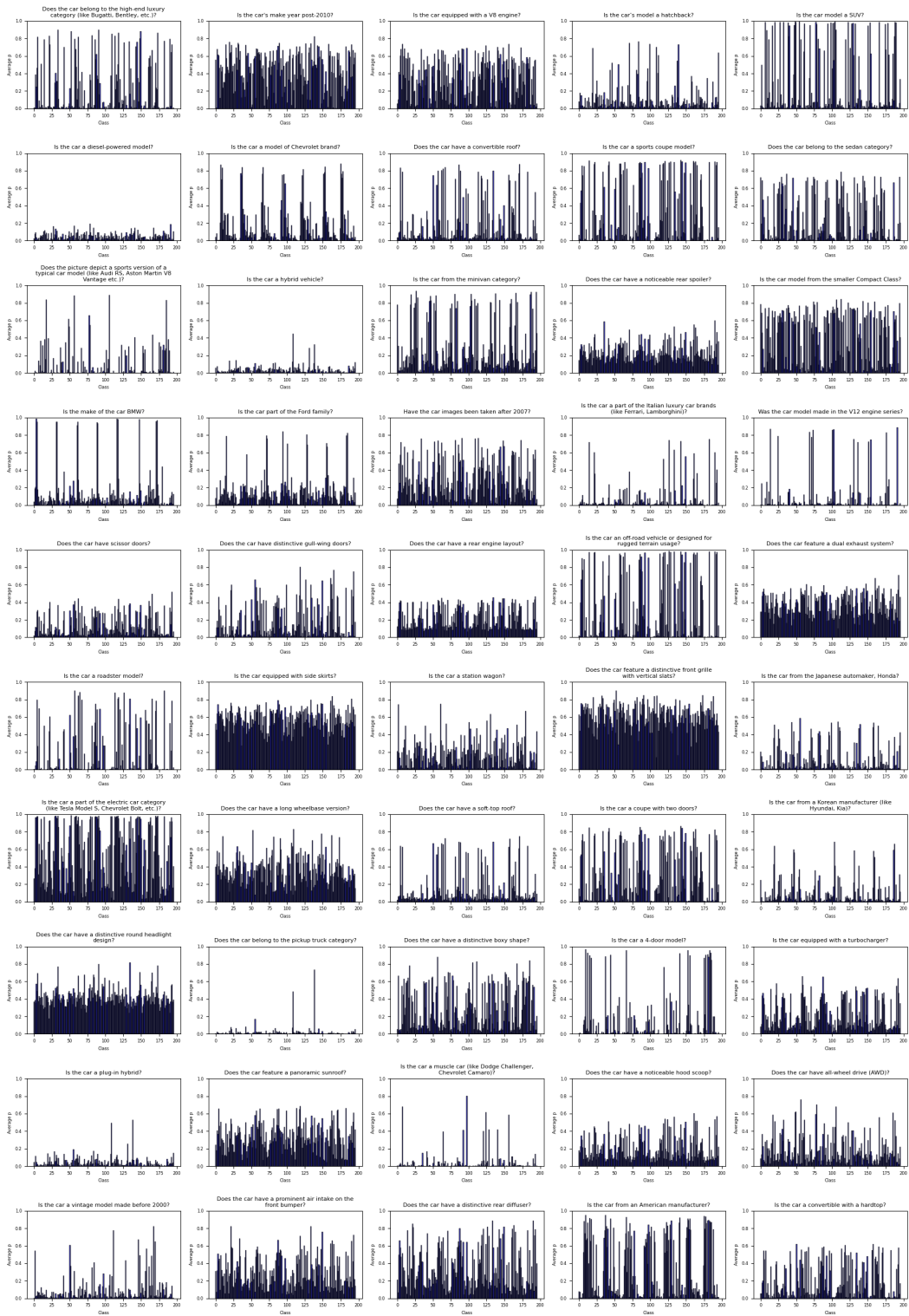


Figure 3: Visualization of the average logit distribution for StanfordCars.

1404
 1405
 1406
 1407
 1408
 1409
 1410
 1411
 1412
 1413
 1414
 1415
 1416
 1417
 1418
 1419
 1420
 1421
 1422
 1423
 1424
 1425
 1426
 1427
 1428
 1429
 1430
 1431
 1432
 1433
 1434
 1435
 1436
 1437
 1438
 1439
 1440
 1441
 1442
 1443
 1444
 1445
 1446
 1447
 1448
 1449
 1450
 1451
 1452
 1453
 1454
 1455
 1456
 1457



Figure 4: Visualization of the average logit distribution for the multi-aspect of the OxfordPets.

1458
 1459
 1460
 1461
 1462
 1463
 1464
 1465
 1466
 1467
 1468
 1469
 1470
 1471
 1472
 1473
 1474
 1475
 1476
 1477
 1478
 1479
 1480
 1481
 1482
 1483
 1484
 1485
 1486
 1487
 1488
 1489
 1490
 1491
 1492
 1493
 1494
 1495
 1496
 1497
 1498
 1499
 1500
 1501
 1502
 1503
 1504
 1505
 1506
 1507
 1508
 1509
 1510
 1511



Figure 5: Visualization of the average logit distribution for the multi-aspect of the DTD.

1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565

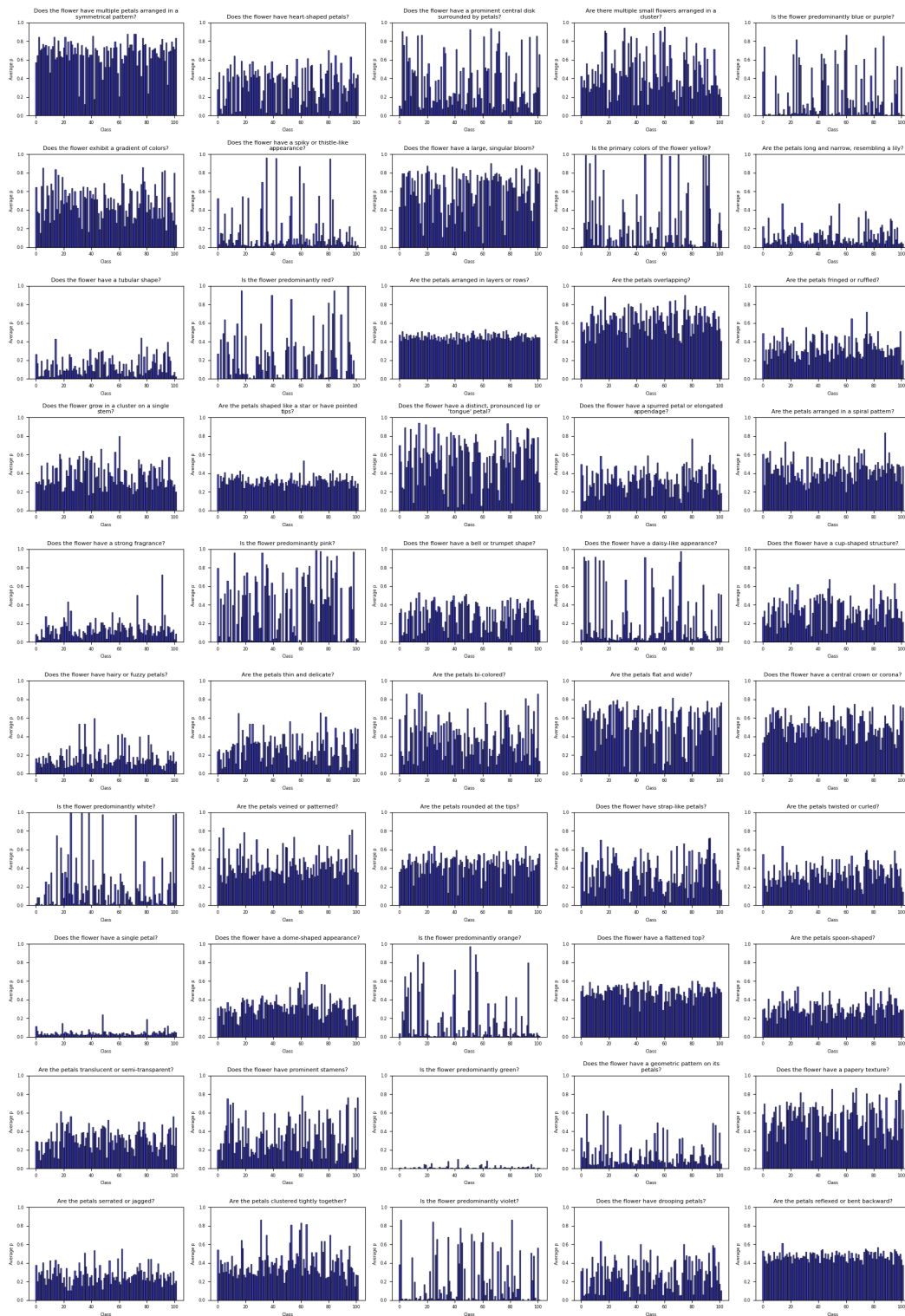


Figure 6: Visualization of the average logit distribution for the multi-aspect of the 102Flowers.

1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619

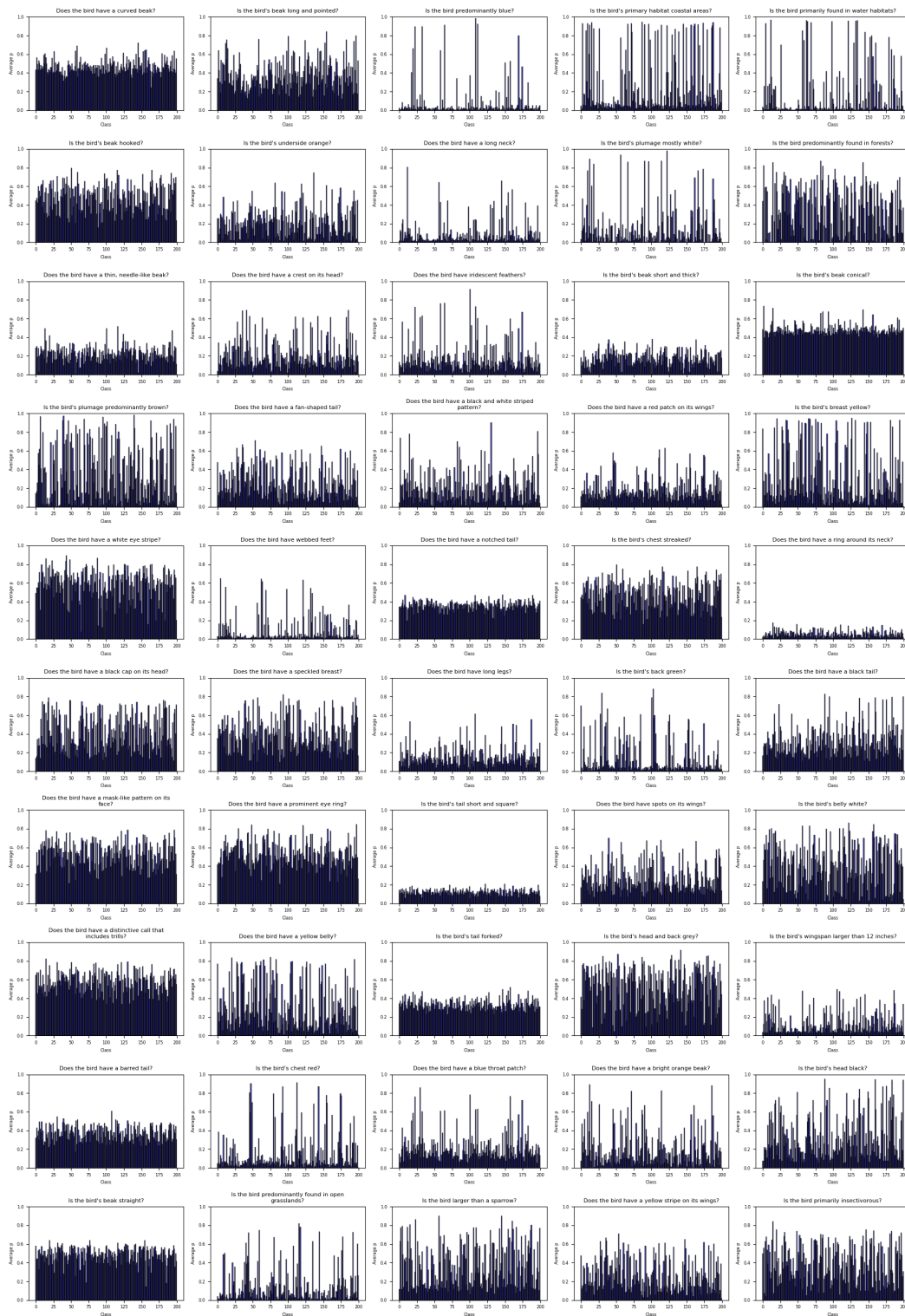


Figure 7: Visualization of the average logit distribution for the multi-aspect of the CUB200.

1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673

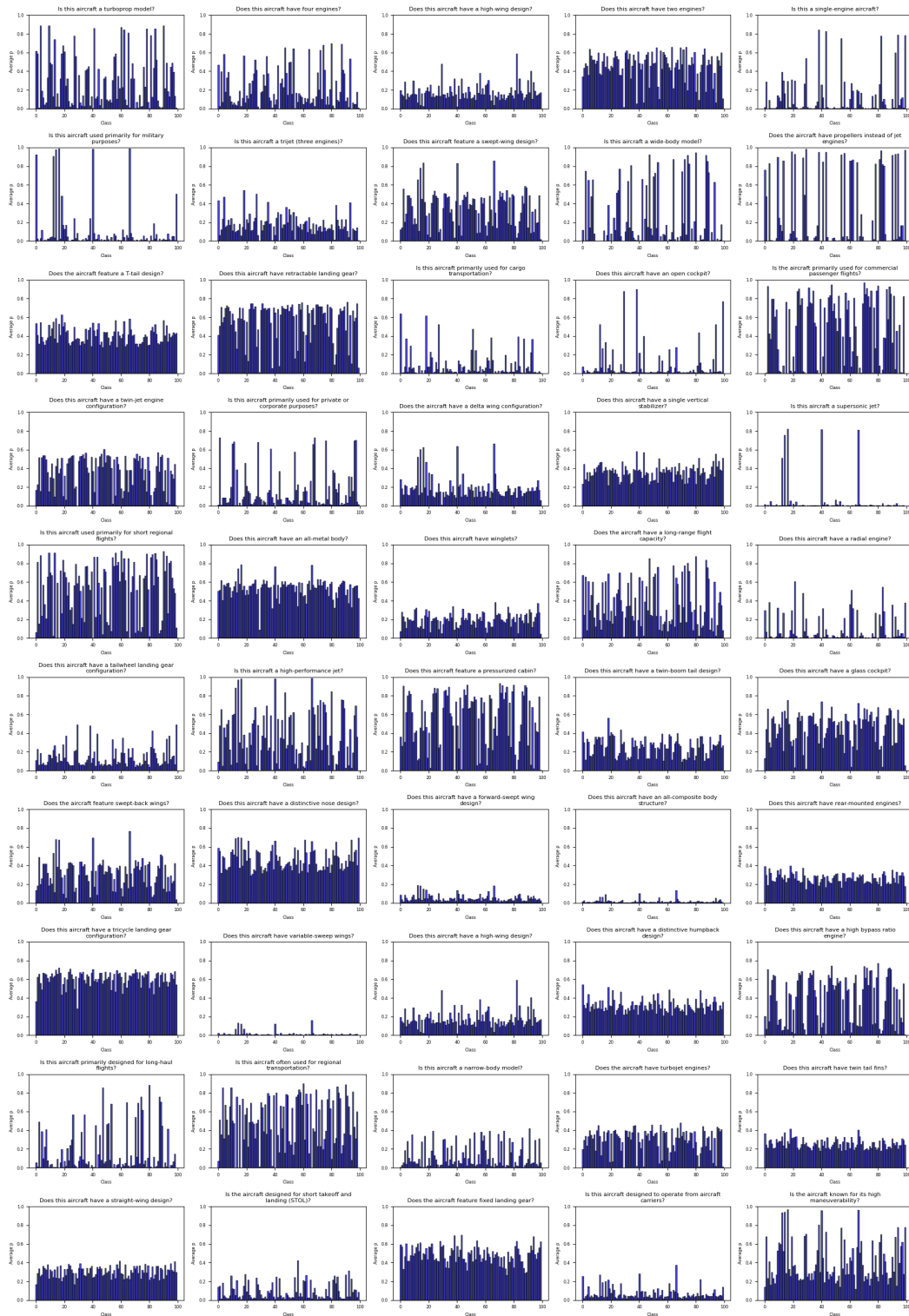


Figure 8: Visualization of the average logit distribution for the multi-aspect of the FGVC-Aircraft.

1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727

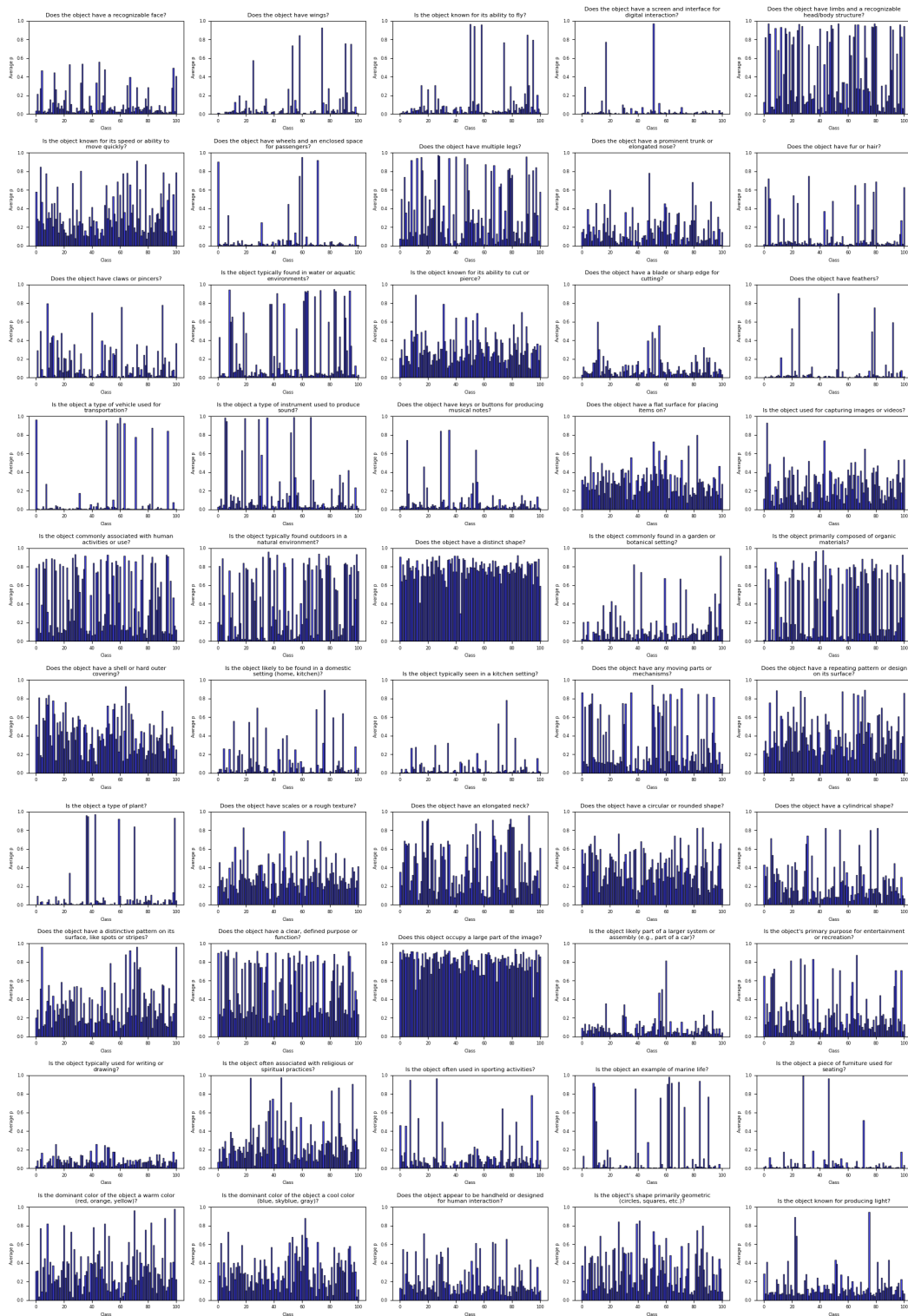


Figure 9: Visualization of the average logit distribution for the multi-aspect of the Caltech101.

1728
 1729
 1730
 1731
 1732
 1733
 1734
 1735
 1736
 1737
 1738
 1739
 1740
 1741
 1742
 1743
 1744
 1745
 1746
 1747
 1748
 1749
 1750
 1751
 1752
 1753
 1754
 1755
 1756
 1757
 1758
 1759
 1760
 1761
 1762
 1763
 1764
 1765
 1766
 1767
 1768
 1769
 1770
 1771
 1772
 1773
 1774
 1775
 1776
 1777
 1778
 1779
 1780
 1781

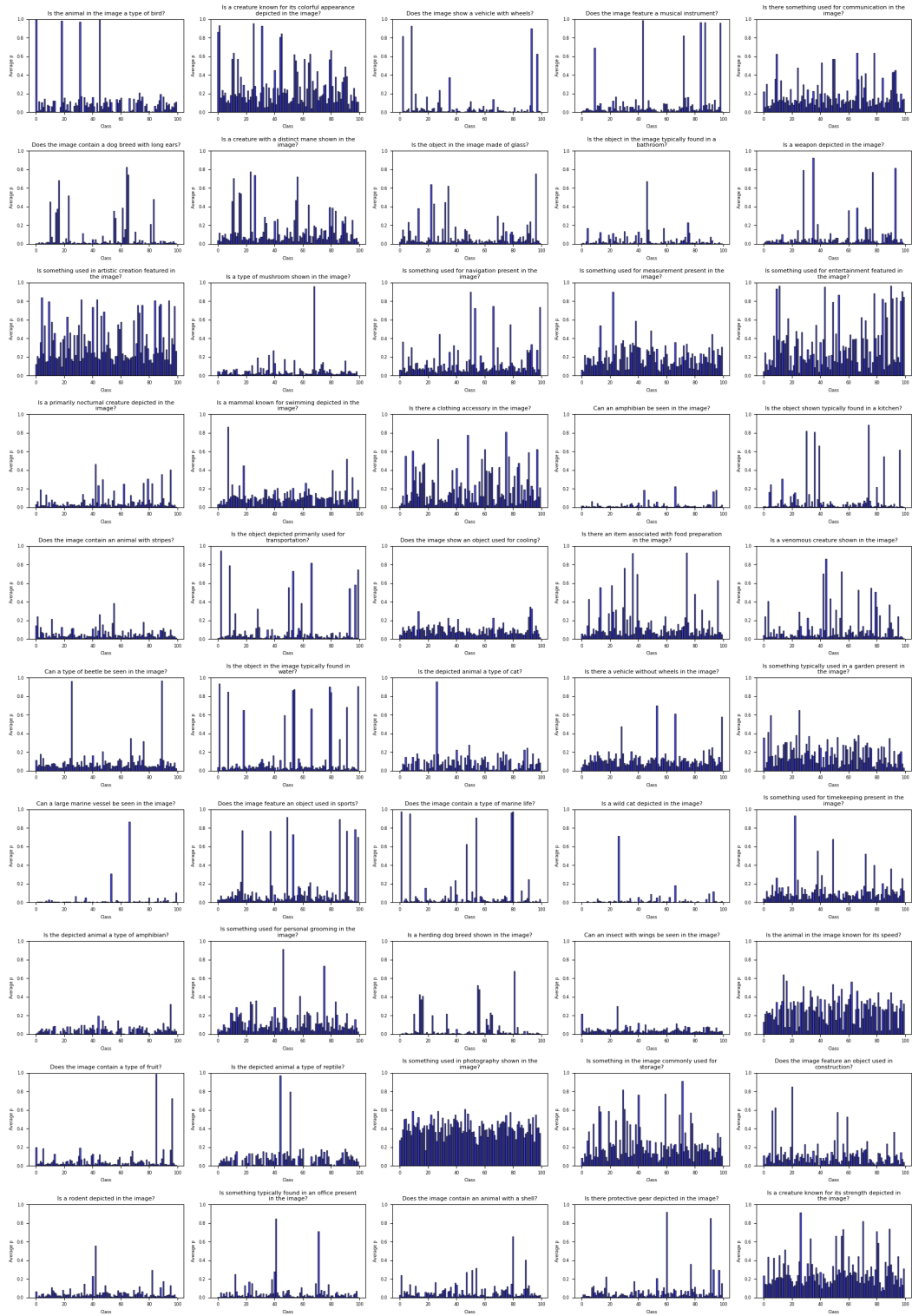


Figure 10: Visualization of the average logit distribution for the multi-aspect of the Mini-ImageNet.

1782
 1783
 1784
 1785
 1786
 1787
 1788
 1789
 1790
 1791
 1792
 1793
 1794
 1795
 1796
 1797
 1798
 1799
 1800
 1801
 1802
 1803
 1804
 1805
 1806
 1807
 1808
 1809
 1810
 1811
 1812
 1813
 1814
 1815
 1816
 1817
 1818
 1819
 1820
 1821
 1822
 1823
 1824
 1825
 1826
 1827
 1828
 1829
 1830
 1831
 1832
 1833
 1834
 1835

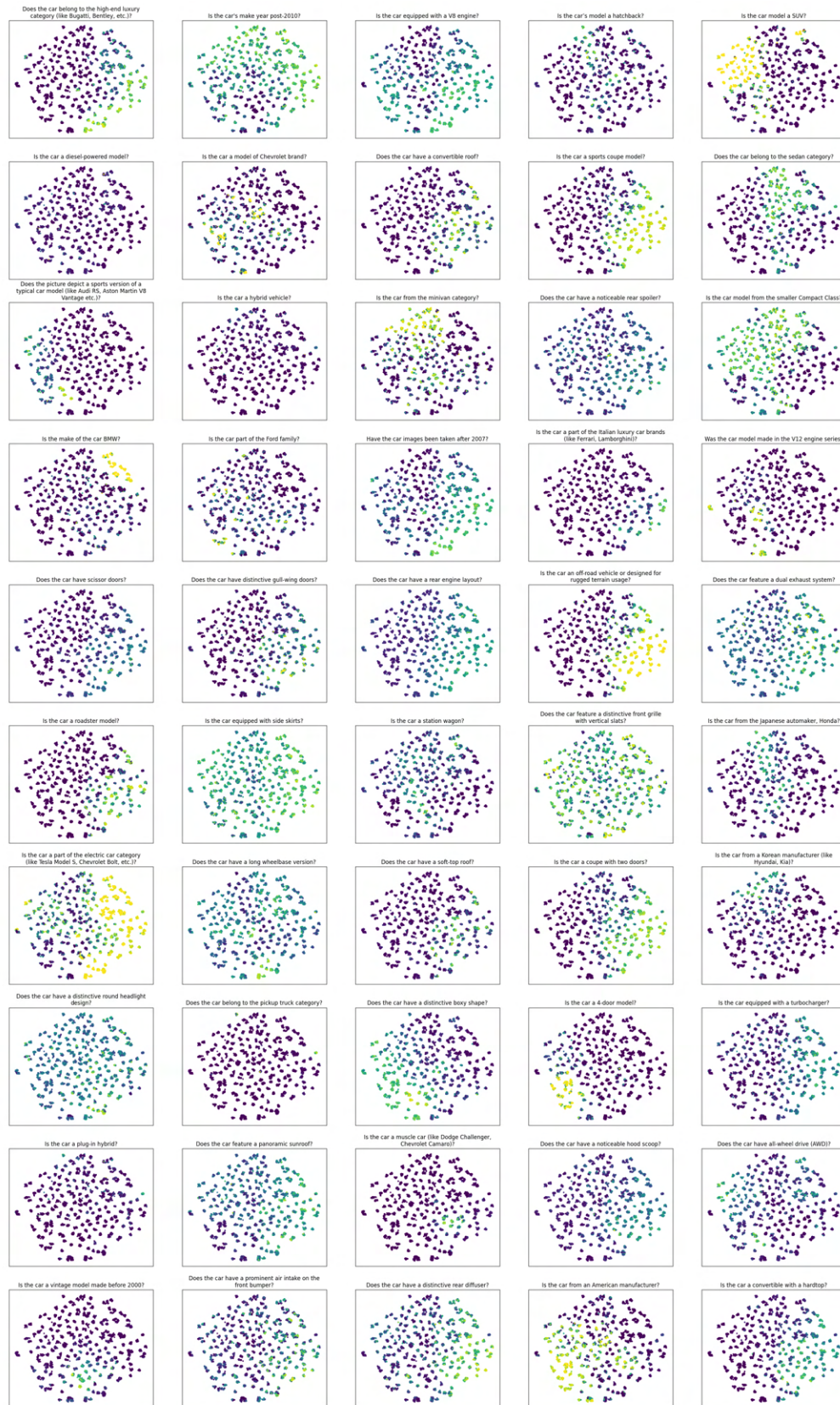


Figure 11: Visualization of ground truth t-SNE embeddings for the multi-aspect of the StanfordCars.

1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889

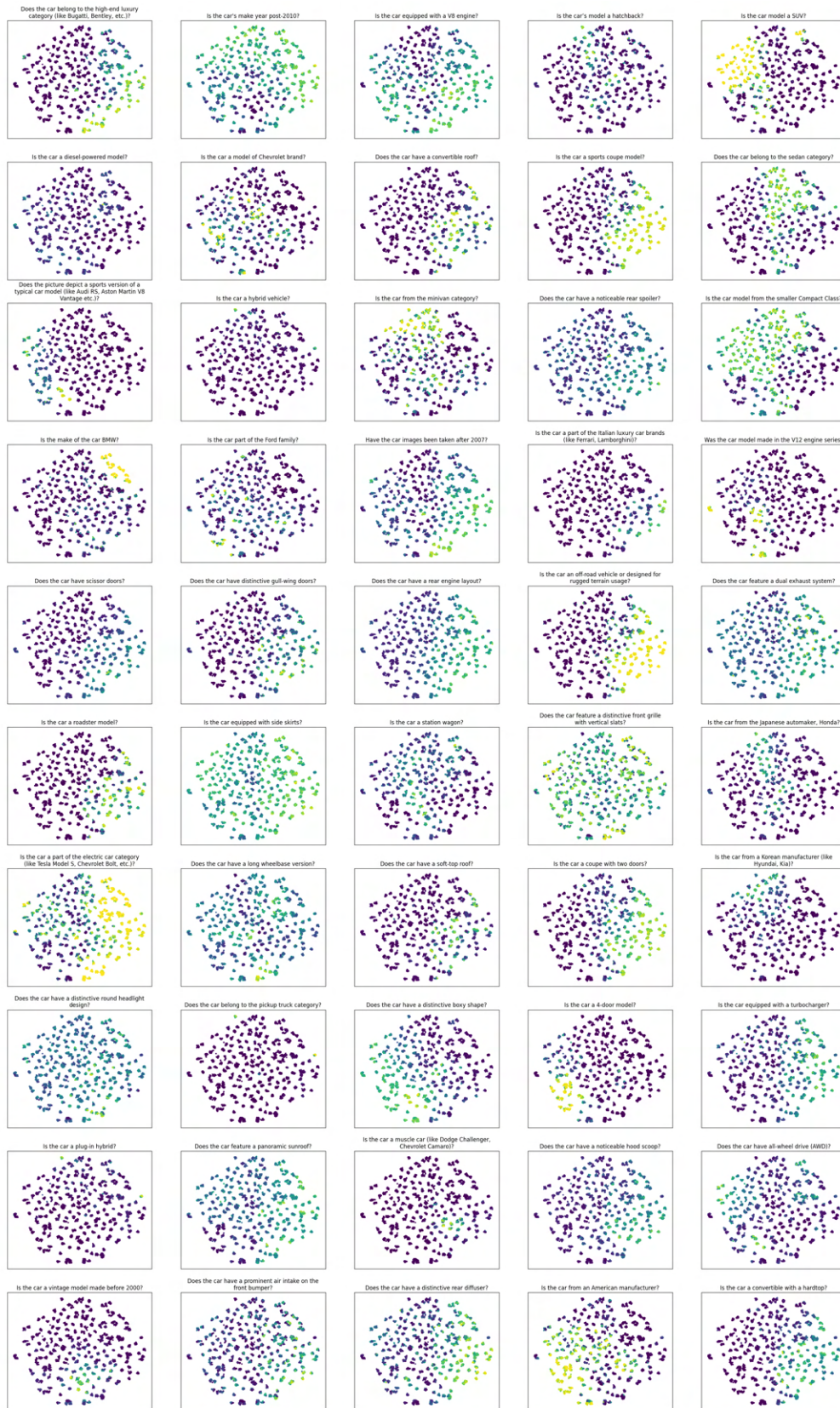


Figure 12: Visualization of predicted result t-SNE embeddings for the multi-aspect of the StanfordCars.

1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943

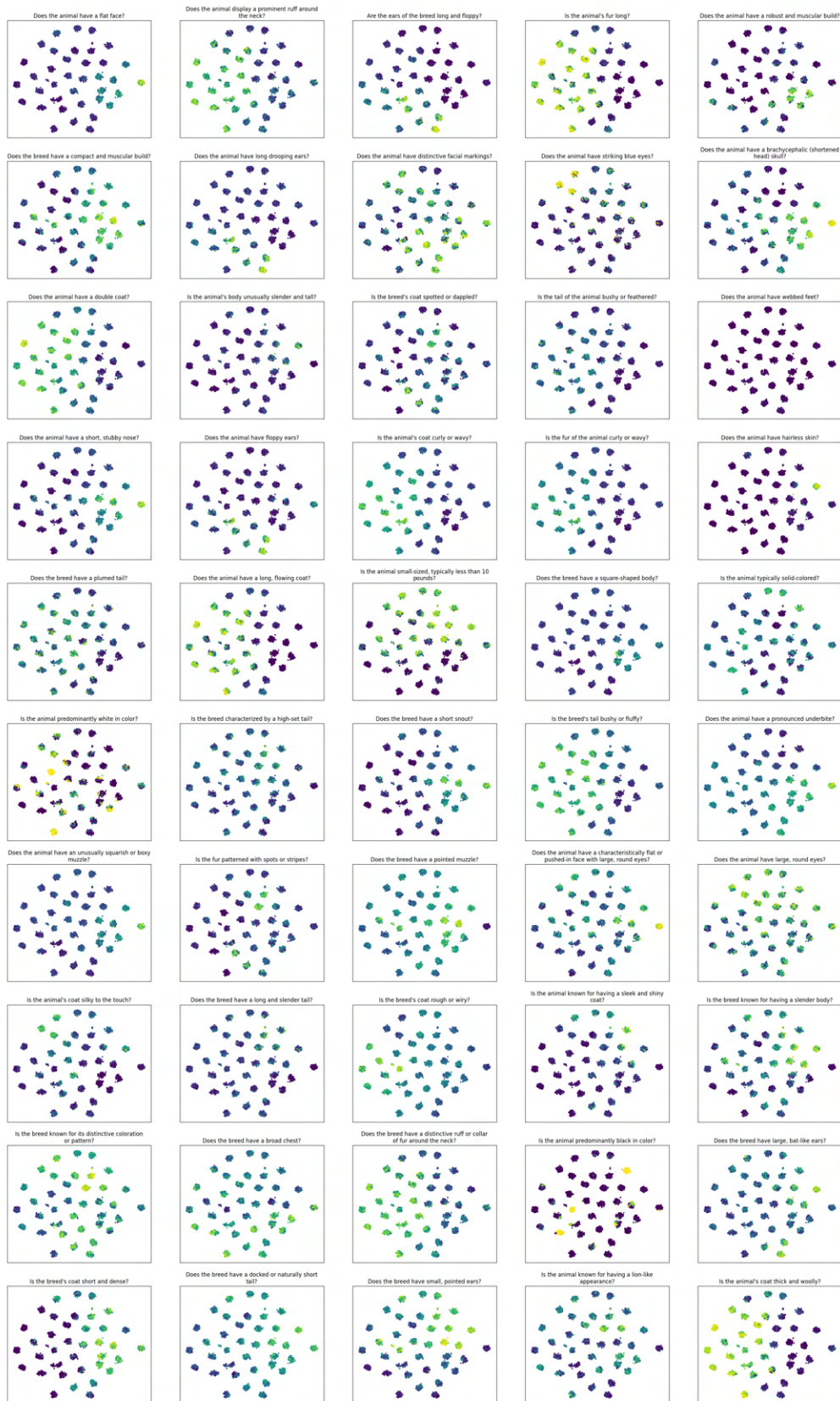


Figure 13: Visualization of ground truth t-SNE embeddings for the multi-aspect of the Oxford-Pets.

1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997

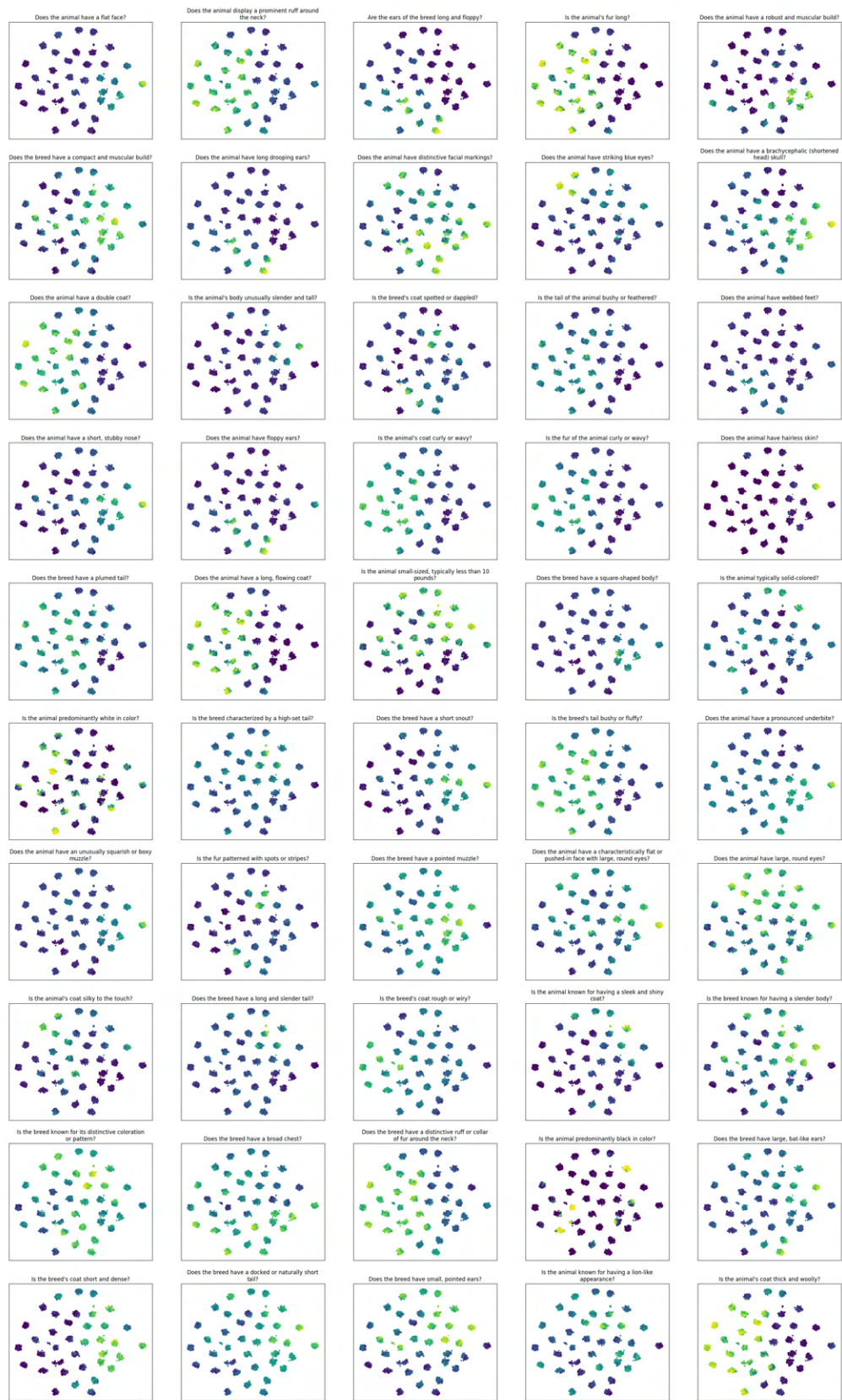


Figure 14: Visualization of predicted result t-SNE embeddings for the multi-aspect of the OxfordPets.

1998
 1999
 2000
 2001
 2002
 2003
 2004
 2005
 2006
 2007
 2008
 2009
 2010
 2011
 2012
 2013
 2014
 2015
 2016
 2017
 2018
 2019
 2020
 2021
 2022
 2023
 2024
 2025
 2026
 2027
 2028
 2029
 2030
 2031
 2032
 2033
 2034
 2035
 2036
 2037
 2038
 2039
 2040
 2041
 2042
 2043
 2044
 2045
 2046
 2047
 2048
 2049
 2050
 2051

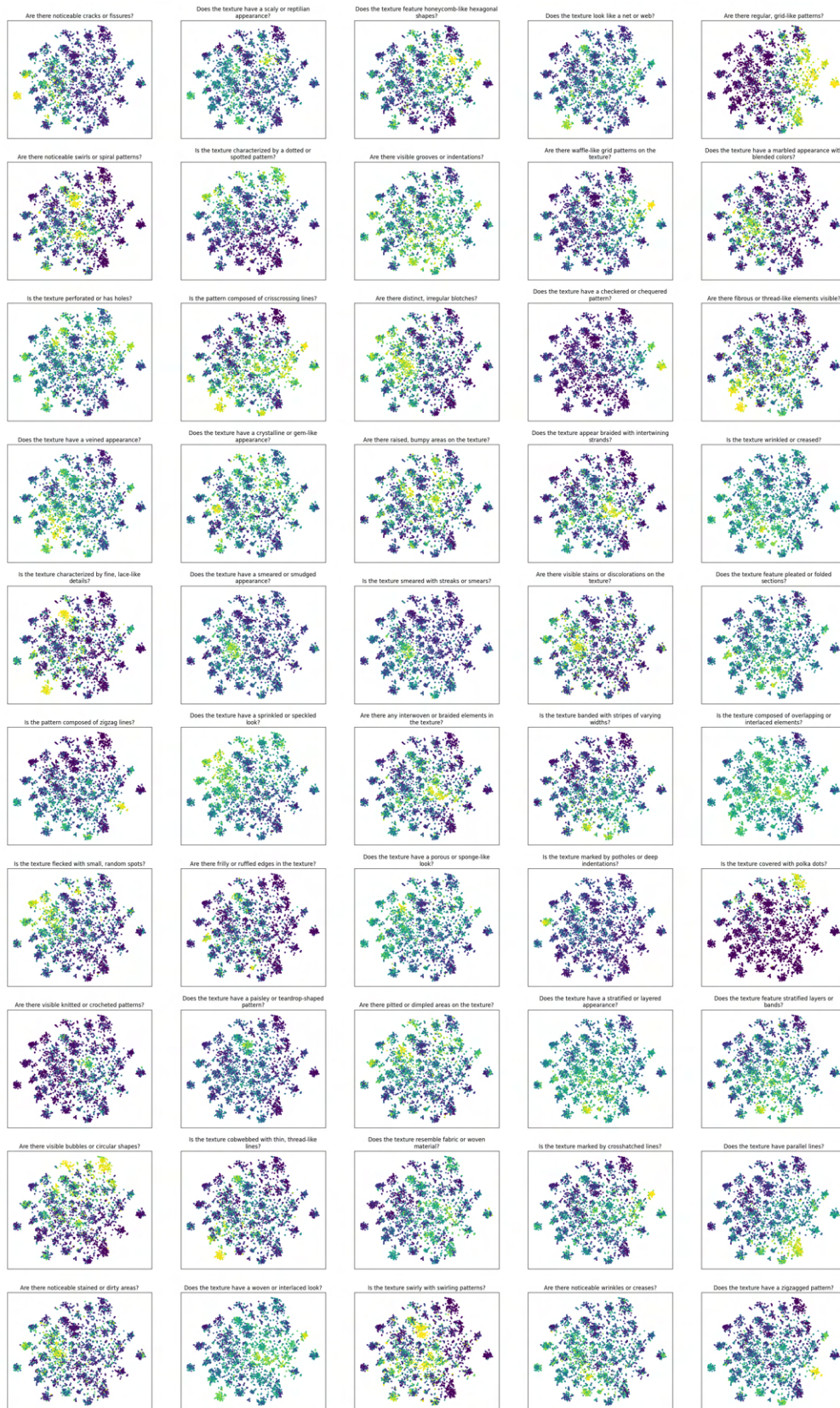


Figure 15: Visualization of ground truth t-SNE embeddings for the multi-aspect of the StanfordCars.

2052
 2053
 2054
 2055
 2056
 2057
 2058
 2059
 2060
 2061
 2062
 2063
 2064
 2065
 2066
 2067
 2068
 2069
 2070
 2071
 2072
 2073
 2074
 2075
 2076
 2077
 2078
 2079
 2080
 2081
 2082
 2083
 2084
 2085
 2086
 2087
 2088
 2089
 2090
 2091
 2092
 2093
 2094
 2095
 2096
 2097
 2098
 2099
 2100
 2101
 2102
 2103
 2104
 2105

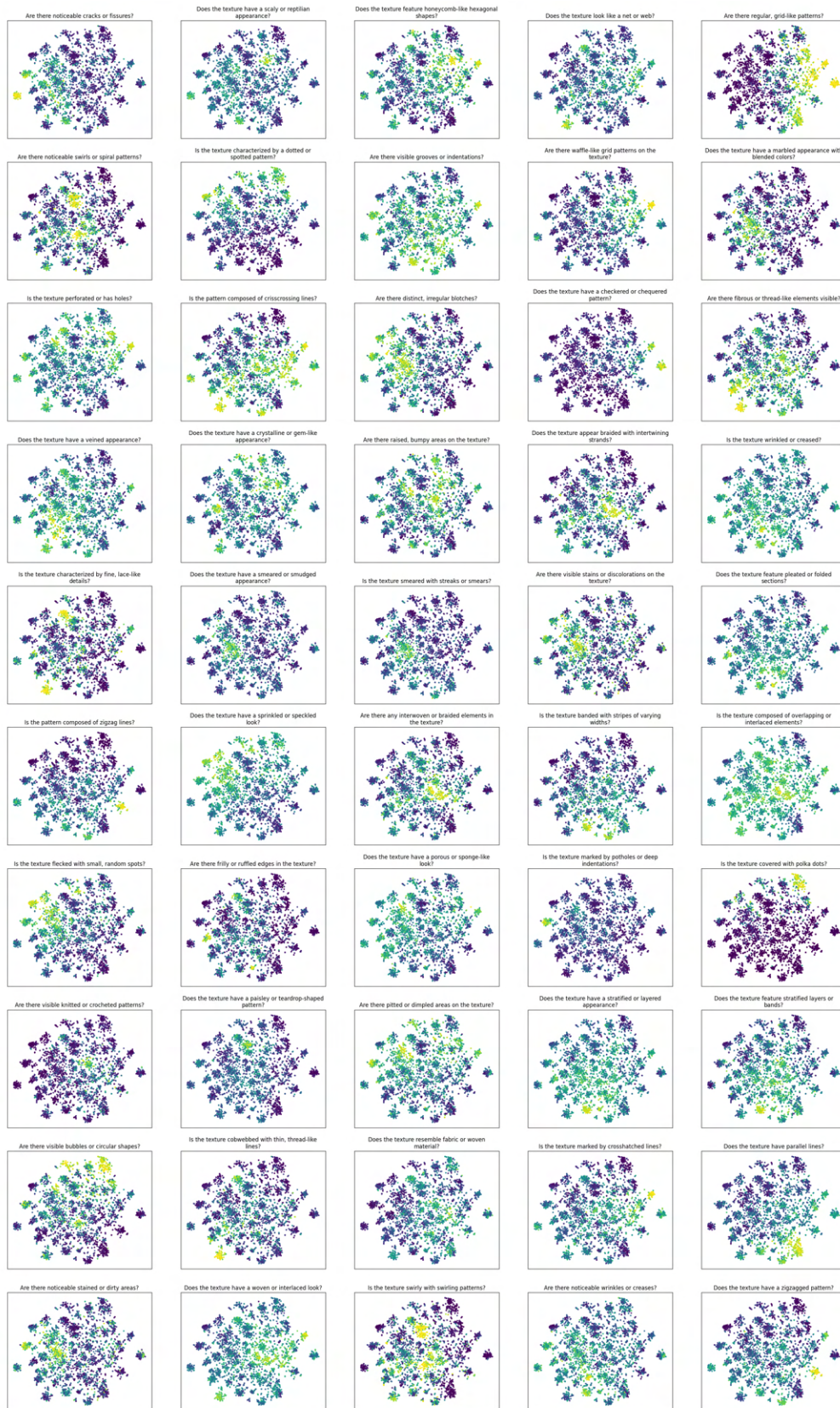


Figure 16: Visualization of predicted result t-SNE embeddings for the multi-aspect of the StanfordCars.

2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159

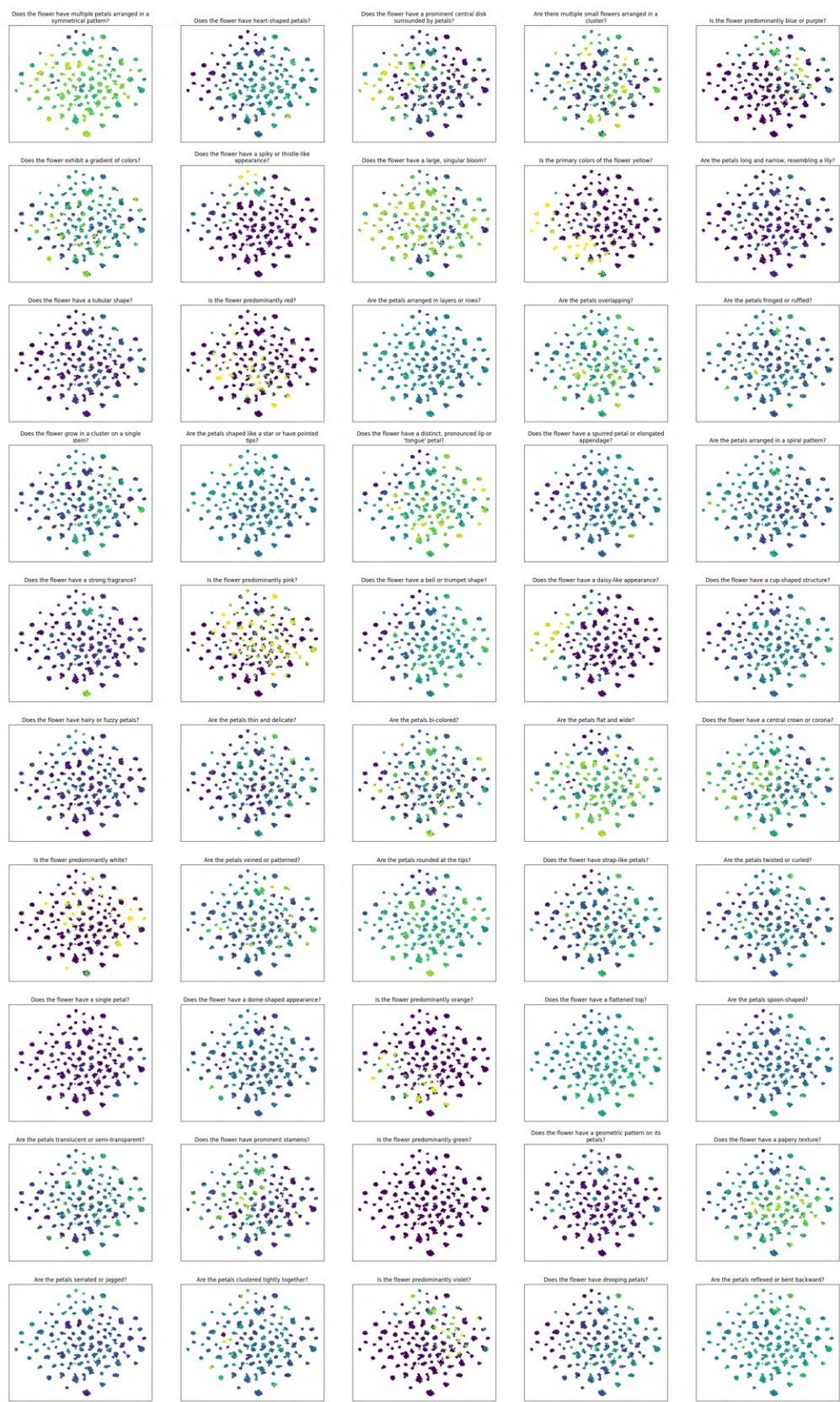


Figure 17: Visualization of ground truth t-SNE embeddings for the multi-aspect of the DTD.

2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213

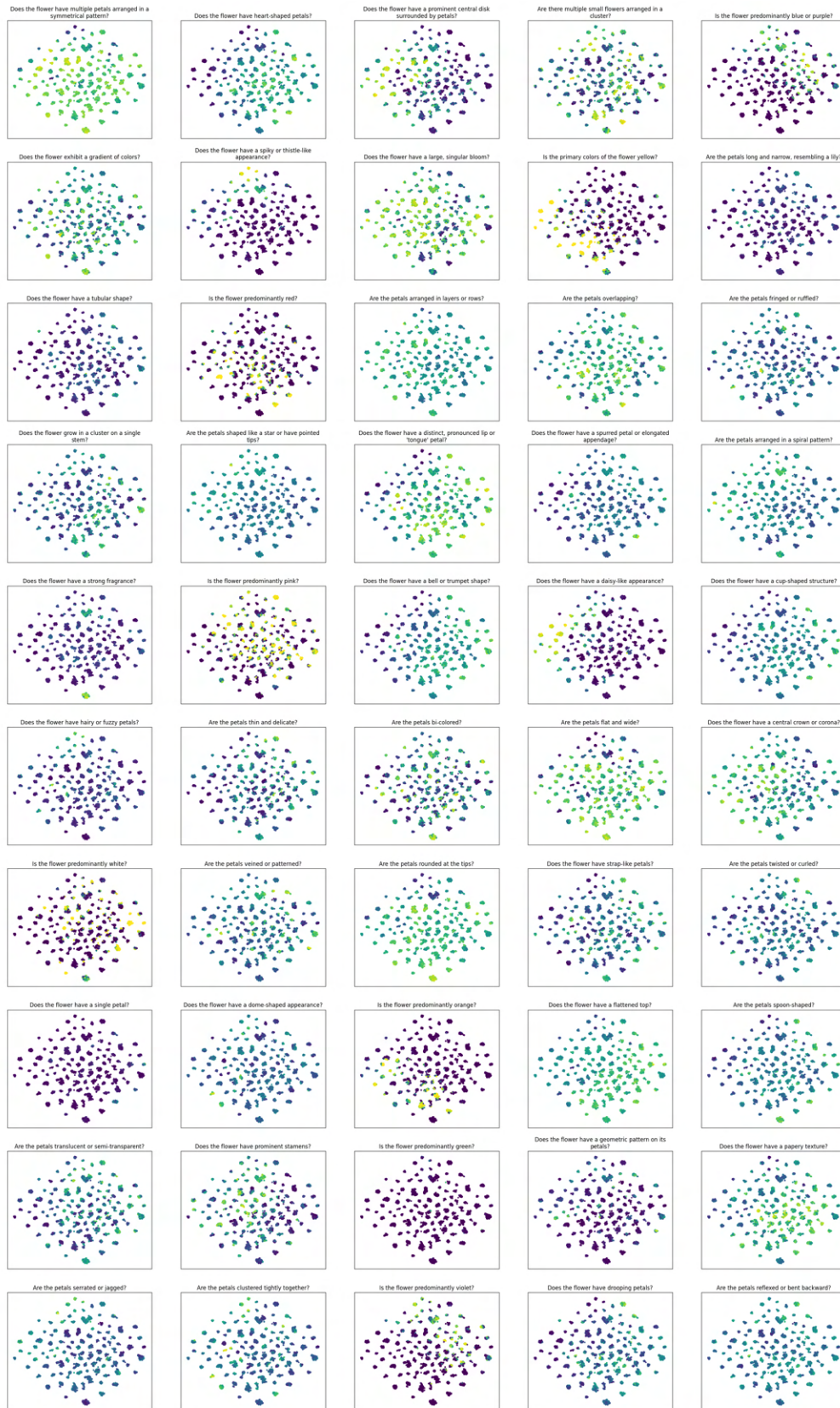


Figure 18: Visualization of predicted result t-SNE embeddings for the multi-aspect of the DTD.

2214
2215
2216
2217
2218
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2230
2231
2232
2233
2234
2235
2236
2237
2238
2239
2240
2241
2242
2243
2244
2245
2246
2247
2248
2249
2250
2251
2252
2253
2254
2255
2256
2257
2258
2259
2260
2261
2262
2263
2264
2265
2266
2267

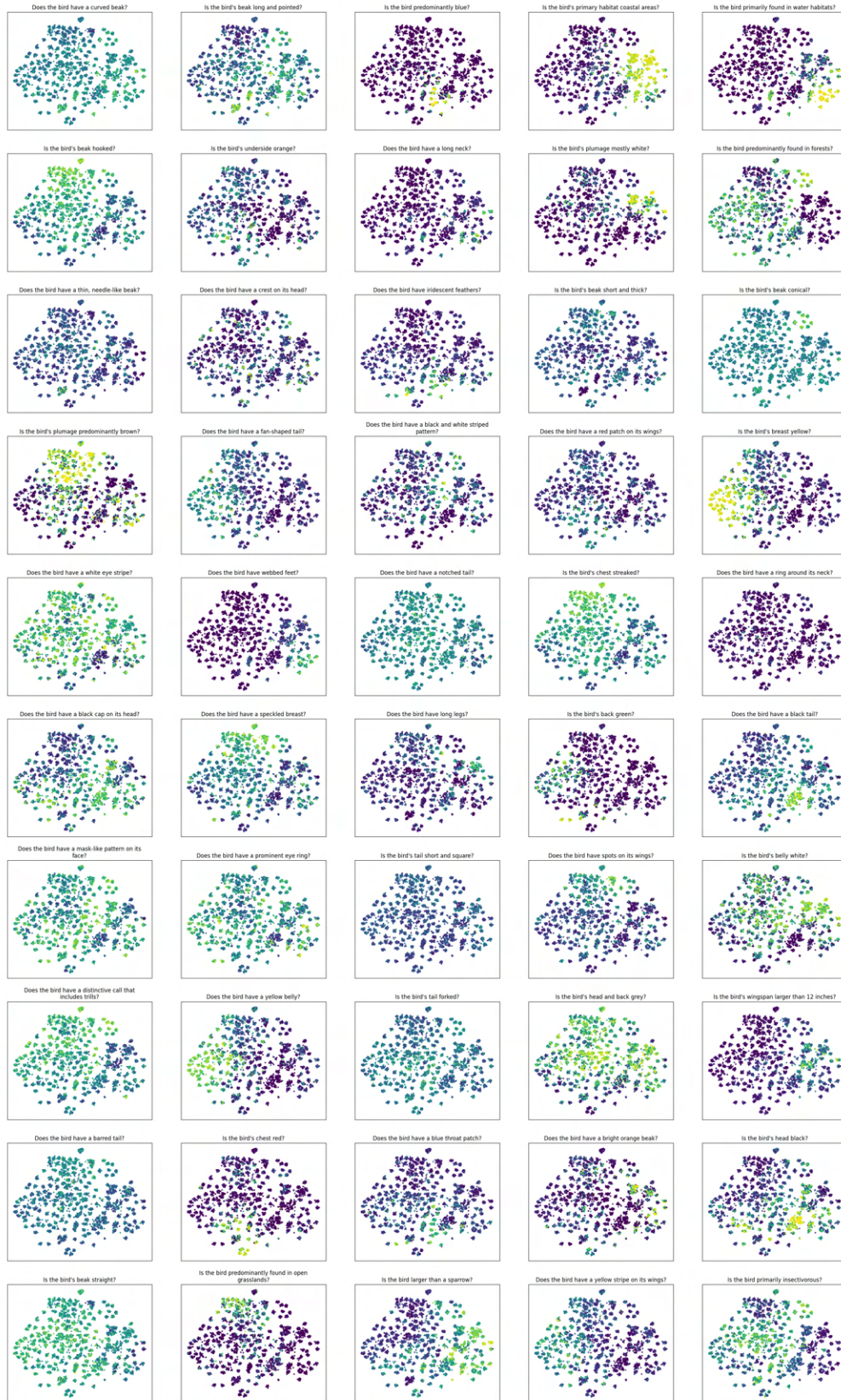


Figure 19: Visualization of ground truth t-SNE embeddings for the multi-aspect of the CUB200.

2268
2269
2270
2271
2272
2273
2274
2275
2276
2277
2278
2279
2280
2281
2282
2283
2284
2285
2286
2287
2288
2289
2290
2291
2292
2293
2294
2295
2296
2297
2298
2299
2300
2301
2302
2303
2304
2305
2306
2307
2308
2309
2310
2311
2312
2313
2314
2315
2316
2317
2318
2319
2320
2321

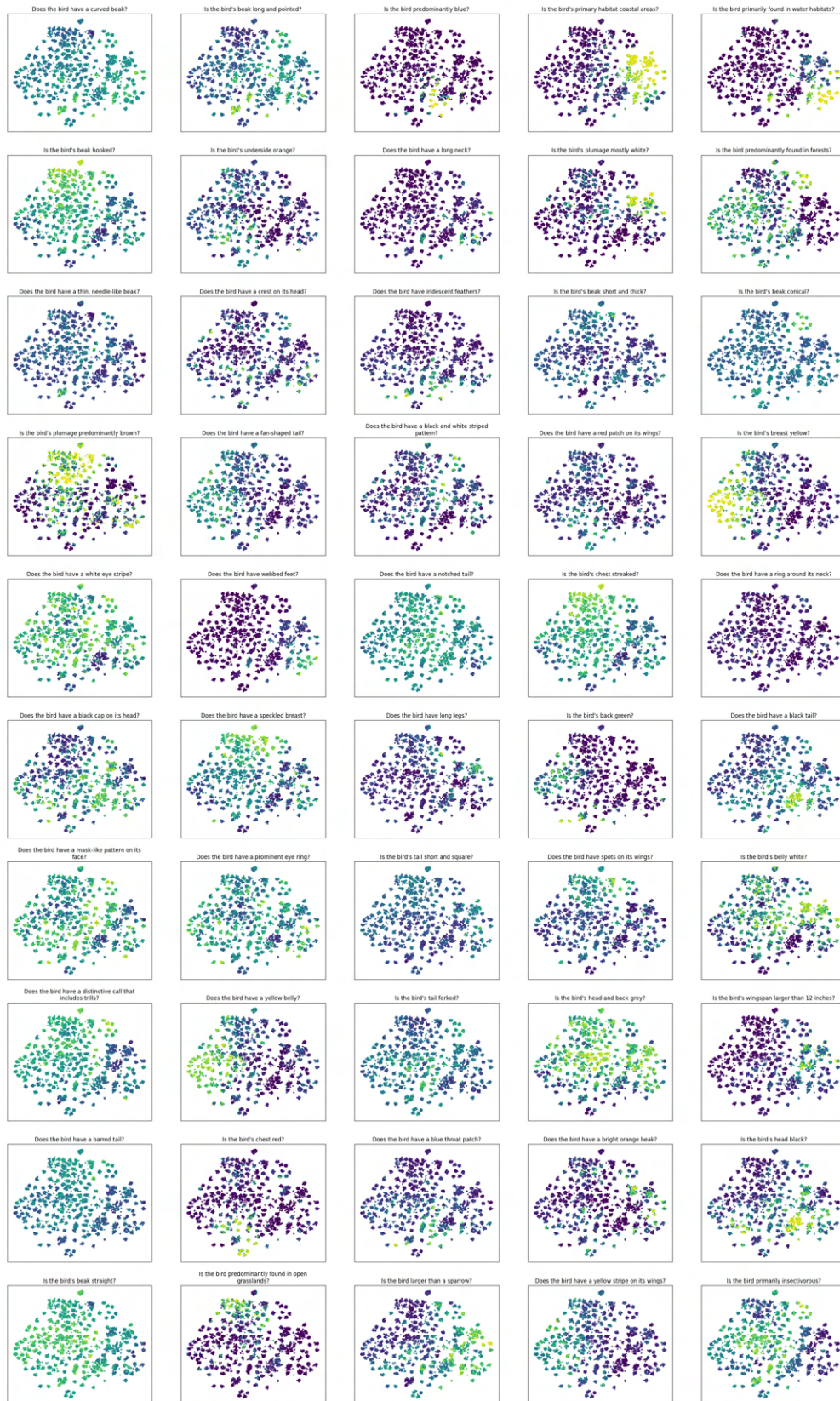


Figure 20: Visualization of predicted result t-SNE embeddings for the multi-aspect of the CUB200.

2322
2323
2324
2325
2326
2327
2328
2329
2330
2331
2332
2333
2334
2335
2336
2337
2338
2339
2340
2341
2342
2343
2344
2345
2346
2347
2348
2349
2350
2351
2352
2353
2354
2355
2356
2357
2358
2359
2360
2361
2362
2363
2364
2365
2366
2367
2368
2369
2370
2371
2372
2373
2374
2375

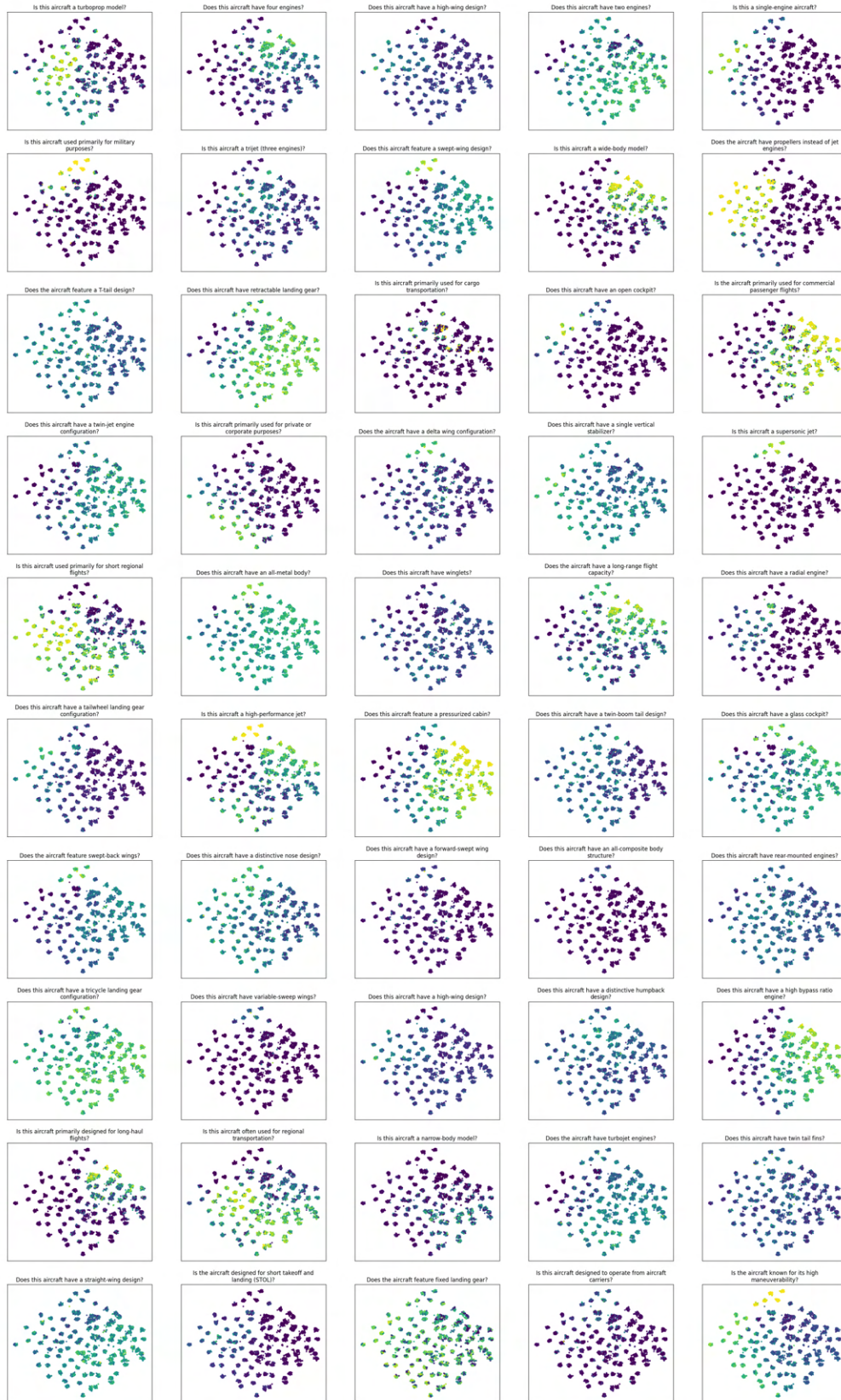


Figure 21: Visualization of ground truth t-SNE embeddings for the multi-aspect of the FGVC-Aircraft.

2376
2377
2378
2379
2380
2381
2382
2383
2384
2385
2386
2387
2388
2389
2390
2391
2392
2393
2394
2395
2396
2397
2398
2399
2400
2401
2402
2403
2404
2405
2406
2407
2408
2409
2410
2411
2412
2413
2414
2415
2416
2417
2418
2419
2420
2421
2422
2423
2424
2425
2426
2427
2428
2429



Figure 22: Visualization of predicted result t-SNE embeddings for the multi-aspect of the FGVC-Aircraft.

2430
2431
2432
2433
2434
2435
2436
2437
2438
2439
2440
2441
2442
2443
2444
2445
2446
2447
2448
2449
2450
2451
2452
2453
2454
2455
2456
2457
2458
2459
2460
2461
2462
2463
2464
2465
2466
2467
2468
2469
2470
2471
2472
2473
2474
2475
2476
2477
2478
2479
2480
2481
2482
2483

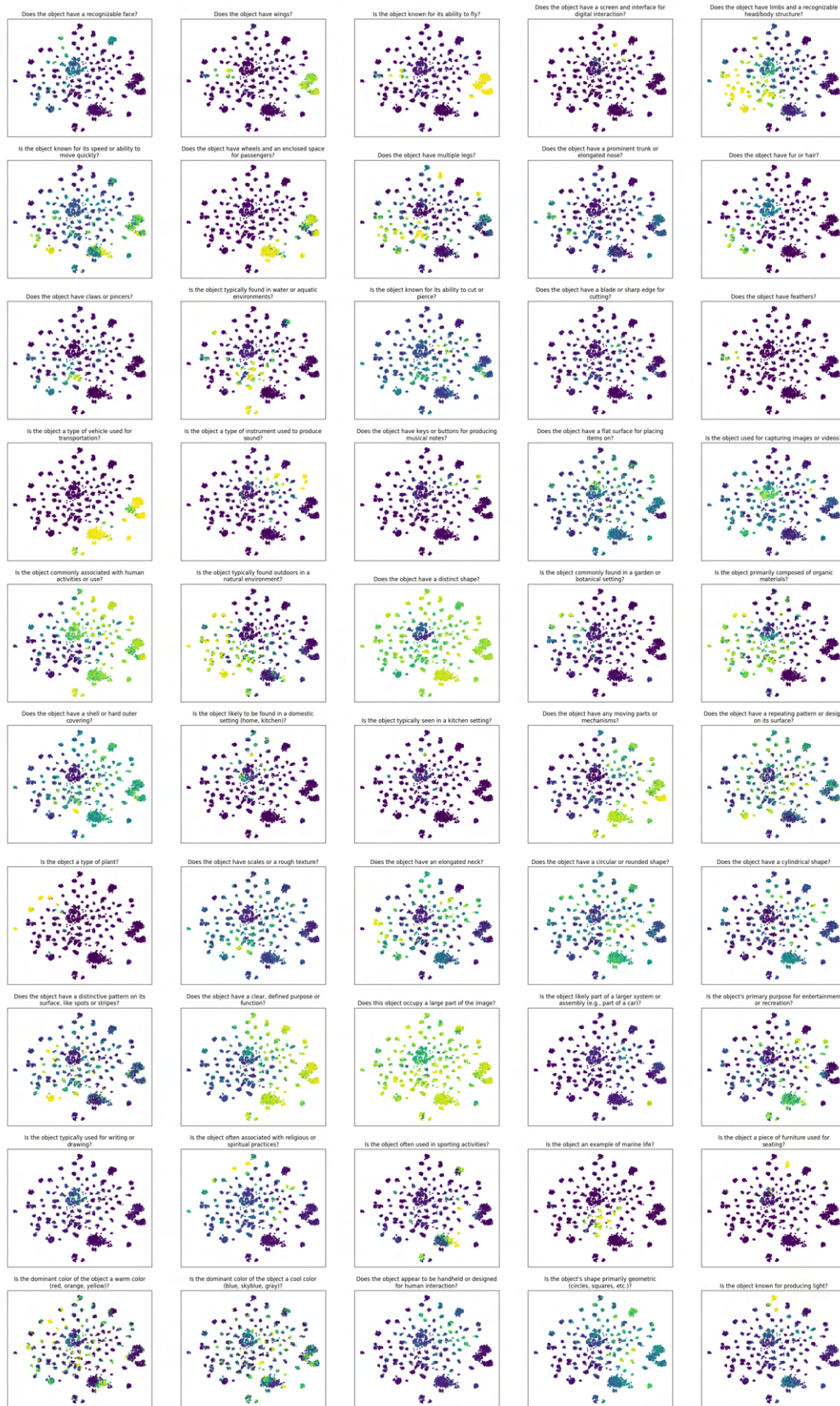


Figure 23: Visualization of ground truth t-SNE embeddings for the multi-aspect of the Caltech101.

2484
2485
2486
2487
2488
2489
2490
2491
2492
2493
2494
2495
2496
2497
2498
2499
2500
2501
2502
2503
2504
2505
2506
2507
2508
2509
2510
2511
2512
2513
2514
2515
2516
2517
2518
2519
2520
2521
2522
2523
2524
2525
2526
2527
2528
2529
2530
2531
2532
2533
2534
2535
2536
2537

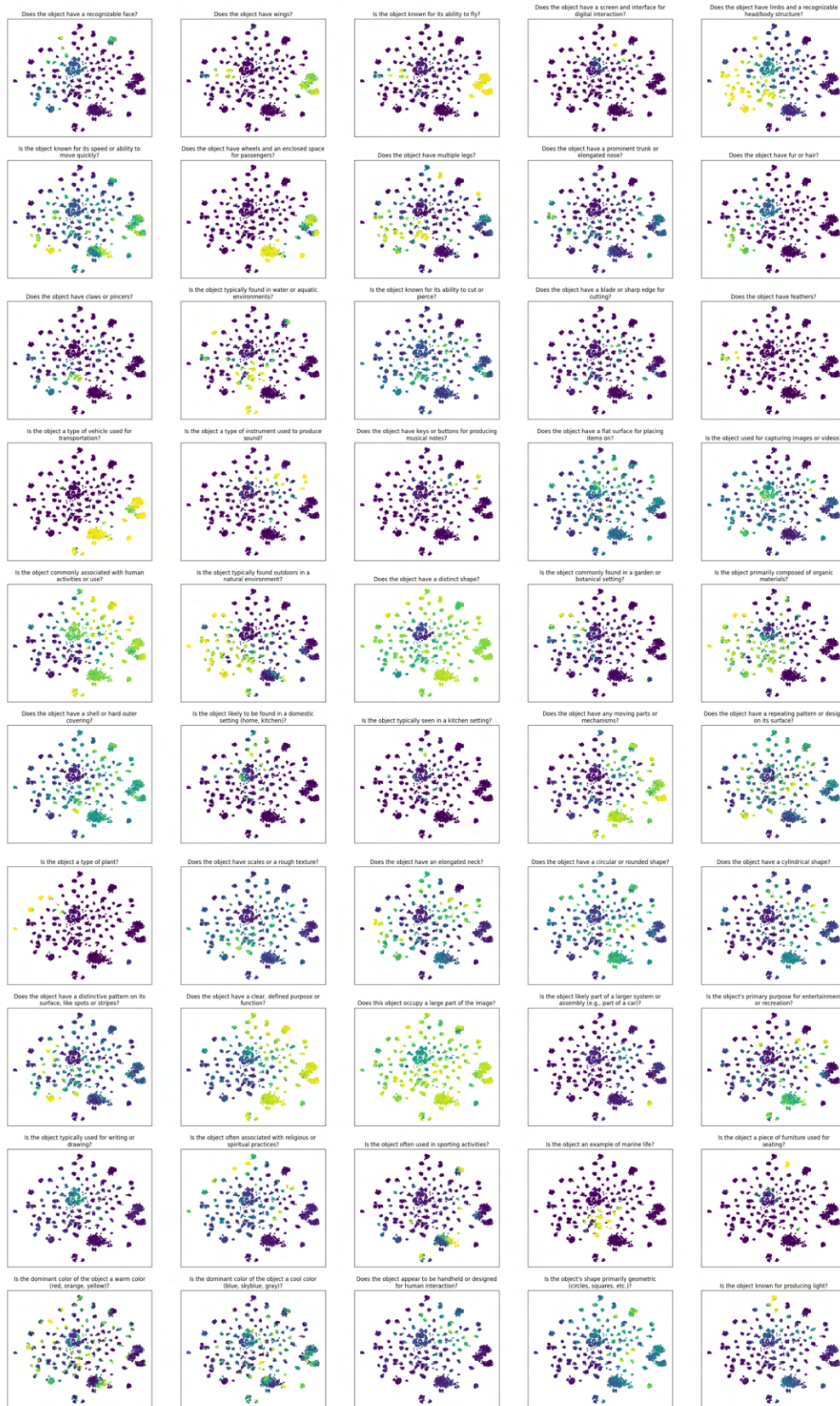


Figure 24: Visualization of predicted result t-SNE embeddings for the multi-aspect of the Cal-tech101.

2538
 2539
 2540
 2541
 2542
 2543
 2544
 2545
 2546
 2547
 2548
 2549
 2550
 2551
 2552
 2553
 2554
 2555
 2556
 2557
 2558
 2559
 2560
 2561
 2562
 2563
 2564
 2565
 2566
 2567
 2568
 2569
 2570
 2571
 2572
 2573
 2574
 2575
 2576
 2577
 2578
 2579
 2580
 2581
 2582
 2583
 2584
 2585
 2586
 2587
 2588
 2589
 2590
 2591

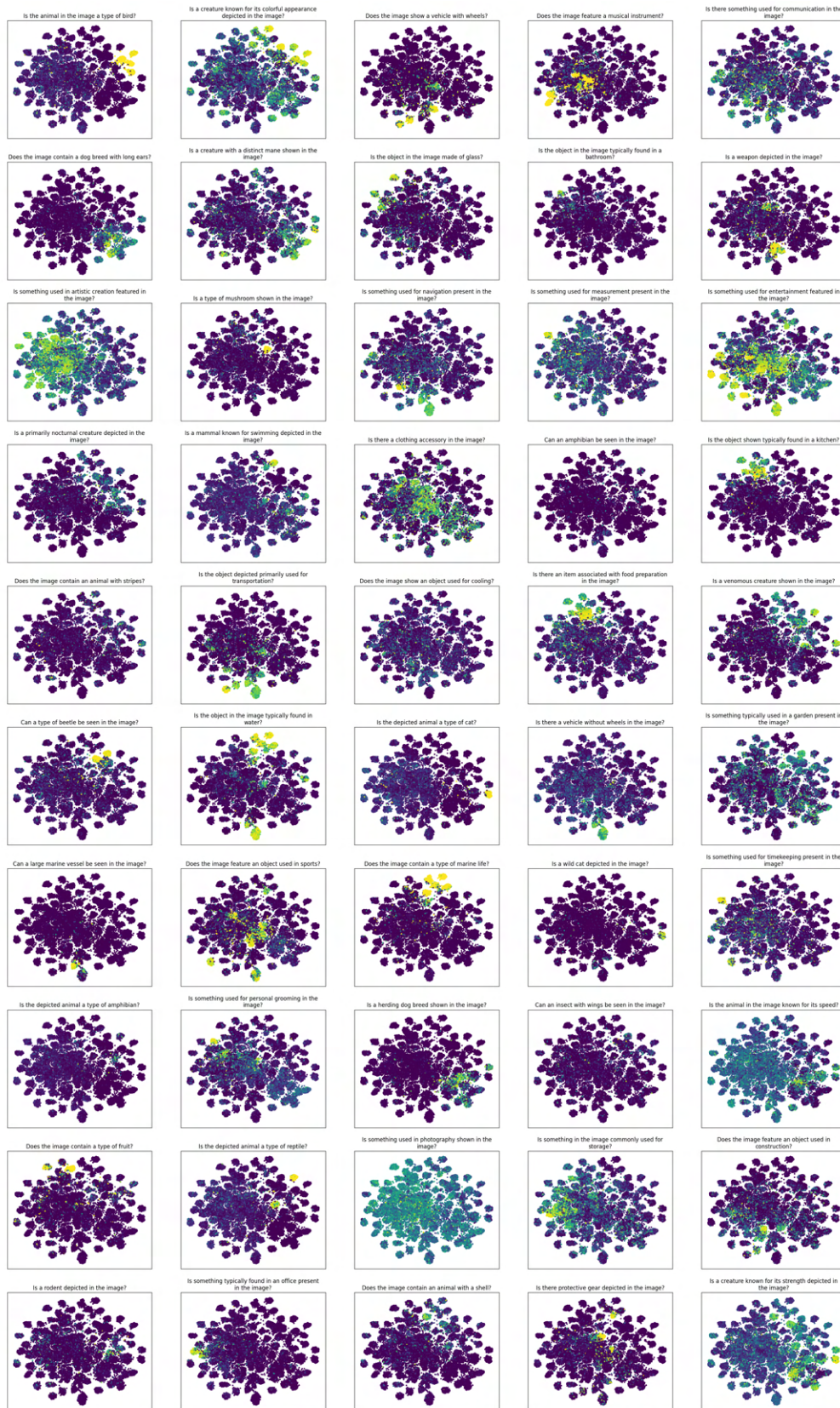


Figure 25: Visualization of ground truth t-SNE embeddings for the multi-aspect of the Mini-ImageNet.

2592
2593
2594
2595
2596
2597
2598
2599
2600
2601
2602
2603
2604
2605
2606
2607
2608
2609
2610
2611
2612
2613
2614
2615
2616
2617
2618
2619
2620
2621
2622
2623
2624
2625
2626
2627
2628
2629
2630
2631
2632
2633
2634
2635
2636
2637
2638
2639
2640
2641
2642
2643
2644
2645

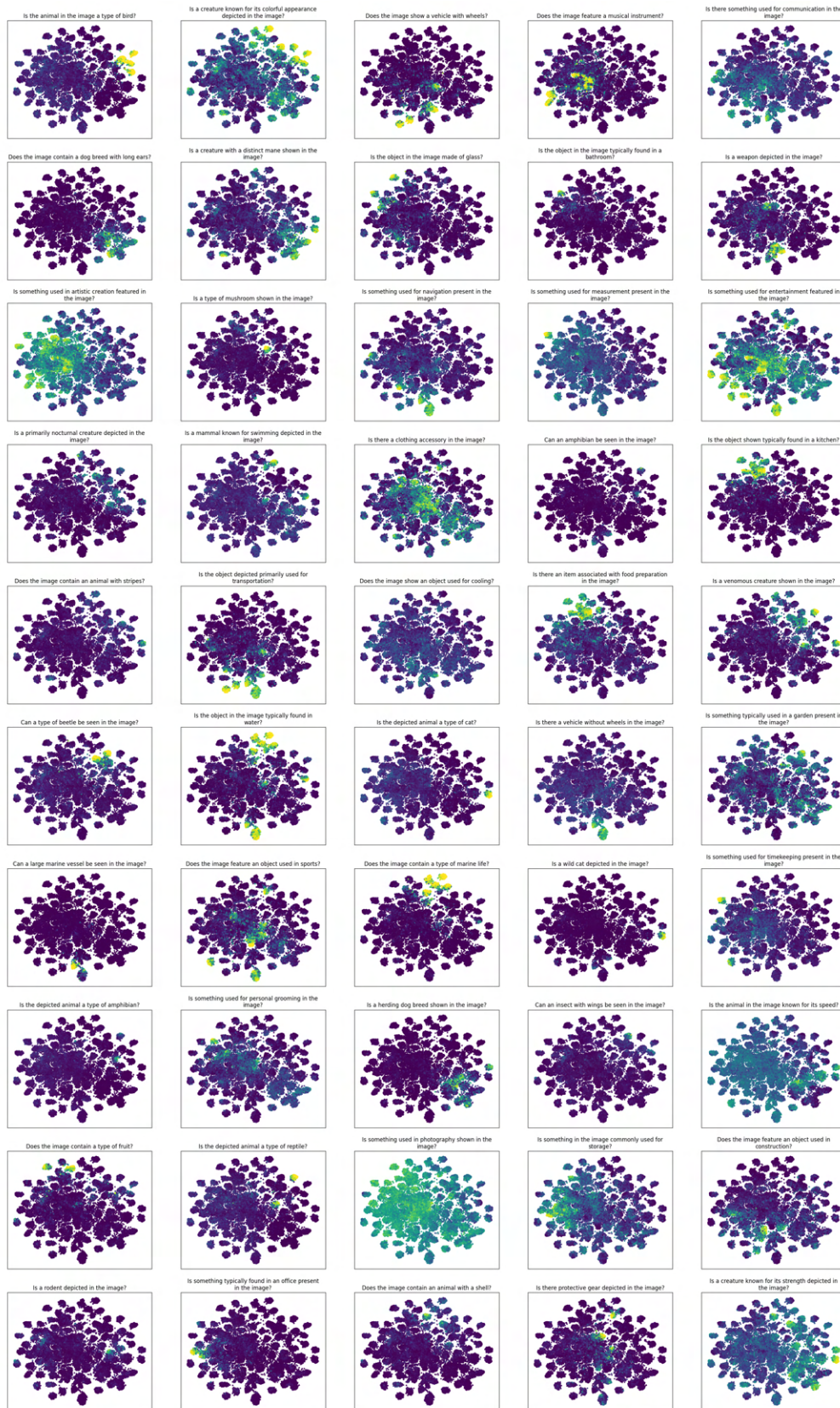


Figure 26: Visualization of predicted result t-SNE embeddings for the multi-aspect of the Mini-ImageNet.

2646
2647
2648
2649
2650
2651
2652
2653
2654
2655
2656
2657
2658
2659
2660
2661
2662
2663
2664
2665
2666
2667
2668
2669
2670
2671
2672
2673
2674
2675
2676
2677
2678
2679
2680
2681
2682
2683
2684
2685
2686
2687
2688
2689
2690
2691
2692
2693
2694
2695
2696
2697
2698
2699

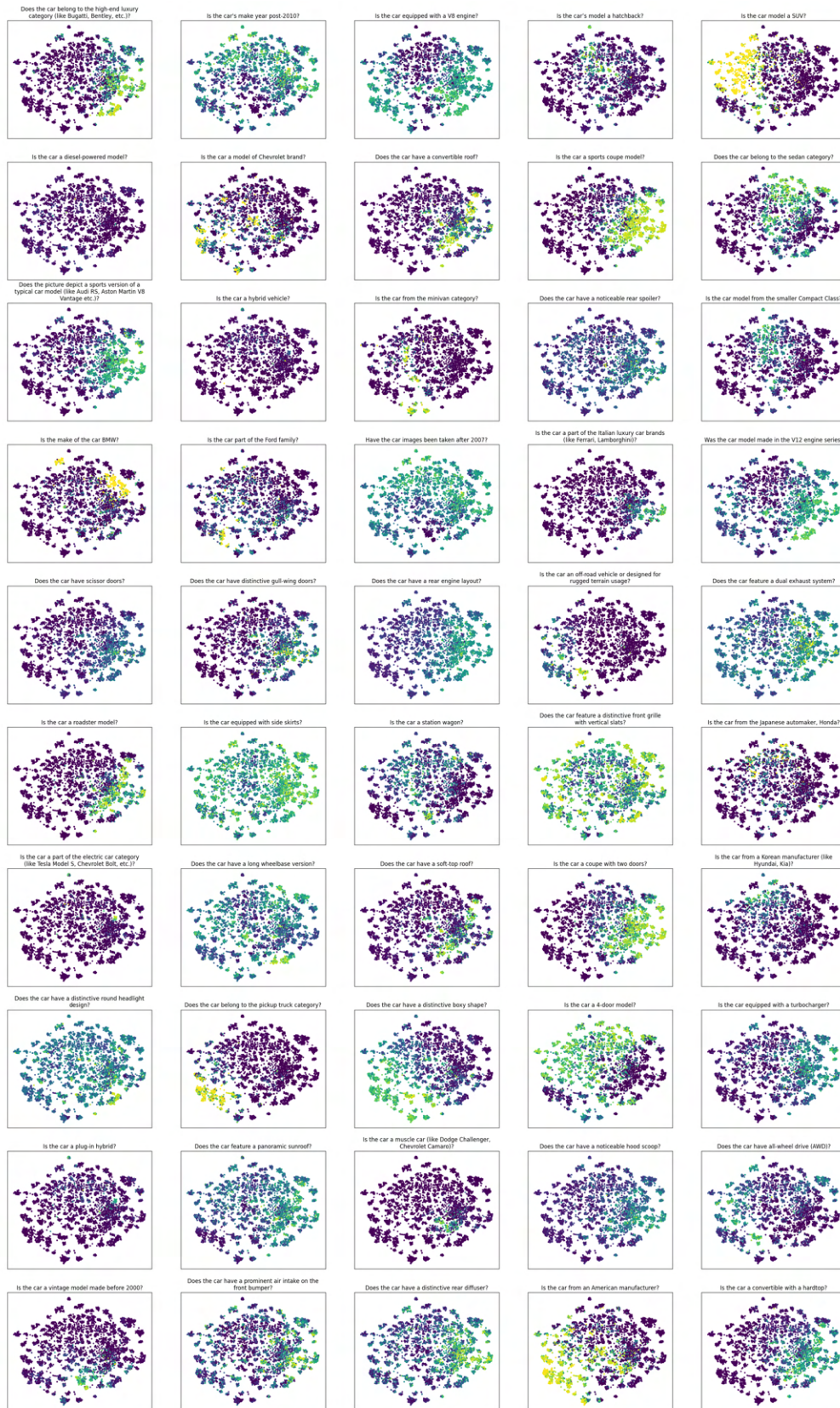


Figure 27: Visualization of ground truth t-SNE embeddings for the multi-aspect of the StanfordCars testing dataset.

2700
2701
2702
2703
2704
2705
2706
2707
2708
2709
2710
2711
2712
2713
2714
2715
2716
2717
2718
2719
2720
2721
2722
2723
2724
2725
2726
2727
2728
2729
2730
2731
2732
2733
2734
2735
2736
2737
2738
2739
2740
2741
2742
2743
2744
2745
2746
2747
2748
2749
2750
2751
2752
2753

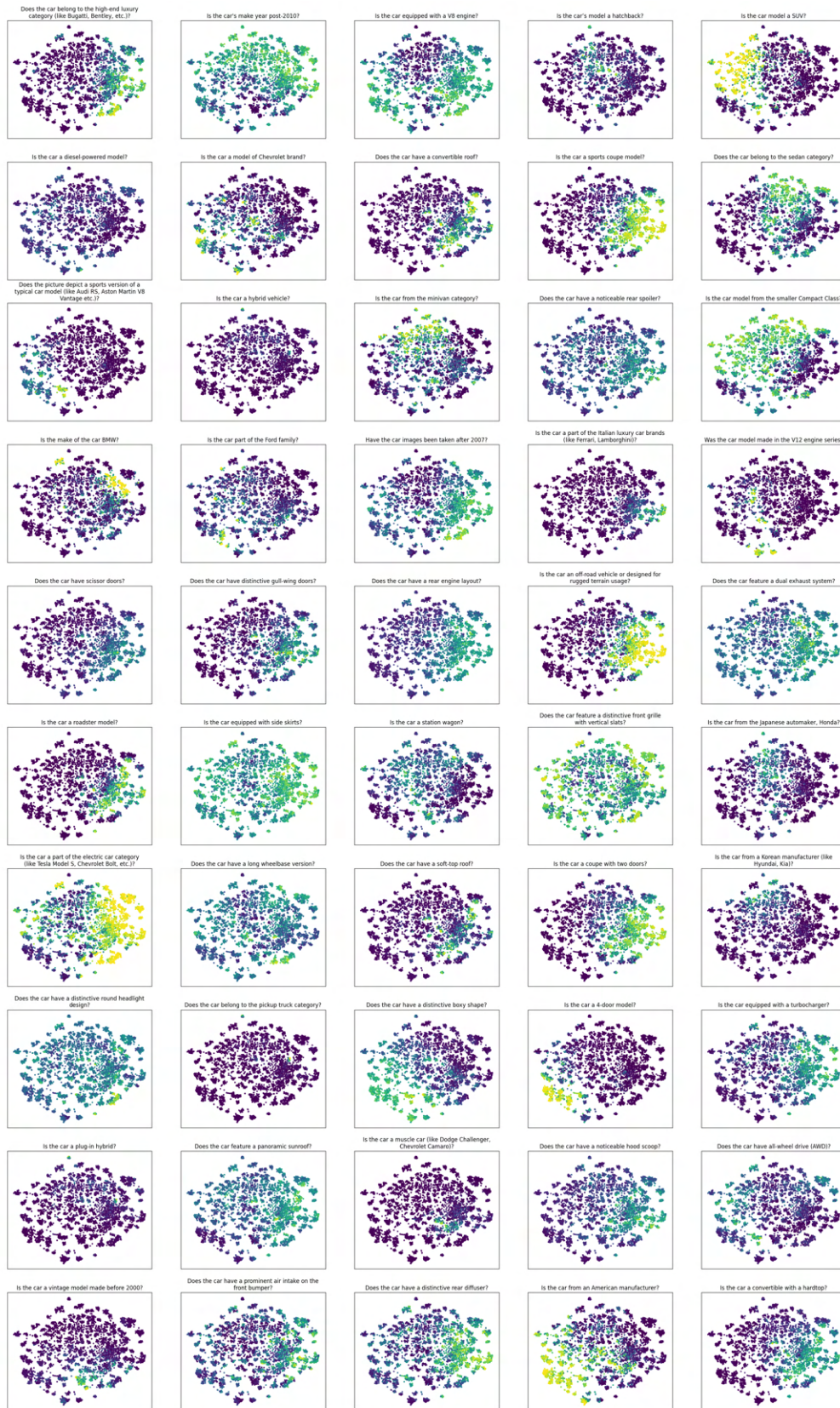


Figure 28: Visualization of predicted result t-SNE embeddings for the multi-aspect of the StanfordCars testing dataset.

2754
2755
2756
2757
2758
2759
2760
2761
2762
2763
2764
2765
2766
2767
2768
2769
2770
2771
2772
2773
2774
2775
2776
2777
2778
2779
2780
2781
2782
2783
2784
2785
2786
2787
2788
2789
2790
2791
2792
2793
2794
2795
2796
2797
2798
2799
2800
2801
2802
2803
2804
2805
2806
2807

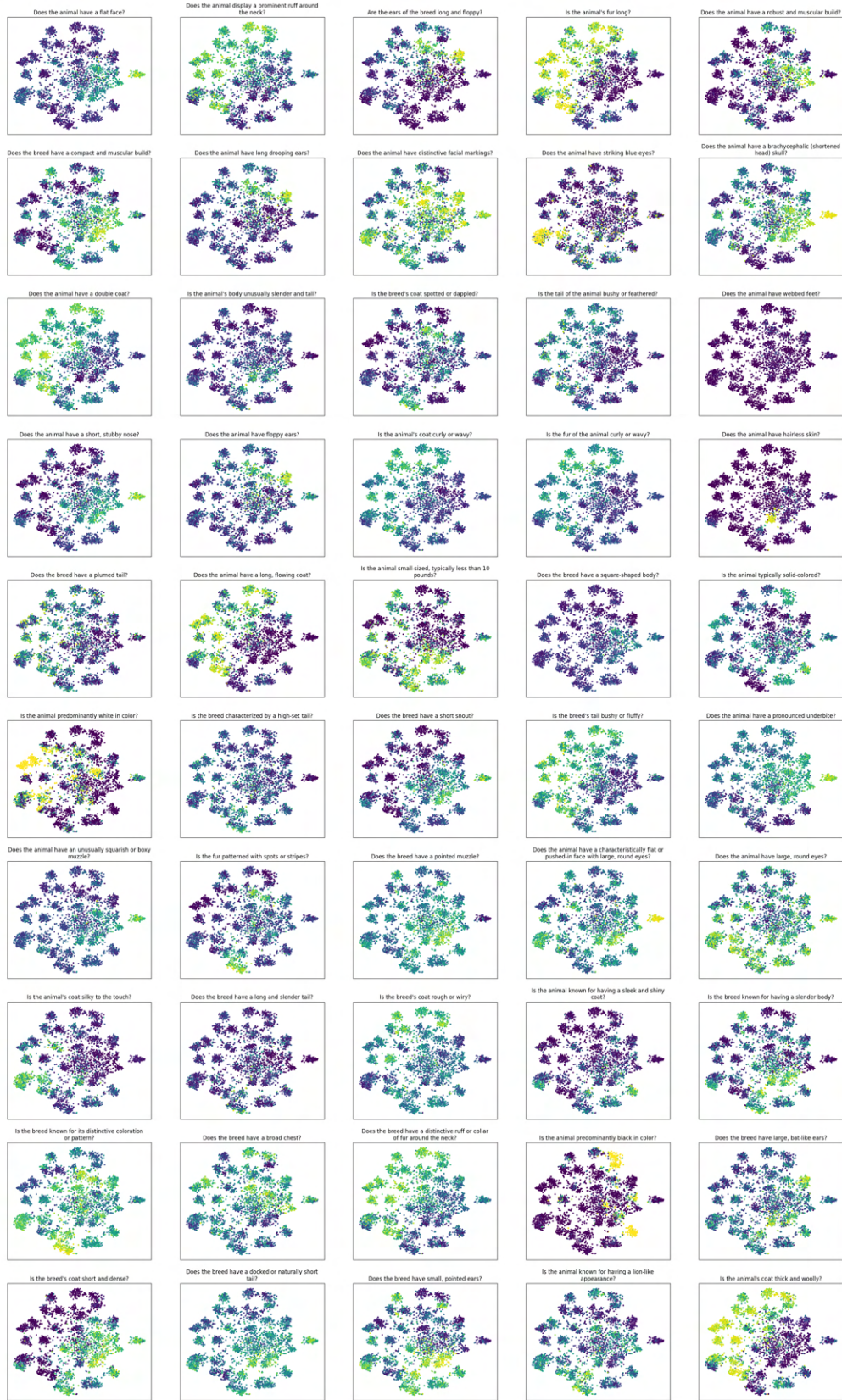


Figure 29: Visualization of ground truth t-SNE embeddings for the multi-aspect of the Oxford-Pets testing dataset.

2808
2809
2810
2811
2812
2813
2814
2815
2816
2817
2818
2819
2820
2821
2822
2823
2824
2825
2826
2827
2828
2829
2830
2831
2832
2833
2834
2835
2836
2837
2838
2839
2840
2841
2842
2843
2844
2845
2846
2847
2848
2849
2850
2851
2852
2853
2854
2855
2856
2857
2858
2859
2860
2861

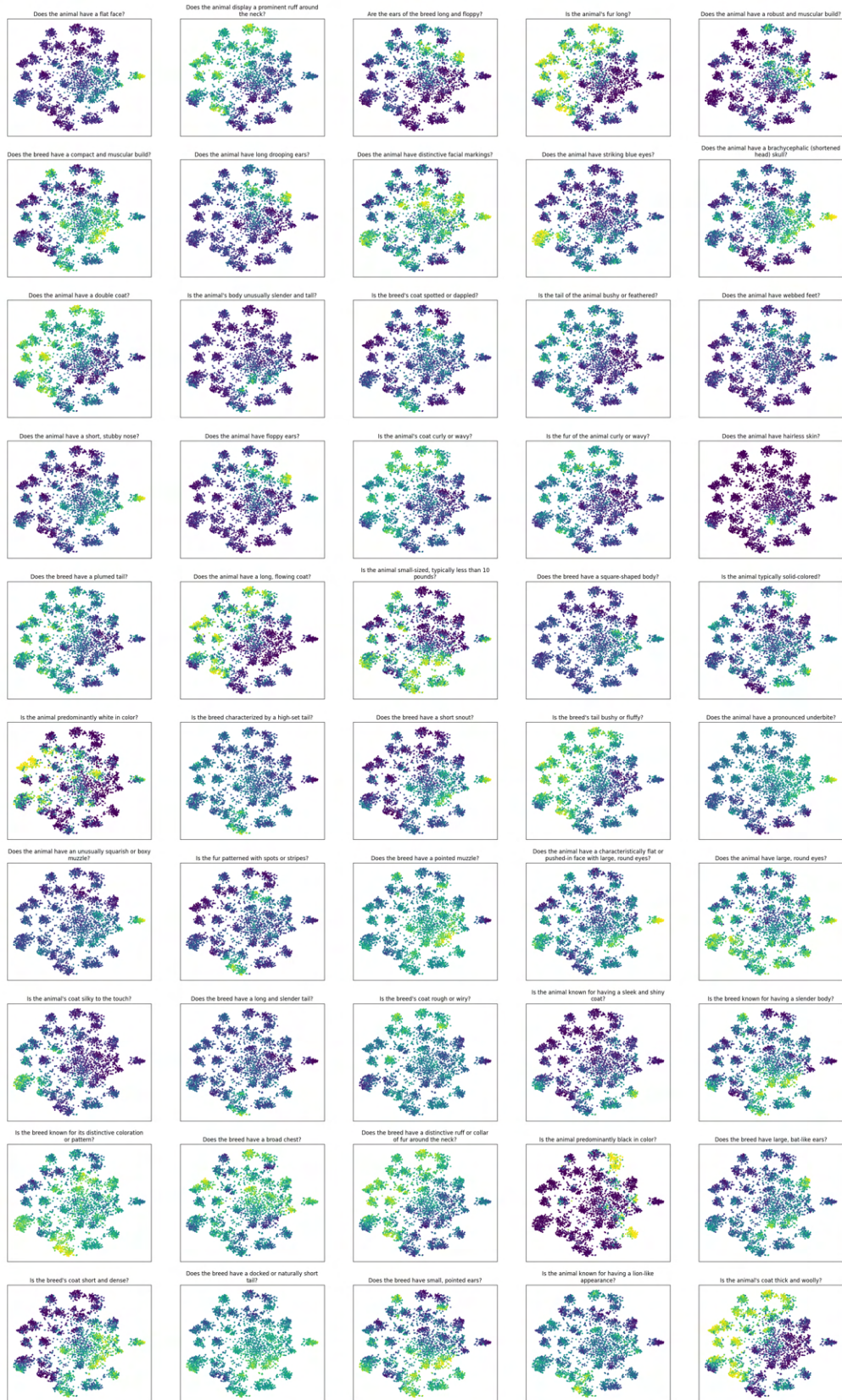


Figure 30: Visualization of predicted result t-SNE embeddings for the multi-aspect of the OxfordPets testing dataset.

2862
2863
2864
2865
2866
2867
2868
2869
2870
2871
2872
2873
2874
2875
2876
2877
2878
2879
2880
2881
2882
2883
2884
2885
2886
2887
2888
2889
2890
2891
2892
2893
2894
2895
2896
2897
2898
2899
2900
2901
2902
2903
2904
2905
2906
2907
2908
2909
2910
2911
2912
2913
2914
2915

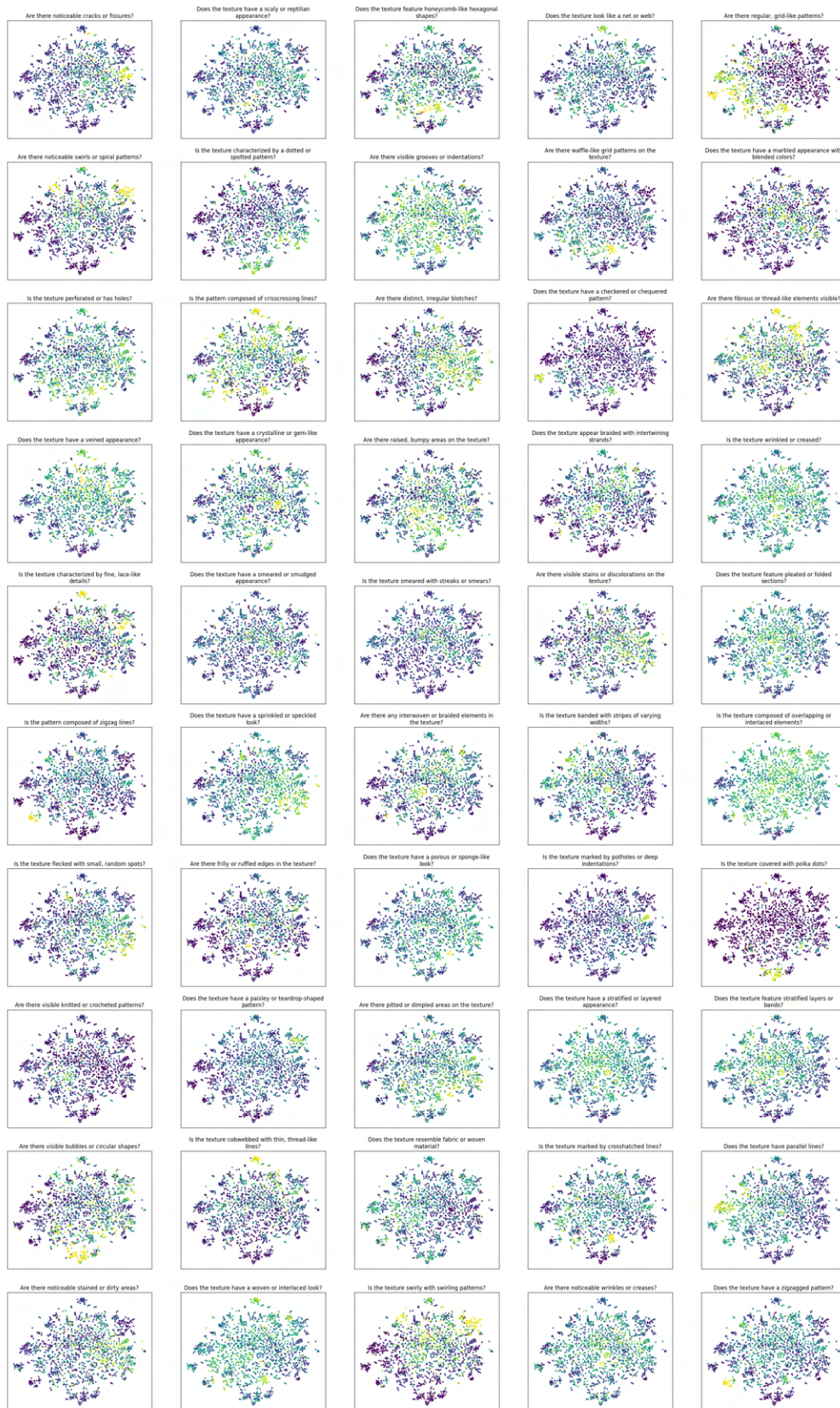


Figure 31: Visualization of ground truth t-SNE embeddings for the multi-aspect of the DTD testing dataset.

2916
 2917
 2918
 2919
 2920
 2921
 2922
 2923
 2924
 2925
 2926
 2927
 2928
 2929
 2930
 2931
 2932
 2933
 2934
 2935
 2936
 2937
 2938
 2939
 2940
 2941
 2942
 2943
 2944
 2945
 2946
 2947
 2948
 2949
 2950
 2951
 2952
 2953
 2954
 2955
 2956
 2957
 2958
 2959
 2960
 2961
 2962
 2963
 2964
 2965
 2966
 2967
 2968
 2969

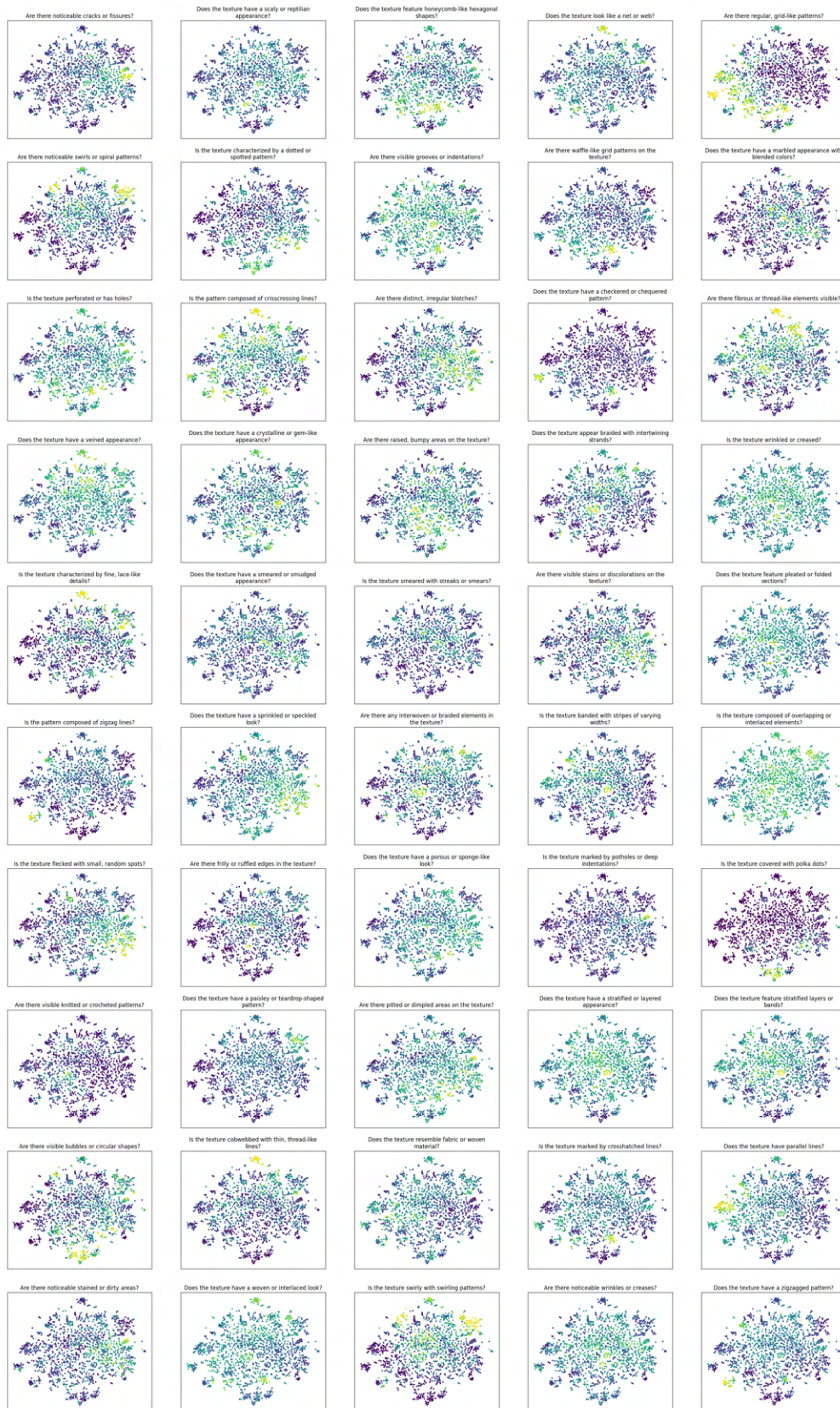


Figure 32: Visualization of predicted result t-SNE embeddings for the multi-aspect of the DTD testing dataset.

2970
2971
2972
2973
2974
2975
2976
2977
2978
2979
2980
2981
2982
2983
2984
2985
2986
2987
2988
2989
2990
2991
2992
2993
2994
2995
2996
2997
2998
2999
3000
3001
3002
3003
3004
3005
3006
3007
3008
3009
3010
3011
3012
3013
3014
3015
3016
3017
3018
3019
3020
3021
3022
3023



Figure 33: Visualization of ground truth t-SNE embeddings for the multi-aspect of the 102Flowers testing dataset.

3024
3025
3026
3027
3028
3029
3030
3031
3032
3033
3034
3035
3036
3037
3038
3039
3040
3041
3042
3043
3044
3045
3046
3047
3048
3049
3050
3051
3052
3053
3054
3055
3056
3057
3058
3059
3060
3061
3062
3063
3064
3065
3066
3067
3068
3069
3070
3071
3072
3073
3074
3075
3076
3077



Figure 34: Visualization of predicted result t-SNE embeddings for the multi-aspect of the 102Flowers testing dataset.

3078
 3079
 3080
 3081
 3082
 3083
 3084
 3085
 3086
 3087
 3088
 3089
 3090
 3091
 3092
 3093
 3094
 3095
 3096
 3097
 3098
 3099
 3100
 3101
 3102
 3103
 3104
 3105
 3106
 3107
 3108
 3109
 3110
 3111
 3112
 3113
 3114
 3115
 3116
 3117
 3118
 3119
 3120
 3121
 3122
 3123
 3124
 3125
 3126
 3127
 3128
 3129
 3130
 3131

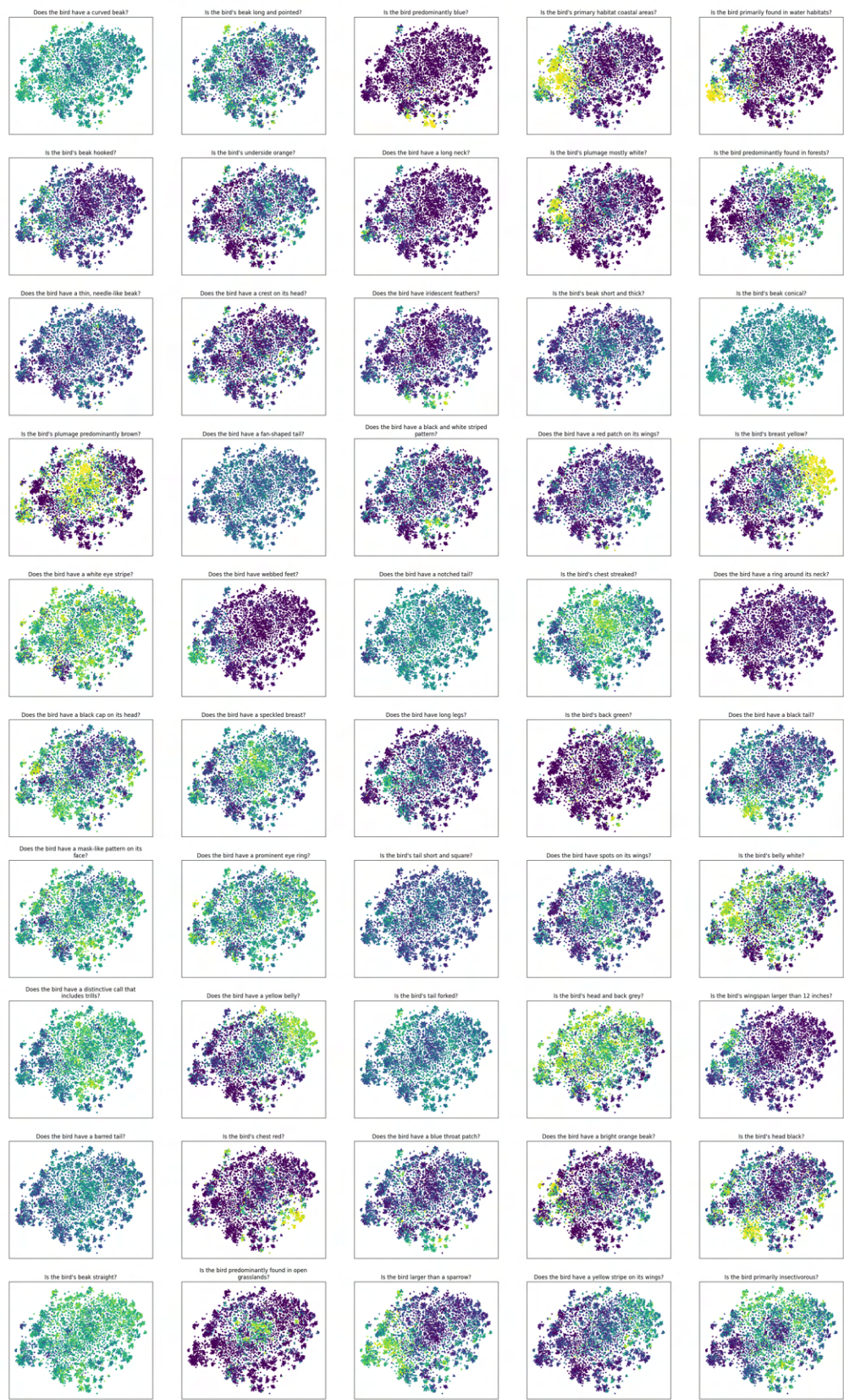


Figure 35: Visualization of ground truth t-SNE embeddings for the multi-aspect of the CUB200 testing dataset.

3132
3133
3134
3135
3136
3137
3138
3139
3140
3141
3142
3143
3144
3145
3146
3147
3148
3149
3150
3151
3152
3153
3154
3155
3156
3157
3158
3159
3160
3161
3162
3163
3164
3165
3166
3167
3168
3169
3170
3171
3172
3173
3174
3175
3176
3177
3178
3179
3180
3181
3182
3183
3184
3185

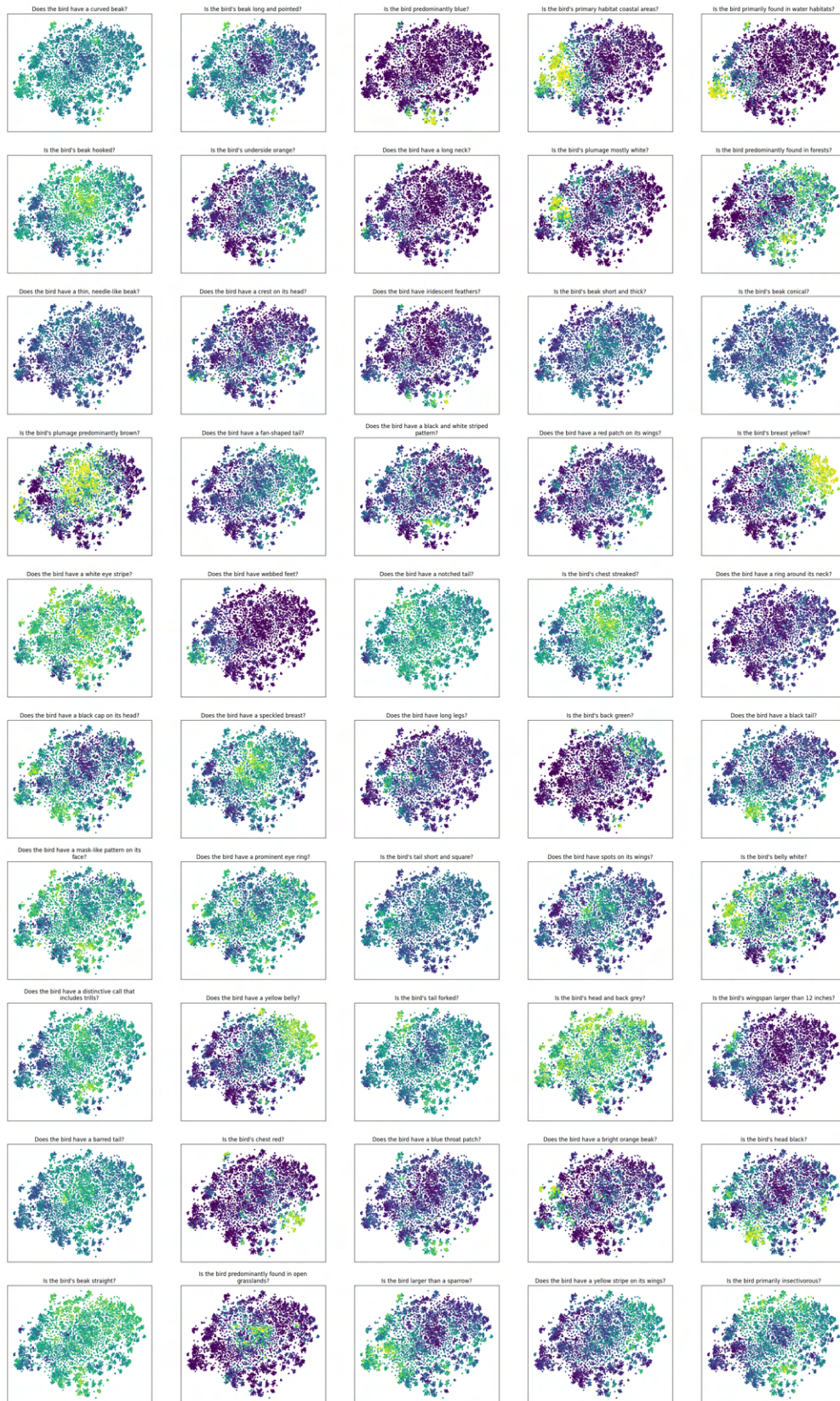


Figure 36: Visualization of predicted result t-SNE embeddings for the multi-aspect of the CUB200 testing dataset.

3186
3187
3188
3189
3190
3191
3192
3193
3194
3195
3196
3197
3198
3199
3200
3201
3202
3203
3204
3205
3206
3207
3208
3209
3210
3211
3212
3213
3214
3215
3216
3217
3218
3219
3220
3221
3222
3223
3224
3225
3226
3227
3228
3229
3230
3231
3232
3233
3234
3235
3236
3237
3238
3239

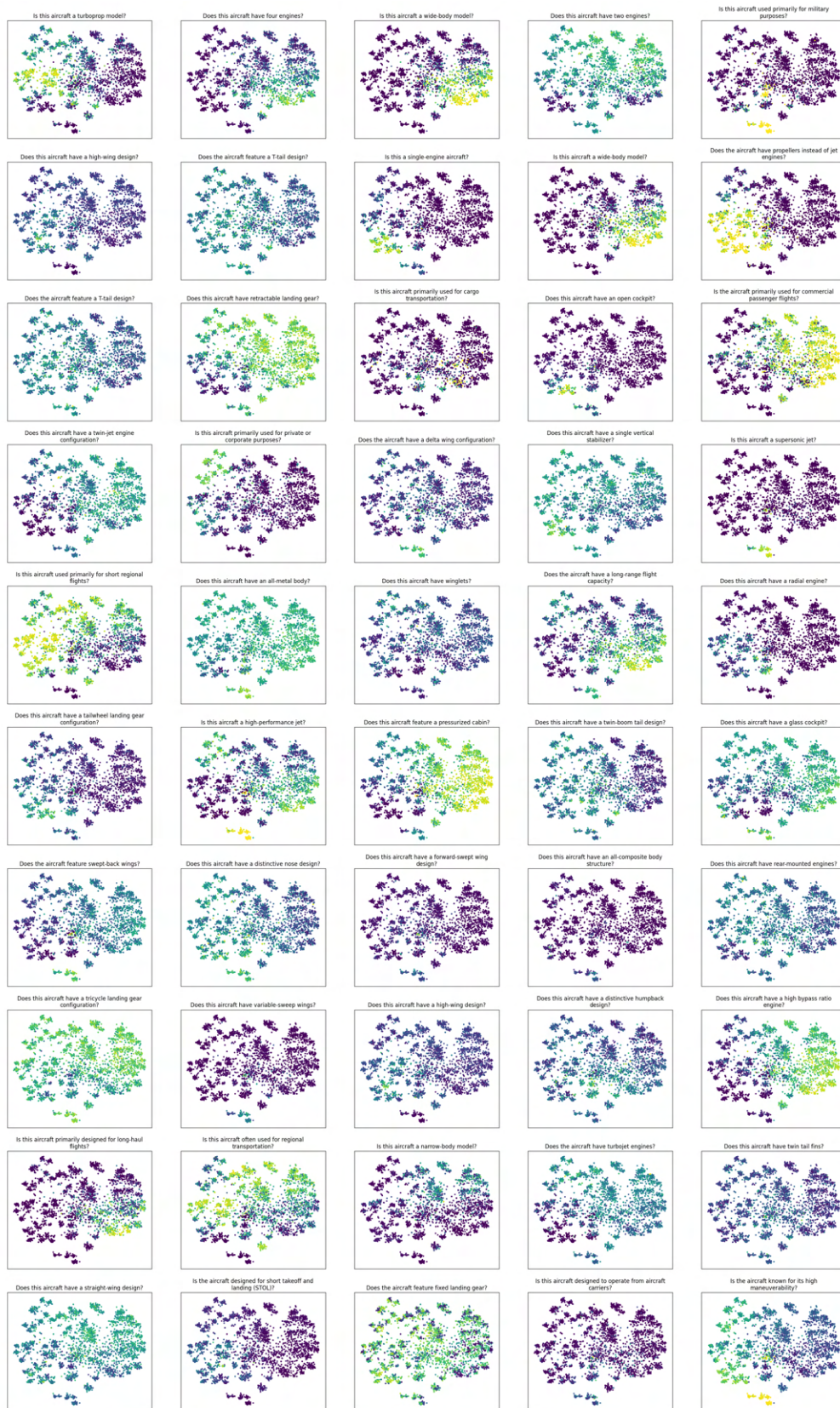


Figure 37: Visualization of ground truth t-SNE embeddings for the multi-aspect of the FGVC-Aircraft testing dataset.

3240
3241
3242
3243
3244
3245
3246
3247
3248
3249
3250
3251
3252
3253
3254
3255
3256
3257
3258
3259
3260
3261
3262
3263
3264
3265
3266
3267
3268
3269
3270
3271
3272
3273
3274
3275
3276
3277
3278
3279
3280
3281
3282
3283
3284
3285
3286
3287
3288
3289
3290
3291
3292
3293

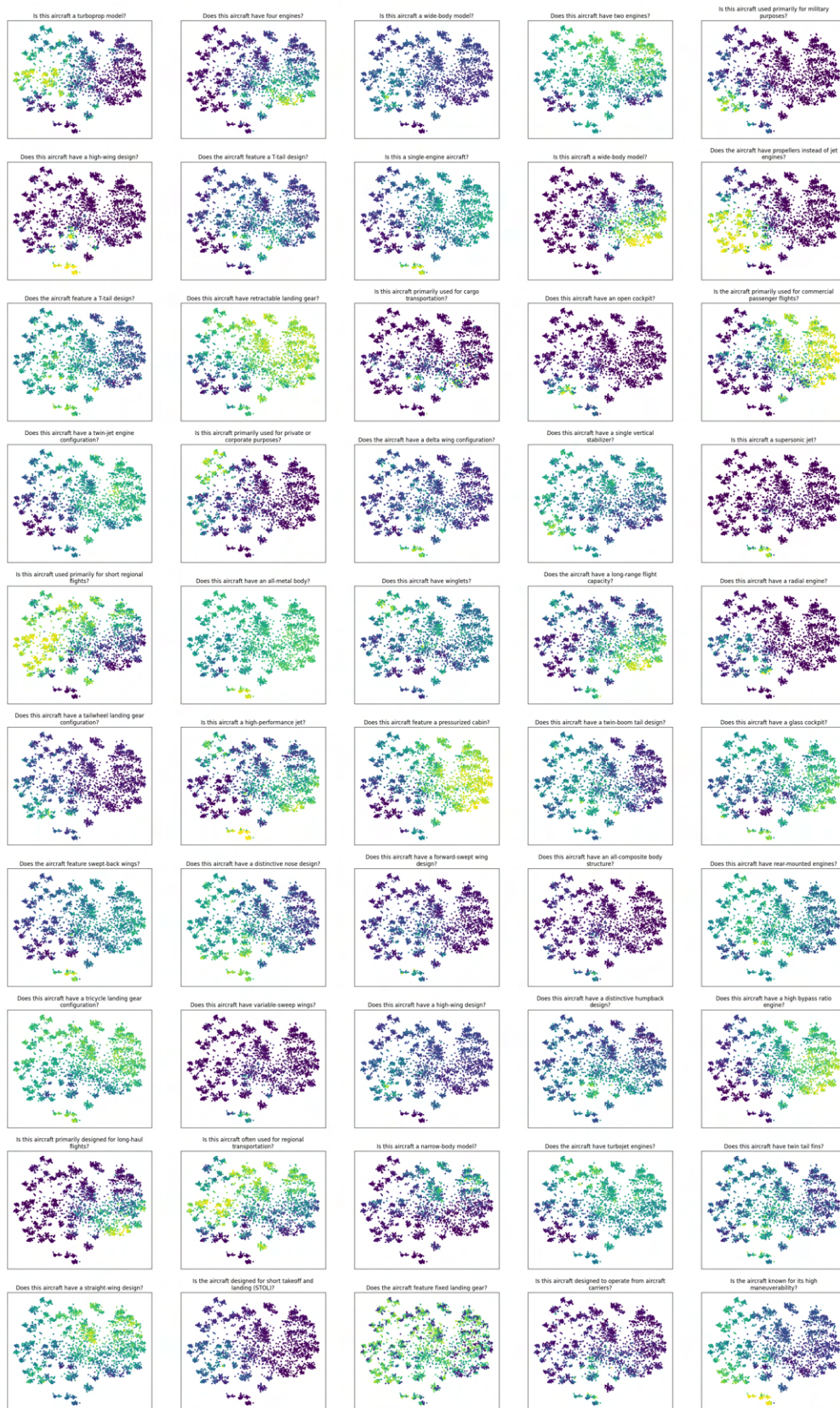


Figure 38: Visualization of predicted result t-SNE embeddings for the multi-aspect of the FGVC-Aircraft testing dataset.

3294
3295
3296
3297
3298
3299
3300
3301
3302
3303
3304
3305
3306
3307
3308
3309
3310
3311
3312
3313
3314
3315
3316
3317
3318
3319
3320
3321
3322
3323
3324
3325
3326
3327
3328
3329
3330
3331
3332
3333
3334
3335
3336
3337
3338
3339
3340
3341
3342
3343
3344
3345
3346
3347

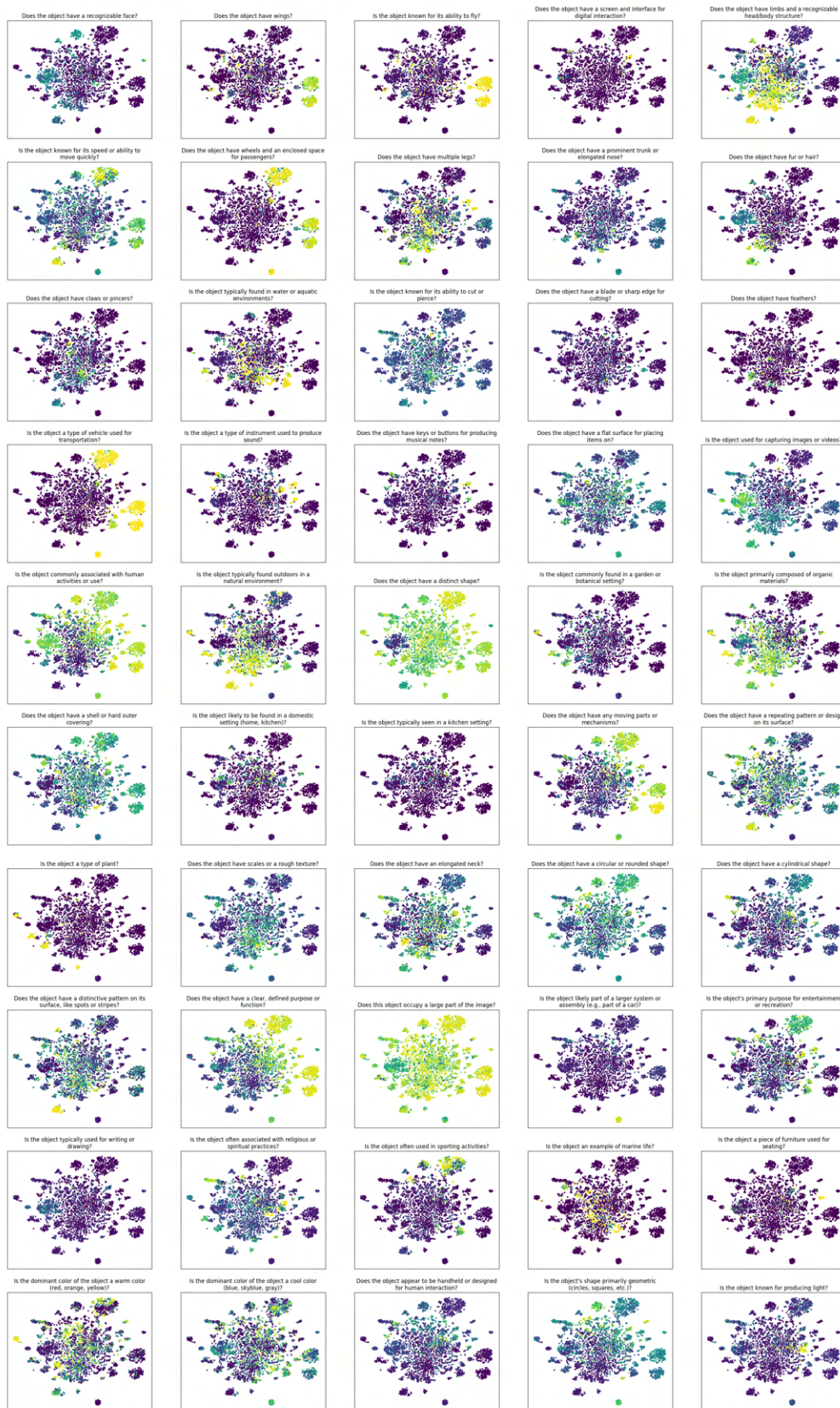


Figure 39: Visualization of ground truth t-SNE embeddings for the multi-aspect of the Caltech101 testing dataset.

3348
 3349
 3350
 3351
 3352
 3353
 3354
 3355
 3356
 3357
 3358
 3359
 3360
 3361
 3362
 3363
 3364
 3365
 3366
 3367
 3368
 3369
 3370
 3371
 3372
 3373
 3374
 3375
 3376
 3377
 3378
 3379
 3380
 3381
 3382
 3383
 3384
 3385
 3386
 3387
 3388
 3389
 3390
 3391
 3392
 3393
 3394
 3395
 3396
 3397
 3398
 3399
 3400
 3401

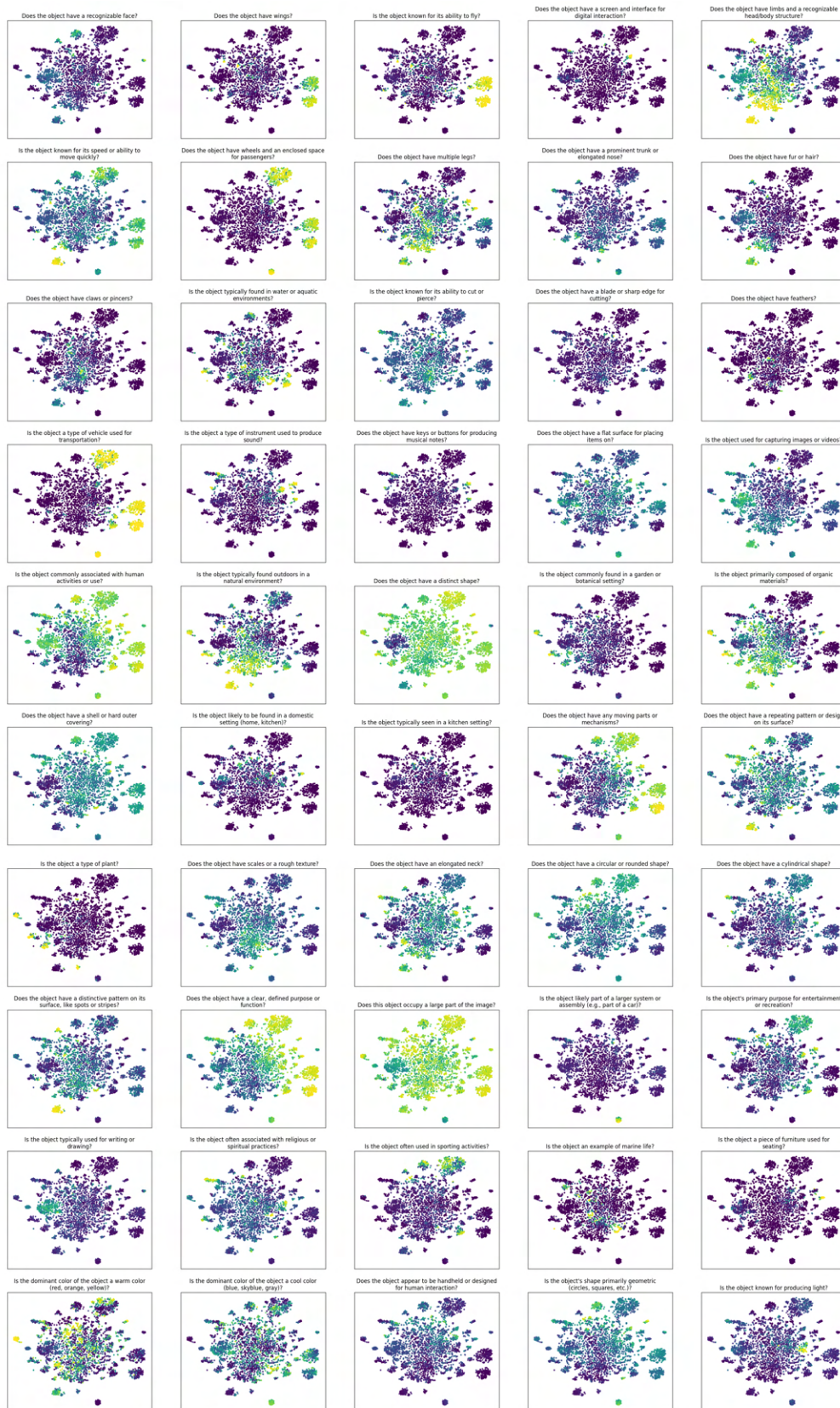


Figure 40: Visualization of predicted result t-SNE embeddings for the multi-aspect of the Caltech101 testing dataset.

3402
 3403
 3404
 3405
 3406
 3407
 3408
 3409
 3410
 3411
 3412
 3413
 3414
 3415
 3416
 3417
 3418
 3419
 3420
 3421
 3422
 3423
 3424
 3425
 3426
 3427
 3428
 3429
 3430
 3431
 3432
 3433
 3434
 3435
 3436
 3437
 3438
 3439
 3440
 3441
 3442
 3443
 3444
 3445
 3446
 3447
 3448
 3449
 3450
 3451
 3452
 3453
 3454
 3455

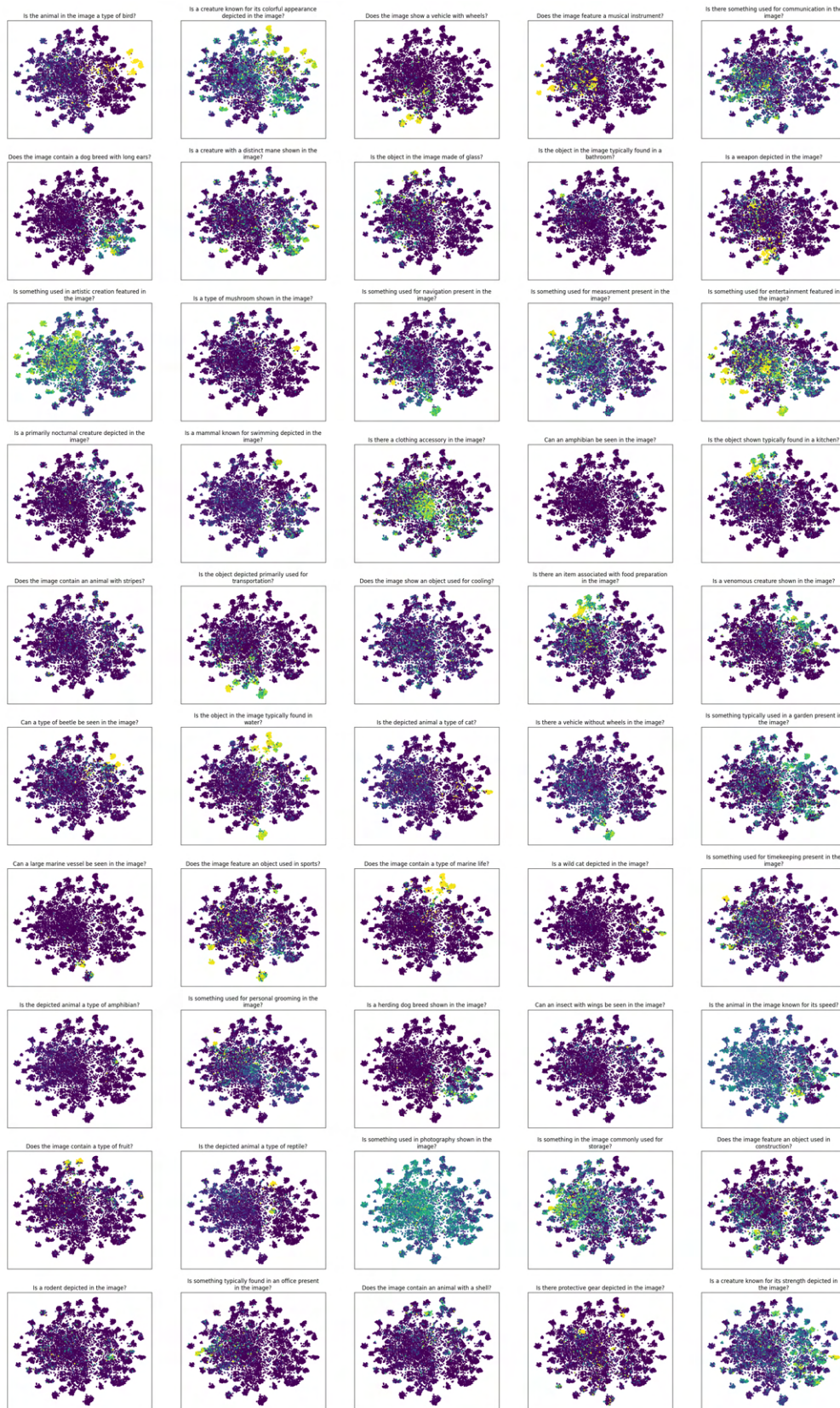


Figure 41: Visualization of ground truth t-SNE embeddings for the multi-aspect of the Mini-ImageNet testing dataset.

3456
3457
3458
3459
3460
3461
3462
3463
3464
3465
3466
3467
3468
3469
3470
3471
3472
3473
3474
3475
3476
3477
3478
3479
3480
3481
3482
3483
3484
3485
3486
3487
3488
3489
3490
3491
3492
3493
3494
3495
3496
3497
3498
3499
3500
3501
3502
3503
3504
3505
3506
3507
3508
3509

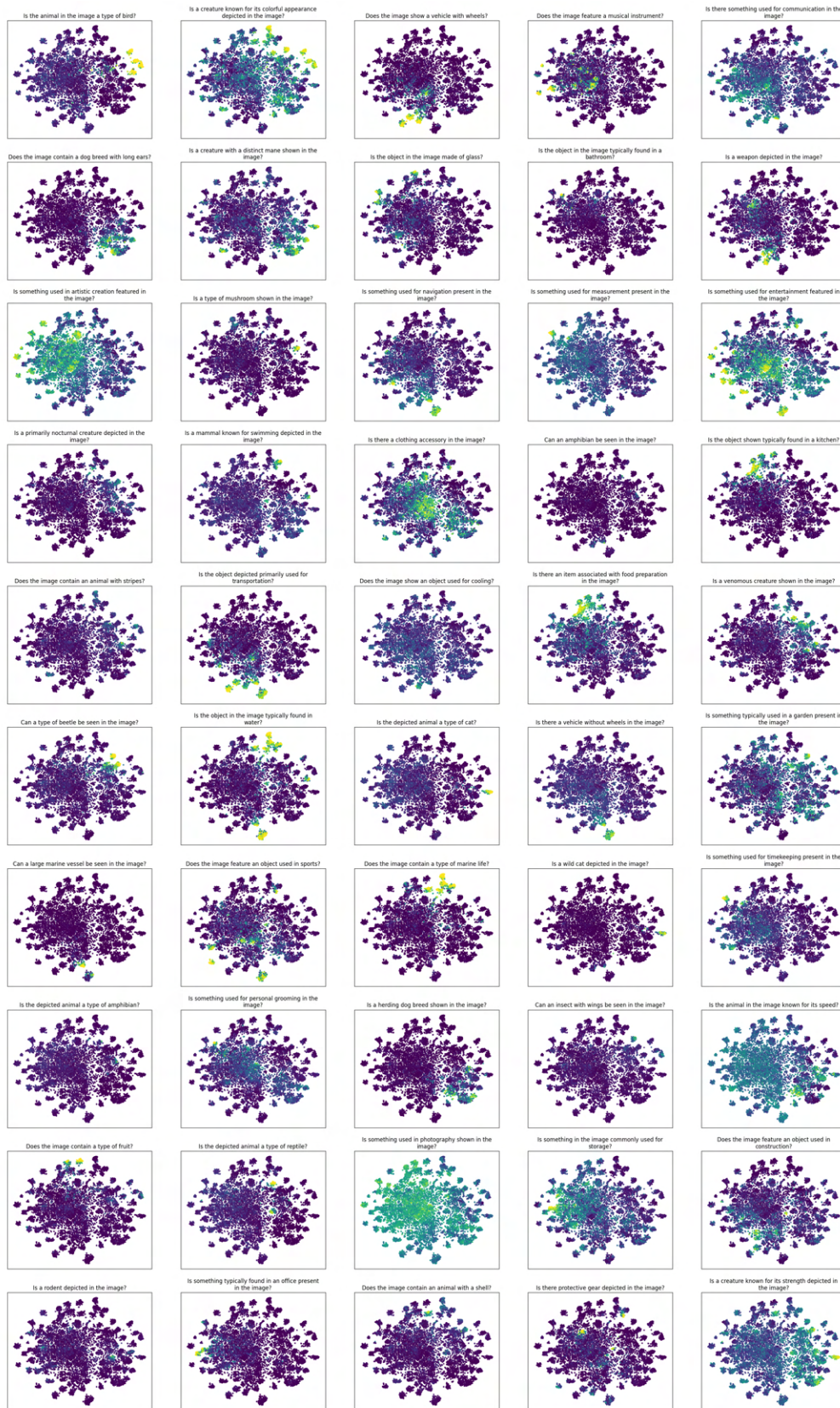


Figure 42: Visualization of predicted result t-SNE embeddings for the multi-aspect of the Mini-ImageNet testing dataset.