

APPENDIX: *What does a platypus look like?*

Generating customized prompts for zero-shot image classification

Overview

- A. CuPL (Full Prompts) vs Standard Prompts**
- B. CuPL Base Prompts**
- C. Evaluation Metric**
- D. Comparison with Definitions and Wikipedia Descriptions**
- E. Ensembling with Standard Prompts**
- F. Additional Qualitative Results**
- G. Open-Source LLM and ChatGPT**
- H. Single Sentence Baseline**
- I. Robustness**
- J. Tagging with ImageNet Parts details**
- K. CuPL Improvement Analysis**
- L. Error Analysis**
- M. Image-Prompt Distribution**
- N. Temperature Analysis**
- O. Example Generated image-prompts**

A. CuPL (Full Prompts) vs Standard Prompts

We detail the hand-written prompt templates used for CuPL (full prompts) versus standard CLIP [42] prompt templates. For CuPL, hand-written prompt templates are needed for the LLM-prompts, while for the standard method hand-written prompt templates are needed for the image-prompts.

Note that many of the hand-written templates for the standard method encode information about the datasets. For example, "a toy {}" demonstrates knowledge that objects are sometimes represented as a toy version of an object rather than as the literal object. CuPL prompts remain much more general (e.g. "Describe what a {} looks like").

Caltech101

CuPL hand-written	Standard hand-written	
Describe what a(n) {} looks like:	a photo of a {}.	a photo of the {}.
Describe a(n) {}:	a painting of a {}.	a painting of the {}.
What are the identifying characteristics of a(n) {}?	a plastic {}.	the plastic {}.
	a sculpture of a {}.	a sculpture of the {}.
	a sketch of a {}.	a sketch of the {}.
	a tattoo of a {}.	a tattoo of the {}.
	a toy {}.	the toy {}.
	a rendition of a {}.	a rendition of the {}.
	a embroidered {}.	the embroidered {}.
	a cartoon {}.	the cartoon {}.
	a {} in a video game.	the {} in a video game.
	a plushie {}.	the plushie {}.
	a origami {}.	the origami {}.
	art of a {}.	art of the {}.
	graffiti of a {}.	graffiti of the {}.
	a drawing of a {}.	a drawing of the {}.
	a doodle of a {}.	a doodle of the {}.

Food101

CuPL hand-written

Describe what {} looks like
Visually describe {}
How can you tell that the food in this photo is {}?

Standard hand-written

a photo of {}, a type of food.

Stanford Cars

CuPL hand-written

How can you identify a(n) {}?
Description of a(n) {}, a type of car.
A caption of a photo of a(n) {}:
What are the primary characteristics of a(n) {}?
Description of the exterior of a(n) {}
What are the identifying characteristics of a(n) {}, a type of car?
Describe an image from the internet of a(n) {}
What does a(n) {} look like?
Describe what a(n) {}, a type of car, looks like

Standard hand-written

a photo of a {}.
a photo of the {}.
a photo of my {}.
i love my {}!
a photo of my dirty {}.
a photo of my clean {}.
a photo of my new {}.
a photo of my old {}.

Oxford Pets

CuPL hand-written

Describe what a pet {} looks like
Visually describe a(n) '{}', a type of pet.

Standard hand-written

a photo of a {}, a type of pet.

ImageNet

CuPL hand-written	Standard hand-written
Describe what a(n) {} looks like How can you identify a(n) {}? What does a(n) look like? Describe an image from the internet of a(n) {} A caption of an image of a(n) {}:	a bad photo of a {}. a photo of many {}. a sculpture of a {}. a photo of the hard to see {}. a low resolution photo of the {}. a rendering of a {}. graffiti of a {}. a bad photo of the {}. a cropped photo of the {}. a tattoo of a {}. the embroidered {}. a photo of a hard to see {}. a bright photo of a {}. a photo of a clean {}. a photo of a dirty {}. a dark photo of the {}. a drawing of a {}. a photo of my {}. the plastic {}. a photo of the cool {}. a close-up photo of a {}. a black and white photo of the {}. a painting of the {}. a painting of a {}. a pixelated photo of the {}. a sculpture of the {}. a bright photo of the {}. a cropped photo of a {}. a plastic {}. a photo of the dirty {}. a jpeg corrupted photo of a {}. a blurry photo of the {}. a photo of the {}. a good photo of the {}. a rendering of the {}. a {} in a video game. a photo of one {}. a doodle of a {}. a close-up photo of the {}. a photo of a {} the origami {}. the {} in a video game. a sketch of a {}. a doodle of the {}. a origami {}. a low resolution photo of a {}. the toy {}. a rendition of the {}. a photo of the clean {}. a photo of a large {}. a rendition of a {}. a photo of a nice {}. a photo of a weird {}. a blurry photo of a {}. a cartoon {}. art of a {}. a sketch of the {}. a embroidered {}. a pixelated photo of a {}. itap of the {}. a jpeg corrupted photo of the {}. a good photo of a {}. a plushie {}. a photo of the nice {}. a photo of the small {}. a photo of the weird {}. the cartoon {}. art of the {}. a drawing of the {}. a photo of the large {}. a black and white photo of a {}. the plushie {}. a dark photo of a {}. itap of a {}. graffiti of the {}. a toy {}. itap of my {}. a photo of a cool {}. a photo of a small {}. a tattoo of the {}.

FGVC Aircraft

CuPL hand-written	Standard hand-written
Describe a(n) {} aircraft Describe the {} aircraft	a photo of a {}, a type of aircraft. a photo of the {}, a type of aircraft.

DTD

CuPL hand-written	Standard hand-written
What does “{}” material look like? What does a “{}” surface look like? What does a “{}” texture look like? What does a “{}” object look like? What does a “{}” thing look like? What does a “{}” pattern look like?	a photo of a {} texture. a photo of a {} pattern. a photo of a {} thing. a photo of a {} object. a photo of the {} texture. a photo of the {} pattern. a photo of the {} thing. a photo of the {} object.

SUN397

CuPL hand-written	Standard hand-written
Describe what a(n) {} looks like How can you identify a(n) {}? Describe a photo of a(n) {}	a photo of a {}. a photo of the {}.

Kinetics-700

CuPL hand-written

Describe the action "{}"
What does a person {} look like?
What does the act of {} look like?
Describe "{}"

Standard hand-written

a photo of a person {}.
a photo of a person using {}.
a photo of a person doing {}.
a photo of a person during {}.
a photo of a person performing {}.
a photo of a person practicing {}.
a video of {}.
a video of a person {}.
a video of a person using {}.
a video of a person doing {}.
a video of a person during {}.
a video of a person performing {}.
a video of a person practicing {}.
an example of {}.
an example of a person {}.
an example of a person using {}.
an example of a person doing {}.
an example of a person during {}.
an example of a person performing {}.
an example of a person practicing {}.
a demonstration of {}.
a demonstration of a person {}.
a demonstration of a person using {}.
a demonstration of a person doing {}.
a demonstration of a person during {}.
a demonstration of a person performing {}.
a demonstration of a person practicing {}.

UCF 101

CuPL hand-written

Describe the action of {}
What does the act of {} look like?
What does a person doing {} look like?
Describe “{}”
Describe the action “{}”

Standard hand-written

a photo of a person {}.
a video of a person {}.
a example of a person {}.
a demonstration of a person {}.
a photo of the person {}.
a video of the person {}.
a example of the person {}.
a demonstration of the person {}.
a photo of a person using {}.
a video of a person using {}.
a example of a person using {}.
a demonstration of a person using {}.
a photo of the person using {}.
a video of the person using {}.
a example of the person using {}.
a demonstration of the person using {}.
a photo of a person doing {}.
a video of a person doing {}.
a example of a person doing {}.
a demonstration of a person doing {}.
a photo of the person doing {}.
a video of the person doing {}.
a example of the person doing {}.
a demonstration of the person doing {}.
a photo of a person during {}.
a video of a person during {}.
a example of a person during {}.
a demonstration of a person during {}.
a photo of the person during {}.
a video of the person during {}.
a example of the person during {}.
a demonstration of the person during {}.
a photo of a person performing {}.
a video of a person performing {}.
a example of a person performing {}.
a demonstration of a person performing {}.
a photo of the person performing {}.
a video of the person performing {}.
a example of the person performing {}.
a demonstration of the person performing {}.
a photo of a person practicing {}.
a video of a person practicing {}.
a example of a person practicing {}.
a demonstration of a person practicing {}.
a photo of the person practicing {}.
a video of the person practicing {}.
a example of the person practicing {}.
a demonstration of the person practicing {}.

RESISC 45

CuPL hand-written

Describe a satellite photo of a(n) {}
Describe a(n) {} as it would appear in an aerial image
How can you identify a(n) {} in an aerial photo?
Describe the satellite photo of a(n) {}
Describe an aerial photo of a(n) {}

Standard hand-written

satellite imagery of {}.
aerial imagery of {}.
satellite photo of {}.
aerial photo of {}.
satellite view of {}.
aerial view of {}.
satellite imagery of a {}.
aerial imagery of a {}.
satellite photo of a {}.
aerial photo of a {}.
satellite view of a {}.
aerial view of a {}.
satellite imagery of the {}.
aerial imagery of the {}.
satellite photo of the {}.
aerial photo of the {}.
satellite view of the {}.
aerial view of the {}.

Birdsnap

CuPL hand-written

Describe what the bird {} looks like:
Describe the bird {}:
What are the identifying characteristics of the bird {}?

Standard hand-written

a photo of a {}, a type of bird.

Flowers 102

CuPL hand-written

Describe how to identify a(n) {}, a type of flower
What does a(n) {} flower look like?

Standard hand-written

a photo of a {}, a type of flower.'

CIFAR-10

CuPL hand-written

Describe what a(n) {} looks like
Describe a(n) {}:
What are the identifying characteristics of a(n) {}?

Standard hand-written

a photo of a {}.
a blurry photo of a {}.
a black and white photo of a {}.
a low contrast photo of a {}.
a high contrast photo of a {}.
a bad photo of a {}.
a good photo of a {}.
a photo of a small {}.
a photo of a big {}.
a photo of the {}.
a blurry photo of the {}.
a black and white photo of the {}.
a low contrast photo of the {}.
a high contrast photo of the {}.
a bad photo of the {}.
a good photo of the {}.
a photo of the small {}.
a photo of the big {}.

CIFAR-100

CuPL hand-written

Describe a photo of a(n) {}:
What are the identifying characteristics of a(n) {}?
Describe what a(n) {} looks like:
Describe a(n) {}:

Standard hand-written

a photo of a {}.
a blurry photo of a {}.
a black and white photo of a {}.
a low contrast photo of a {}.
a high contrast photo of a {}.
a bad photo of a {}.
a good photo of a {}.
a photo of a small {}.
a photo of a big {}.
a photo of the {}.
a blurry photo of the {}.
a black and white photo of the {}.
a low contrast photo of the {}.
a high contrast photo of the {}.
a bad photo of the {}.
a good photo of the {}.
a photo of the small {}.
a photo of the big {}.

B. CuPL Base Prompts

The three general sentences used in the base prompt setting are:

Describe what a/the ___ looks like:
 Describe a/the ___ :
 What are the identifying characteristics of a/the ___ ?

Here we specify the type filled in for each of the examined datasets, as well as the article used for that dataset (‘a(n)’ or ‘the’):

Dataset	Base LLM-prompt type specification
ImageNet	a(n) {}
DTD	the texture {}
StanfordCars	the car {}
SUN397	a(n) {}
Food 101	the food {}
FGVC Aircraft	the aircraft {}
Oxford Pets	a pet {}
Caltech101	a(n) {}
CIFAR-10	a(n) {}
CIFAR-100	a(n) {}
Flowers 102	the flower {}
Kinetics-700	the action of {}
UCF101	the action of {}
RESISC45	a satellite photo of {}
Birdsnap	the bird {}

C. Evaluation Metric

ImageNet	DTD	Stanford Cars	SUN397	Food101	FGVC Aircraft	Oxford Pets	Caltech101	Flowers 102	UCF101	Kinetics-700	RESISC45	CIFAR-10	CIFAR-100	Birdsnap
Acc.	Acc.	Acc.	Acc.	Acc.	Mean per class	Mean per class	Mean per class	Mean per class	Acc.	Mean (top1, top5)	Acc.	Acc.	Acc.	Acc.

D. Comparison with Definitions and Wikipedia Descriptions

WordNet Definitions and Wikipedia Descriptions. We also consider two additional methods of obtaining descriptive sentences for each ImageNet category, other than using an LLM. Firstly, we compare CuPL (full) image-prompts with image-prompts generated using definitions of each ImageNet category. Because each ImageNet category is derived from the WordNet database [34], we can use the WordNet definition of each word.

We preprocess these definitions so they are of the form “A(n) {} is a ...” as not all WordNet definitions contain the name of the word itself. We also add a period to the end of each definition, as we find this increases performance. As shown in Table 3, ImageNet Top-1 accuracy with WordNet definition prompts is below that of CuPL or standard prompts. In addition to lower accuracy, this method uses significantly more hand-constructed sentences as it requires 1000 unique hand-written definitions compared to 175 unique hand-written image-prompt templates for the standard method and 45 unique hand-written LLM-prompt templates for CuPL (full).

Secondly, we compare against prompts generated from Wikipedia articles corresponding to each ImageNet category, as collected in [7]. Note that the Wikipedia article does not always exactly match the natural language name of the class used

Standard	CuPL	WordNet	Wiki
75.54	76.69	73.44	68.20

Table 3. **ImageNet Top-1 accuracy for different methods of generating image-prompts.**

by [42]. Additionally, 80 categories map to more than one Wikipedia article (e.g. the category associated with the natural language word “patio” is mapped to the articles for “patio” and “terrace”). In this case, we select the first associated article. We preprocess these by removing the first line (the name of the article), and then extracting the first sentence, including the final period. We find both of these preprocessing steps lead to increase in accuracy. We also truncate this sentence to the maximum allowed input length of CLIP. For the 24 ImageNet categories that do not have an associated Wikipedia page, we use the name of the category as the image-prompts. As shown in Table 3, we find Wikipedia to be less effective than standard prompts, CuPL prompts, or WordNet definitions.

E. Ensembling with Standard Prompts.

We also consider using LLM generated prompts from CuPL (full) in addition to hand-written prompts. We do this by averaging together all the text embeddings of the CuPL prompts and hand-written prompts. As shown in Table 4, we find that for some datasets, ensembling both types of prompts outperforms CuPL prompts on their own, while for others CuPL prompts perform better. For all datasets, this ensemble performs better than standard prompts alone. However, this ensembling method requires all the hand-written effort and domain knowledge of the standard approach.

	ImageNet	DTD	Stanford Cars	SUN397	Food101	FGVC Aircraft	Oxford Pets	Caltech101	Flowers 102	UCF101	Kinetics-700	RESISC45	CIFAR-10	CIFAR-100	Birdsnap	mean
Ensemble	76.51	61.60	77.66	73.51	93.42	36.47	93.71	93.87	79.73	78.16	61.50	73.03	95.88	79.33	51.09	75.03
Δ std	+0.97	+6.40	+0.13	+4.20	+0.34	+3.59	+0.38	+0.63	+1.20	+0.71	+1.43	+1.93	+0.29	+1.07	+0.66	
Δ CuPL	-0.18	-0.1	+0.03	+0.20	+0.06	+0.36	-0.10	+0.42	+0.06	+0.20	+0.87	+1.34	+0.04	+0.76	-0.02	

Table 4. **Performance of the ensemble of CuPL (full) and the standard, hand-written prompts in CLIP [42].** This ensemble outperforms the standard hand-written prompts for all examined datasets (difference shown with Δ std), and outperforms CuPL (full) for 11 datasets (difference shown with Δ CuPL)

F. Additional Qualitative Results

In addition to Figure 5, we present additional qualitative results (Figure 7), demonstrating the difference in importance between regions when using CuPL prompts versus baseline prompts. As in Figure 5, we find that CuPL prompts allow the model to place more importance on semantically relevant regions of the image when distinguishing between two visually similar categories.

G. Open-Source LLMs and ChatGPT

While GPT-3 [5] demonstrates higher performance on a number of tasks compared to smaller open-source models, open-source models are sometime more accessible. We therefore show improvement using GPT-J-6B [55], a small open-source model available on HuggingFace [58]. We find that we are able to surpass human written prompts with prompts generated by this model as shown in Table 5, though we still fall short of those generated by GPT-3. Additionally, we employ a number of strategies to increase the accuracy of the lower quality GPT-J-6B generations. First, we generate at a lower

	ImageNet Acc
standard	75.54
CuPL (GPT-J-6B)	75.62
CuPL (ChatGPT)	76.09
CuPL (GPT-3)	76.69

Table 5. **CuPL with an open-source model and ChatGPT.** CuPL is able to improve over hand-written baselines even for smaller models.









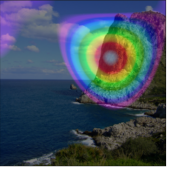





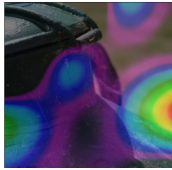



Original Image	Region Importance with CuPL Prompts		Region Importance with Baseline Prompts		Example Image of Distractor Class
	(A)	(B)	(C)	(D)	
GT Label: Eastern Diamondback Rattlesnake	Prompt eastern diamondback: "A eastern diamondback rattlesnake is a snake with a diamond-shaped pattern on its back ."	Prompt sidewinder rattlesnake: "A sidewinder rattlesnake is a type of rattlesnake that has a unique method of locomotion."	Prompt eastern diamondback: A photo of a eastern diamondback rattlesnake.	Prompt sidewinder rattlesnake: A photo of a sidewinder rattlesnake.	Example: Sidewinder Rattlesnake
					
	Prediction: Eastern Diamondback ✓		Prediction: Sidewinder ✗		
GT Label: Promontory	Prompt promontory: "A promontory is a landform that protrudes into a body of water ."	Prompt cliff: "A cliff is a high, steep rock face or slope."	Prompt promontory: A photo of a promontory.	Prompt cliff: A photo of a cliff.	Example: Cliff
					
	Prediction: Promontory ✓		Prediction: Cliff ✗		
GT Label: Milk Can	Prompt milk can: "A milk can looks like a large metal can with a lid ."	Prompt water jug: "A water jug is typically a plastic or glass container with a spout and a handle."	Prompt milk can: A photo of a milk can.	Prompt water jug: A photo of a water jug.	Example: Water Jug
					
	Prediction: Milk Can ✓		Prediction: Water Jug ✗		

Figure 7. More examples as in Figure 5

temperature (0.3) to prevent irrelevant or nonsensical generations, which we find occur more frequently in smaller models. Additionally, we generate 5 times more Image-prompts per LLM-prompt than we do when using GPT-3 [5]. We also add punctuation to the end of all LLM-prompts to encourage the LLM to begin new sentences. Finally, we filter out any Image-prompts which do not contain the name of the ImageNet category they are meant to describe, as well as remove a number of unicode characters from the generations (e.g. ‘u2019’). We present these findings as a way to make CuPL a more accessible options until large high performance models become available to the public.

Additionally, we present CuPL with prompts generated with ChatGPT, a popular chat model with a similar design to GPT-3 [5]. This model is less expensive to run, 10 times cheaper than Davinci-002. As shown in Table 5, it outperforms hand-written prompts by 0.55%. We do modify the original LLM prompts (i.e. “What does a {} look like?”), by formatting it as follows: “Based on what you know as an AI Language model, answer the following question in one sentence:” + LLM prompt + “. Do not mention that you are a language model.” This is because we found that ChatGPT often mentions that it is an AI language model before answering the question or sometimes refused to answer the question. Even with this prompt addition, there were a number of generations mentioning that it is an AI model or refusing to answer, so we filtered out any

responses that contained the words “AI” or “sorry”. We use all the same hyperparameters as the original CuPL, except with max context length increased to 100. We present this as an inexpensive alternative to CuPL full as we were able to generate all ImageNet prompts with ChatGPT for less than \$5 USD.

H. Single Sentence Baseline

	ImageNet Acc
a photo of a {}	73.46
CuPL (1 hand-written)	75.71
CuPL (1 generated)	74.24

Table 6. **Single sentence baselines.** Comparison of a single hand-written Image-prompt template with a single hand-written LLM-prompt as well as a single CuPL generated Image-prompt.

like’. For this experiment, we generate at a temperature of 0.3 as we find that higher temperatures are only helpful when we are able to ensemble many diverse prompts, not when we are limited to one. We find that CuPL outperforms a single hand-written template under both of these settings, as shown in Table 6.

One of the primary benefits of CuPL is that it decreases the amount of necessary hand-engineering. However, this could also be done by decreasing the number of total hand-written templates used, which comes at a loss in performance. We present this as a ‘low effort’ baseline, where we use only the hand constructed template of ‘a photo of a {}’. We compare this with two CuPL baselines. The first is the baseline in which we also only construct one hand written template: ‘Describe what a {} looks like’. We then use this to generate 10 Image-prompts. The second baseline is a single CuPL generated sentence, generated with the prompt ‘Describe what a {} looks

I. Robustness

In addition to the previously mentioned benefits of open-vocabulary models, one of the important advances made by CLIP [42] is an increased robustness on out-of-distribution data. Fine-tuning has been shown to degrade performance on out-of-distribution tasks [59], however zero-shot CLIP is robust to these distribution shifts. We show improvement on two common distribution shifts in Table 7, demonstrating that CuPL maintains the robustness of CLIP.

	ImageNet [13]	ImageNet-V2 [44]	ImageNet-Sketch [56]
Standard	75.54	69.86	59.60
CuPL	76.69	70.85	60.05

Table 7. **Robustness of CuPL.** CuPL accuracy on two common ImageNet variants, using CuPL ImageNet Image-prompts. CuPL improves performance on both of these variants, demonstrating that CuPL improves accuracy on in-distribution tasks, while maintaining robustness to distribution shifts.

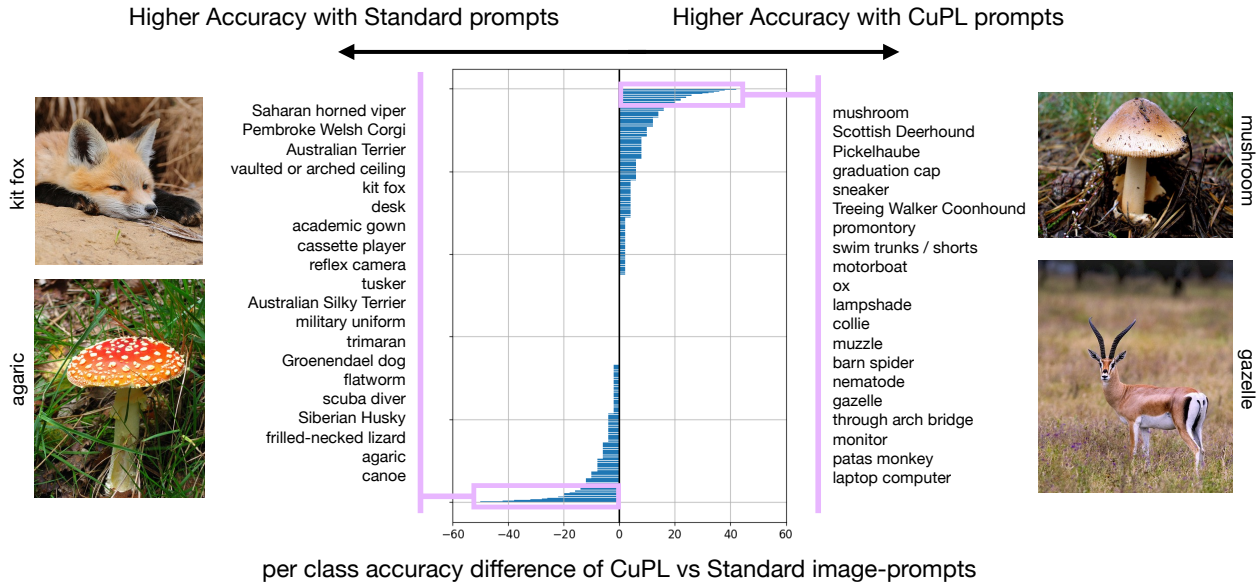
J. Tagging with ImageNet Parts details

As demonstrated in Figure 6, we find a strong correlation between the number of mentions of parts of ImageNet classes (‘head’, ‘body’, ‘legs’, etc) in the CuPL prompts and the importance of those parts visually. In this section we detail how we tag these parts in the text of CuPL prompts. The reason that this is difficult is because we cannot simply count the instances of the word ‘head’ to understand the importance of the head in identifying a class. This is because the head contains many other parts, and these may be mentioned instead. For example the face may be important for identification. This is within the ‘head’ segmentation of the PartImageNet [18] category, so should be counted as a mention of the head. To account for this, we take advantage of WordNet [34] which annotates the parts of various objects. Therefore we can use WordNet to obtain a list of parts for each category in PartImageNet. It is also important to note that while ‘face’ is a part of ‘head’, there are also subparts of face. We consider 2 layers of depth to gather parts for a given category.

K. CuPL Improvement Analysis

K.1. Per Class Accuracy CuPL versus Baseline

In addition to total accuracy gains, we present the per class accuracy shift between the image-prompts used in [42] and CuPL, shown in Figure 8. As demonstrated, the accuracy gains seen in CuPL are not distributed uniformly through the ImageNet classes, with some classes seeing ~40 percentage point accuracy gains, and others seeing ~40 percentage point



per class accuracy difference of CuPL vs Standard image-prompts

Figure 8. **Distribution of ImageNet per class accuracy difference of CuPL image-prompts versus standard image-prompts.** As shown, the accuracy gains of CuPL are not uniform across all classes. Rather, we see large gains for some classes, and losses for others. In addition, we list the classes which see the largest accuracy gains when switching to CuPL prompts (with “mushroom” having the largest gain), and the 20 classes with the largest accuracy losses (with “canoe” having the largest loss.)

accuracy losses when compared against class accuracy with standard prompts. In other words, while CuPL sees a higher accuracy overall when compared to the standard method, the images which are correctly predicted by the standard prompts are not a subset of the images which are correctly predicted by CuPL. In fact CuPL sees just over a 1 percentage point gain when compared to standard prompts, but differs in its predictions from the standard method for 11.50% of predictions (with CuPL correct for 4.48% of these, standard correct for 3.32%, and neither correct for 3.70%).

Figure 8 also shows the classes with the 20 greatest accuracy gains and losses when comparing class accuracy with standard image-prompts [42] versus with CuPL image-prompts. Interestingly, for many of the classes which see a large accuracy gain, we see a corresponding class with a large accuracy loss that is either similar to the initial class or likely to co-occur with it (e.g. agaric/mushroom, academic gown/graduation cap, military uniform/Pickelhaube, desk/monitor).

K.2. visual similarity analysis

In Figure 8, we provide initial analysis on the categories where the CuPL algorithm improves the most over standard hand-written prompts. We find that the improvement is not uniformly distributed, but rather some classes see a large improvement, while others see a decrease in per class accuracy. Interestingly, there are often two similar categories where one sees a large increase in accuracy and the other sees a decrease. For example, the ‘mushroom’ class has an approximately 40pp increase, while ‘agaric’ (a subclass of mushroom) is one of the classes with the largest drop in accuracy.

In order to better understand this phenomenon, we examine the change in accuracy between CuPL and the standard method of prompting in the image embedding space. Thus we are able to visualize the close relationship between categories like ‘agaric’ and ‘mushroom’. In order to be able to visualize the high dimensional CLIP image embedding in two dimensions, we utilize the t-distributed stochastic neighbor embedding algorithm [54]. Figure 9 visualizes image features (reduced into two dimensions) in relation to CuPL improvement.

As was suggested by Figure 8, we see in Figure 9 that when there is a class that has a large increase in accuracy with CuPL prompts (‘mushroom’, ‘graduation cap’, ‘monitor’) there is often a decrease in class accuracy for a visually related class. This means that when choosing between two similar or co-occurring classes, CuPL has a different distribution of classification than the standard method (e.g. the standard method prefers ‘canoe’ over ‘paddle’ much more strongly than CuPL). This suggests that the overall accuracy improvement of CuPL over the standard method may come (at least in part) from better distinguishing between two visually similar classes. While it may be over-correcting from the mistakes of the standard method (as demonstrated by the drop in accuracy in one of the two similar classes), the CuPL predictions appear to be overall more accurate, as demonstrated by the overall higher accuracy.

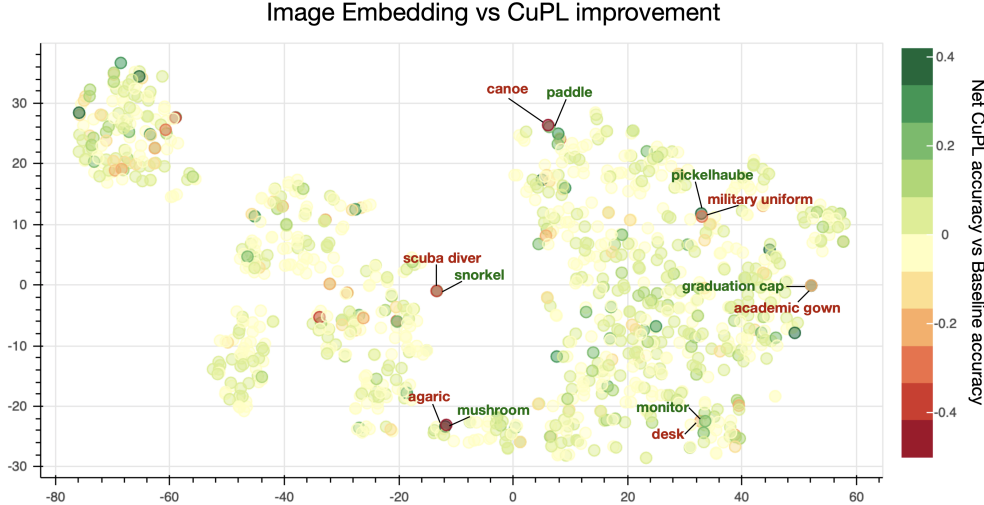


Figure 9. **Visualizing of image embedding of ImageNet classes compared to CuPL improvement on that class.** Each point on this figure represents the average image embedding of an ImageNet class, which has been reduced to two dimensions using t-sne [54]. We see that when there is a class with a large improvement compared to the baseline, it is often visually similar to a class which has a decrease in accuracy. This suggests that CuPL’s improved accuracy may be due in part to an increased ability to distinguish similar classes compared to the baseline.

K.3. Co-occurrences between objects

Many of the frequently confused pairs in Figure 9 are objects that are likely to occur in an image (i.e. ‘canoe’-‘paddle’ or ‘graduation cap’-‘academic gown’ or ‘monitor’-‘desk’). One potential benefit of CuPL captions is that they are able to capture co-occurrences as well. For example, one CuPL prompt for the ‘canoe’ class is *A canoe is typically a narrow boat with pointed ends that is propelled with a paddle.* This caption contains the word ‘paddle’ which is frequently confused with ‘canoe’ and likely to be present in images, even where the correct label is ‘canoe’. We therefore investigate the effectiveness of CuPL captions on images which contain more than one ImageNet object.

We attain this by using the ImageNet-ReaL dataset [3] which relabels ImageNet images with all applicable labels, so an image with both a ‘canoe’ and a ‘paddle’ would have both labels. We then tag images as having multiple ImageNet objects or only one ImageNet object based on the ReaL dataset. Finally, we compute ImageNet accuracy across each of these two sets (using standard ImageNet labels). Results are given in Table K.3.

	One class present (85.1% of ims)	Multiple classes present (14.9% of ims)
Standard	79.89	51.59
CuPL	80.79	53.58

Table 8. **Standard versus CuPL accuracy based on number of ImageNet classes present in image.** We use the ImageNet-ReaL dataset [3] to find images which have more than one applicable ImageNet label. We then present the accuracy for standard prompts and CuPL prompts using standard ImageNet labels, split by images which contain only one possible ImageNet class and images which may contain multiple classes.

L. Error Analysis

In Figure 10, we present an error analysis of our model using two different metrics. The first is an analysis between the model prediction and the correct class using the visual similarity of these labels. To capture this, we first attain an average visual embedding of each class by taking the mean of each image in that class and then normalizing that mean. Then for each class we rank how similar each of the other 999 classes are by the distance between these embeddings. If the models makes an incorrect prediction, but it predicts the class with the closest embedding to the correct label, then we refer to this as an *image offset* of 1.

Additionally, we examine the prediction errors in terms of the linguistic similarity of the labels. We do this with the WordNet [34] similarity of two labels. For example, if the label of the prediction and the ground-truth label share the same

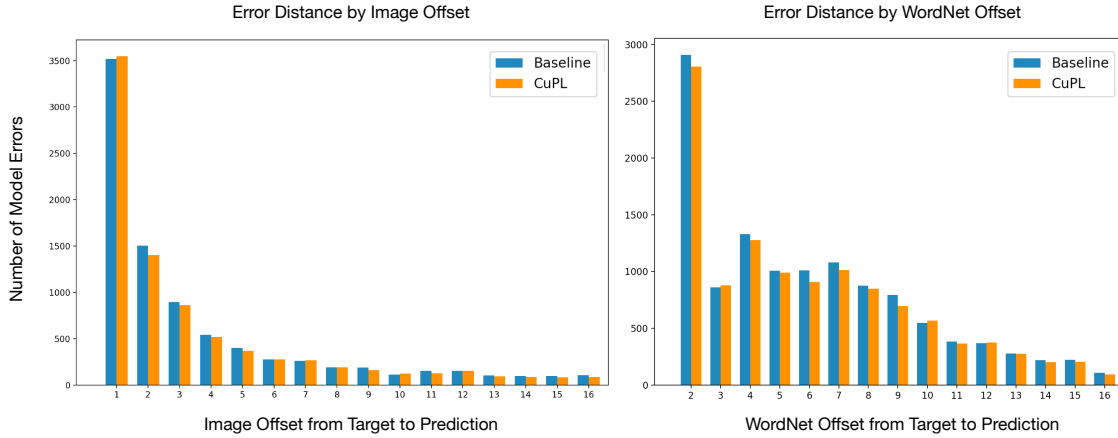


Figure 10. **Error analysis of CuPL and baseline comparing models errors visually and linguistically to ground truth labels.** We compare the errors made by CuPL to the errors made by the baseline methods. We find that the errors made by CuPL are more likely to be the most visually similar class to the ground truth label when compared to the errors made by the baseline. However, the errors made by the baseline are more likely to be the most linguistically similar class according to the WordNet [34] heirarchy. This suggests that visual information is being extracted from the CuPL descriptions to make visually consistent predictions.

parent in the WordNet tree, that is a *WordNet offset* of 2.

While slight, there is a difference in the errors made by CuPL compared to the errors made by the baseline as shown in Figure 10. CuPL is more likely to have an error that has an image offset of 1 than the baseline. However, the baseline is more likely to have an error that has a WordNet offset of 1 than CuPL. This implies that CuPL may be taking advantage of the visually descriptive language of the captions, as even when the model makes errors, they tend to favor categories that are visually similar to the ground truth. However, the baseline method does not have visual descriptions in its Image-prompts which may lead to its errors aligning more linguistically with the ground-truth.

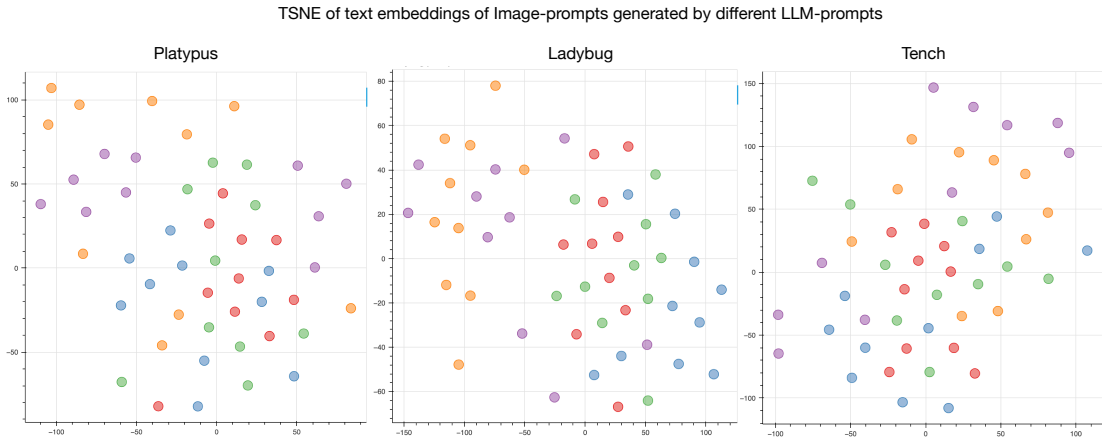


Figure 11. **Visualization of embeddings of Image-prompts generated with various LLM-prompts.** Each point represents the image embedding of one Image-prompts for the stated category with has been reduced to two dimensions using t-sne [54]. Image-prompts with the same color are generated by the same LLM-prompts.

M. Image-prompt Distribution

When generating Image-prompts, we use an ensemble of prompts generated by different LLM-prompts (e.g. ‘What does a {} look like?’). As we find that the diversity of Image-prompts is correlated with accuracy, it is valuable to understand how different LLM-prompts affect the diversity of Image-prompts. To accomplish this, we visualize the text embedding of Image-prompts for a selection of classes using t-sne [54] dimension reduction. We then color Image-prompts that were

generated by the same LLM-prompt. As shown in Figure 11, we find that while there is a slight clustering of Image-prompts by LLM-prompts, there is also a large amount of overlap.

N. Temperature Analysis

Diversity of Prompts. We also examine the impact of the diversity of image-prompts on ImageNet accuracy. We adjust this parameter by changing the temperature of the GPT-3 model. This value changes the likelihood of selecting lower probability tokens and makes sentences more diverse from each other. As demonstrated in Figure 12, more diverse prompts lead to higher ImageNet accuracy. Note these comparisons are done with a single LLM-prompt to save computational cost.

We provide several further analyses of the effect of temperature in prompt generation. First, we visualize the distribution of Image-prompts that have been generated with a variety of different temperatures. We do this by selecting prompts for 3 different ImageNet classes at 3 different temperatures. We then perform dimensionality reduction on the text embeddings of these prompts in order to visualize their distribution. As shown in Figure 13, Image-prompts generated with a temperature of 0.1 are clustered in a few different locations. Image-prompts generated with a temperature of 0.5 are more widely, and Image-prompts generated with a temperature of 0.9 have a similar, but even wider distribution.

Additionally in Section O, we give all generated prompts for the ImageNet class ‘Tench’ at three different temperatures. At the lowest temperature, the generated Image-prompts are nearly identical when generated with the same LLM-prompts. As the temperature increases, so does the difference in the generated prompts.

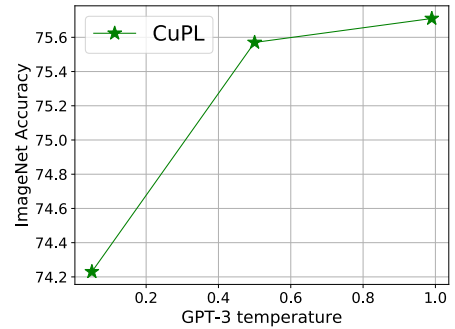


Figure 12. **Effect of LLM temperature.** More prompt diversity leads to higher performance.

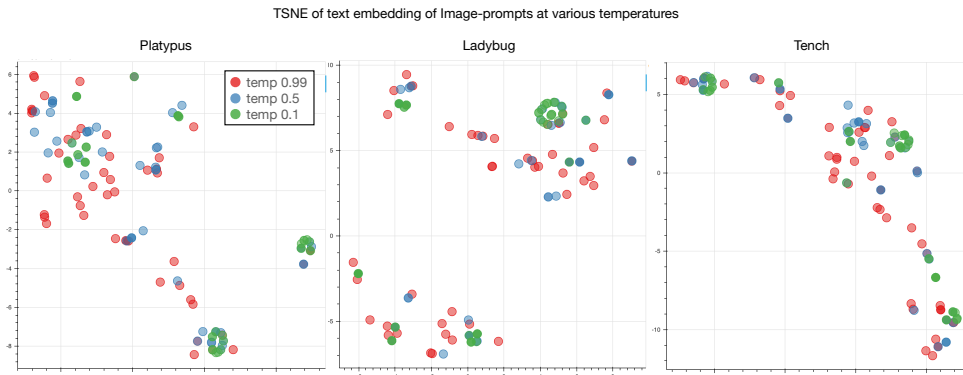


Figure 13. **Visualization of embeddings of Image-prompts generated with various temperatures.** Each point represents the image embedding of one Image-prompts for the stated category with has been reduced to two dimensions using t-sne [54]. Prompts generated with a higher temperature cover a wider distribution.

O. Example Generated image-prompts

A selection of LLM-generated image-prompts for a subset of ImageNet categories. We give all 50 image-prompts for the first ImageNet category of “Tench” and then 10 randomly selected prompts for a number of randomly selected ImageNet categories.

O.1. All generated image-prompts for “Tench” category

Temperature = 0.99

- "A tench is a freshwater fish of the carp family.",
- "A tench is a freshwater fish that is typically brown or olive in color.",
- "A tench is a fresh water fish that can grow up to 2 feet in length.",
- "A tench is a freshwater fish of the family Cyprinidae.",

"A tench is a freshwater fish of the carp family.",
"A tench is a freshwater fish with a dark green back and light-colored sides.",
"Tench are a freshwater fish found in Europe.",
"A tench is a small freshwater fish in the carp family.",
"A tench is a heavyset freshwater fish with a mottled brown body and a small, flat head.",
"A tench is a freshwater fish that looks similar to a carp.",
"A tench is a freshwater fish in the carp family.",
"A tench is a freshwater fish of the Cyprinidae family.",
"The tench is a freshwater fish of the Cyprinidae family.",
"The tench is a freshwater fish in the family Cyprinidae.",
"The easiest way to identify a tench is by its herringbone-patterned scales.",
"A tench is a freshwater fish of the carp family.",
"Tench are a freshwater fish found in Europe.",
"Tench have a large, slimy body with scales that have a green hue.",
"The tench is a freshwater fish belonging to the carp family.",
"A tench is a freshwater fish of the Cynoglossidae family.",
"A tench is a freshwater fish in the carp family.",
"Tench are freshwater fish with Olive Green backs, shading to Yellowish on the sides.",
"A tench looks like a green freshwater fish with a brownish hue.",
"A tench looks like a freshwater fish with a dark olive-green back, fading to yellowish-brown on the sides.",
"A tench usually has olive-green skin with dark spots, and a orange-yellow underbelly.",
"Tench are a freshwater fish that can grow up to 70cm long! They have olive-brown skin with dark spots, and their meat is white and firm.",
"A tench is a freshwater fish with a sturdy body and a greenish-brown coloration.",
"A tench is a freshwater fish that can grow up to about two feet long.",
"A tench is a freshwater fish in the carp family.",
"A tench is a large, freshwater fish with a thick body and large head.",
"The image is of a tench fish swimming in water.",
"The image is of a tench fish swimming in a pond.",
"The tench is a freshwater fish native to Europe.",
"This image shows a large, dark green tench swimming in a pond.",
"An image of a tench from the internet would likely show a dark green fish with a lighter underside.",
"The image is of a tench fish.",
"The image is of a tench fish on a white background.",
"A tench is a freshwater fish of the Cyprinidae family.",
"The image is of a tench swimming in a murky pond.",
"In the image, a tench swims in a pond with lily pads.",
" A tench in a river.",
"A tench (Tinca tinca) is a freshwater fish in the carp family that is found throughout Europe.",
" Tench (Tinca tinca), a member of the carp family (Cyprinidae), native to Eurasia.",
" A tench, a freshwater fish in the family Cyprinidae.",
" The tench (Tinca tinca) is a freshwater fish of the cyprinid family found throughout Eurasia.",
" A tench in a Finnish lake.",
"A tench (Tinca tinca) is a freshwater fish belonging to the carp family (Cyprinidae).",
"A tench in a fishpond.",
" The common tench is a freshwater fish of the cyprinid family found throughout Eurasia.",
"Tench (Tinca tinca) in a pond."

Temperature = 0.5

"A tench is a freshwater fish that is typically greenish-brown in color with a brassy sheen.",
"A tench is a freshwater fish that typically has a dark green back, light brown sides, and a white belly.",
"A tench is a freshwater fish that can grow up to two feet long.",

"Tench are a freshwater fish found in Europe.",
"A tench is a freshwater fish that is typically olive green in color with dark spots.",
"A tench is a freshwater fish that is typically olive green in color with a brownish tint.",
"A tench is a freshwater fish of the Cyprinidae family.",
"Tench are a freshwater fish found in Europe.",
"A tench is a freshwater fish that has a dark green back, light brown sides, and a white belly.",
"A tench is a freshwater fish that is typically olive-green in color with a brownish dorsal fin.",
"A tench is a freshwater fish of the cyprinid family.",
"A tench is a freshwater fish of the carp family.",
"A tench is a freshwater fish that is typically olive green in color with a brownish back.",
"The tench is a freshwater fish of the carp family Cyprinidae.",
"A tench is a freshwater fish that is typically greenish-brown in color.",
"A tench is a freshwater fish of the carp family.",
"A tench is a freshwater fish of the carp family.",
"Tench have olive green backs and flanks, with yellowish bellies.",
"A tench is a freshwater fish of the carp family.",
"A tench is a freshwater fish of the cyprinid family.",
"A tench is a freshwater fish that can grow up to 30 inches long.",
"A tench is a freshwater fish that is typically greenish-brown in color.",
"A tench is a freshwater fish that can grow up to about two feet long.",
"A tench is a freshwater fish with a brownish-green back and sides, and a yellowish-brown belly.",
"A tench is a freshwater fish that can grow up to two feet long.",
"A tench is a freshwater fish that looks similar to a carp.",
"A tench is a freshwater fish that is typically greenish-brown in color.",
"A tench is a freshwater fish that is part of the carp family.",
"A tench is a freshwater fish of the cyprinid family.",
"A tench is a freshwater fish that can grow up to two feet long.",
"The image is of a tench fish swimming in water.",
"The image is of a tench fish swimming in a pond.",
"The image is of a tench fish swimming in a pond.",
"The image is of a tench fish swimming in a pond.",
"The image is of a tench fish swimming in a pond.",
"In the image, a tench is swimming in a pond with lily pads.",
"The image is of a tench fish swimming in a pond.",
"The image is of a tench fish swimming in a pond.",
"The image is of a tench fish swimming in a pond.",
"The image is of a tench fish swimming in a pond.",
"A tench fish, native to Europe, characterized by its greenish-brown color and spots.",
" A tench (Tinca tinca) in a pond.",
"A tench (Tinca tinca) is a freshwater fish belonging to the carp family (Cyprinidae).",
"A tench (Tinca tinca) is a freshwater fish in the carp family.",
"A tench (Tinca tinca) is a freshwater fish in the carp family (Cyprinidae).",
" A tench in a river.",
"A tench (Tinca tinca) is a freshwater fish in the carp family (Cyprinidae).",
"A tench (Tinca tinca) in a pond.",
" A tench (Tinca tinca) in a garden pond.",
" A tench in a river."

Temperature = 0.1

"A tench is a freshwater fish that is typically olive green in color with dark spots.",
"A tench is a freshwater fish that can grow up to 30 inches long.",
"A tench is a freshwater fish that can grow to a length of over two feet.",
"A tench is a freshwater fish that can grow up to two feet long.",
"A tench is a freshwater fish that can grow up to two feet long.",
"A tench is a freshwater fish that is typically olive green in color with dark spots.",

"A tench is a freshwater fish that can grow up to two feet long.",
 "A tench is a freshwater fish that can grow up to two feet long.",
 "A tench is a freshwater fish that can grow up to two feet long.",
 "A tench is a freshwater fish that is typically olive green in color with dark spots.",
 "A tench is a freshwater fish of the carp family.",
 "A tench is a freshwater fish of the carp family.",
 "A tench is a freshwater fish of the carp family.",
 "A tench is a freshwater fish of the carp family.",
 "A tench is a freshwater fish of the carp family.",
 "Tench have a dark green back, light olive sides, and a yellowish belly.",
 "A tench is a freshwater fish of the carp family.",
 "A tench is a freshwater fish of the carp family.",
 "A tench is a freshwater fish of the carp family.",
 "A tench is a freshwater fish that can grow up to two feet long.",
 "A tench is a freshwater fish that can grow up to two feet long.",
 "A tench is a freshwater fish that can grow up to two feet long.",
 "A tench is a freshwater fish that can grow up to two feet long.",
 "A tench is a freshwater fish that is typically olive green in color with a brownish dorsal side.",
 "A tench is a freshwater fish that can grow up to two feet long.",
 "A tench is a freshwater fish that can grow up to two feet long.",
 "A tench is a freshwater fish that is typically olive green in color with a brownish dorsal side.",
 "A tench is a freshwater fish that can grow up to two feet long.",
 "A tench is a freshwater fish that can grow up to two feet long.",
 "The image is of a tench fish swimming in a pond.",
 "The image is of a tench fish swimming in a pond.",
 "The image is of a tench fish swimming in a pond.",
 "The image is of a tench fish swimming in a pond.",
 "The image is of a tench fish swimming in a pond.",
 "The image is of a tench fish swimming in a pond.",
 "The image is of a tench fish swimming in a pond.",
 "The image is of a tench fish swimming in a pond.",
 "The image is of a tench fish swimming in a pond.",
 "A tench (Tinca tinca) is a freshwater fish in the carp family.",
 "A tench (Tinca tinca) is a freshwater fish of the carp family (Cyprinidae).",
 "A tench (Tinca tinca) is a freshwater fish in the carp family.",
 "A tench (Tinca tinca) is a freshwater fish in the carp family.",
 "A tench (Tinca tinca) is a freshwater fish in the carp family.",
 "A tench (Tinca tinca) is a freshwater fish in the carp family.",
 "A tench (Tinca tinca) is a freshwater fish in the carp family.",
 "A tench (Tinca tinca) is a freshwater fish in the carp family.",
 "A tench (Tinca tinca) in a garden pond.",
 "A tench (Tinca tinca) in a garden pond.",
 "A tench (Tinca tinca) is a freshwater fish of the carp family (Cyprinidae)."

O.2. Sample generated image-prompts for randomly selected ImageNet categories

"bubble":
 "A bubble is a sustained period of inflated asset prices.",
 "A bubble looks like a sphere of air.",
 "A bubble is often characterized by rapidly increasing prices in an asset or security, followed by a sharp decrease in prices.",
 "A bubble looks like a small, round, thin film of soap filled with air.",
 "A bubble looks like a round sphere of soap film.",
 "\"A bubble being blown in the park.",
 "A close-up of a soap bubble with a thin film of water in between two layers of air.",

"A bubble is a spherical shape made up of a thin film of soap water.",
"A bubble is a circle of air surrounded by water.",
"A bubble looks like a round, thin layer of soap surrounding a pocket of air."

"kit fox":

"A kit fox is a small species of fox, about the size of a domestic cat.",
"You can identify a kit fox by its small size, its big ears, and its long, bushy tail.",
" A kit fox laying on the ground in a desert habitat.",
"A kit fox is a small fox found in North America.",
"A kit fox is a small fox with a sleek coat of fur.",
"A kit fox is a small fox with large ears, a long, black-tipped tail, and pale fur.",
"A kit fox has a reddish coat, with white patches on its chest and throat.",
"This kit fox has a reddish coat and large ears.",
"A kit fox looks like a small fox with a pointed nose, large ears, and a long, bushy tail.",
"A kit fox is a small species of fox."

"toy terrier":

"A toy terrier looks like a very small version of a terrier.",
"The image is of a toy terrier that is mostly white with brown spots.",
"You can identify a toy terrier by looking for a compact, short-legged dog with a short muzzle.",
" Cute little guy.",
"Toy terriers are miniature versions of terriers, such as the Jack Russell Terrier.",
"A toy terrier is a small, lightweight breed of dog.",
"The image is of a small, brown toy terrier.",
"A toy terrier typically has a long, narrow head with pointy ears, and a small, compact body.",
"A toy terrier is a small, short-legged dog with a long body, pointy nose, and large ears.",
"A small, brown and white toy terrier is sitting on a beige couch, looking at the camera."

"mousetrap":

"In the image, there is a mousetrap made of wood and metal.",
"An image of a mousetrap from the internet would most likely show a traditional wooden mousetrap with a metal spring.",
"A mousetrap is a device made to catch and kill mice.",
"A mousetrap is a small device that is used to catch mice.",
"The classic mousetrap consists of a wooden base with a metal spring mounted on one end.",
"The mousetrap is a small wooden box with a metal spring inside.",
"The mousetrap is a simple device that has been used for centuries to catch mice.",
"The classic mousetrap - simple, effective, and deadly.",
"The most common way to identify a mousetrap is by its small size and rectangular shape.",
"A mousetrap typically has a wire or wooden frame that is baited with food and springs open quickly to snap shut on the mouse when it attempts to steal the bait."

"dog sled":

"A dog sled is a vehicle on runners, typically with a thin frame and a flat bottom, that is used to convey goods or passengers over snow or ice.",
"A dog sled is a toboggan pulled by dogs, typically over snow.",
"A dog sled looks like a bed on runners that is pulled by dogs.",
"A dog sled is typically a heavy frame on runners that is pulled by one or more dogs.",
"A dog sled is traditionally a sled pulled by dogs, used for transportation, racing, or other purposes.",
" teamA dog sled team is a group of dogs that are harnessed together to pull a sled.",
"A dog sled looks like a small, open vehicle that is pulled by one or more dogs.",
"Dog sledding in winter.",
" pulled by huskiesThe image is of a dog sled pulled by huskies.",
"A dog sled looks like a large cart that is pulled by a team of dogs."

"geyser":

"A geyser is a hot spring that periodically erupts, shooting a column of water and steam into the air.",
"A geyser looks like a hole in the ground that sometimes spurts hot water and steam into the air.",
"Geysers are hot springs that periodically spout water and steam into the air.",
"The image is of a geyser erupting.",
"A geyser is a hot spring that periodically erupts, spraying water into the air.",
"A geyser typically looks like a cone of rocks with a small hole at the top.",
"A geyser is a hot spring where water intermittently boils, sending a jet of hot water and steam into the air.",
"A geyser is a hot spring that periodically shoots a stream of hot water and steam into the air.",
"The image is of a geyser shooting water high into the air.",
"A geyser looks like a column of water that shoots into the air and then falls back down."

"Schipperke":

"A Schipperke is a small, Belgian breed of dog.",
"A Schipperke is a small black Belgian dog with a rat-like tail.",
"The image is of a black and white dog with pointy ears and a long body.",
"A Schipperke is a small, black, Belgian breed of dog.",
"A Schipperke is a small, black, Belgian breed of dog that closely resembles a fox.",
"A Schipperke is a small Belgian breed of dog that resembles a fox.",
"Schipperkes have a long, black coat and a pointed muzzle.",
"A Schipperke is a small dog breed with a fox-like appearance.",
"It's a photo of a black and tan Schipperke dog standing in front of a brick wall.",
"Black, small, spitz-type dog with a long, fox-like snout, large erect ears, and a long, high-set tail."

"go-kart":

"A go-kart typically looks like a small car or buggy with a small engine in the back.",
"A go-kart is a small vehicle with four wheels, a steering wheel, and a gas pedal.",
"A go-kart is a small vehicle with a steering wheel, pedals, and an engine.",
"Two young girls in go-karts race down a path in a park.",
"A go-kart is a small, lightweight vehicle with four wheels and a simple, open frame.",
"A go-kart is a small, open-wheeled vehicle used for racing.",
"Two kids racing go-karts on a dirt track.",
"A go-kart is a small, racing car.",
"A go-kart typically looks like a small, open-wheeled car.",
"A go-kart is a small, lightweight vehicle with four wheels that is propelled by a small engine."

"black-and-white colobus":

"The Black-and-White Colobus is a type of Old World monkey, found in Africa.",
"A black-and-white colobus has long black fur, and a white face with a black triangle around the eyes.",
"Colobus Monkey in the TreesThis elegant colobus monkey is swinging through the trees in search of food.",
" monkeyIn this image, a black-and-white colobus monkey is shown perched atop a tree branch.",
"The black-and-white colobus monkey is one of the most beautiful and distinctive of all the colobus monkeys.",
"The black-and-white colobus monkey is a species of primate in the Colobidae family.",
"Colobus monkeys are generally black with white patches on their face, back, and sides.",
"The black-and-white colobus is a species of Old World monkey.",
" monkeyThe image is of a black-and-white colobus monkey sitting on a tree branch.",
" monkeyIn the image, the black-and-white colobus monkey is sitting in a tree."

"sock":

"A sock usually has a cuff at the top, and a heel at the bottom.",

" A black sock with a white line running down the middle.",

"A sock is typically a garment worn on the feet and made from a soft material, such as cotton.",

"A sock normally has a heel, toe and a cuff at the top.",

"A sock is a small amount of money that is given to someone without them knowing.",

"A sock is an article of clothing worn on the feet.",

"A sock is a piece of clothing that is worn on the feet.",

"A sock is a tubular garment that covers the foot and ankle.",

"This image is of a blue and white striped sock.",

"There are many ways that you can identify a sock."

"Cocker Spaniel":

"A Cocker Spaniel is a medium sized dog with long, floppy ears, and a silky coat that is usually either brown or black.",

"The Cocker Spaniel has a long floppy ears, a silky coat, and a bushy tail.",

"A Cocker Spaniel has a long, black muzzle and big, brown eyes.",

"The Cocker Spaniel is a breed of dog.",

"An image of a Cocker Spaniel from the internet shows a small brown and white dog with long floppy ears.",

"The image is of a Cocker Spaniel with short, brown fur and long, floppy ears.",

"Cocker spaniels have long, floppy ears and a long, silky coat.",

"A Cocker Spaniel is a small to medium sized dog.",

"The image is of a light brown and white Cocker Spaniel standing on a green grassy field with its head turned to the side.",

"A Cocker Spaniel has a long, silky coat that is usually either black, brown, or golden."

"southern black widow":

"A southern black widow spider perched atop a web.",

"Female southern black widows have a black body with a red hourglass shape on their abdomen.",

"The southern black widow is black with a red hourglass shape on its belly.",

"Female southern black widow spiders are black with a characteristic red hourglass-shaped mark on their ventral abdomen.",

" spiderThe image is of a large, black spider with a red hourglass shape on its abdomen.",

"A southern black widow is a spider that is black with a red hourglass shape on its abdomen.",

"A southern black widow spider is a small, black spider with a red hourglass-shaped mark on its underside.",

"There's an image on the internet of a southern black widow that's really cool.",

"A southern black widow can be identified by its black coloration with a red hourglass shape on its abdomen.",

"A Southern black widow is a type of spider that is black with a red hourglass shape on its belly."

"catamaran":

"A catamaran is a sailboat that has two hulls, or wide bodies, that are connected by beams.",

"The easiest way to identify a catamaran is by its two hulls.",

"A catamaran is a type of boat that has two parallel hulls.",

"A catamaran is a multi-hulled vessel with two parallel hulls of equal size.",

"My dream boat! A sleek catamaran that can zip through the waves.",

"A catamaran is a multi-hulled vessel with two parallel hulls of equal size.",

"The image is of a white catamaran with blue trim.",

"A catamaran is a type of boat that has two hulls, or platforms, that are parallel to each other.",

"The image is of a yellow catamaran with white trim, sitting in calm water.",

"A catamaran is a type of sailing vessel that consists of two parallel hulls of equal size."

"beach":

".",

"Blue skies, white sands, and clear turquoise waters make this beach a paradise.",

"A beach usually has sand and water.",

"Beach identification can be accomplished through the identification of physical characteristics.",

"A beach is a naturally occurring feature of the landscape.",

"The sun sets over the ocean, casting a beautiful orange hue in the sky.",

"A beach is a large body of water with sand or small rocks at the shore.",

"In the image, the beach is Brilliant white with crystal blue waters.",

"The beach looks like a long strip of land next to the ocean.",

"A beach typically looks like a large, flat expanse of sand with some rocks or other natural features nearby."

"rotary dial telephone":

"A rotary dial is a device used to dial telephone numbers.",

"Rotary dial telephone from the mid-20th century.",

"A rotary dial telephone is a phone with a circular dial on the front face.",

"When looking at a rotary telephone, you can tell it is a rotary phone by the Place the phone's receiver on your ear and listen for a dial tone.",

"A rotary dial phone is an older model phone that has a circular device with numbers on it that you rotate with your finger to dial a number.",

"A rotary dial telephone is a type of telephone that uses a mechanical dial to select the telephone number that a user wishes to call.",

"History of the Rotary Dial Telephone.",

"A rotary dial telephone is an old-fashioned telephone that has a round dial on the front of it.",

" rotary dial telephone looks like a classic telephone with a rotary dial on the base unit.",

"The image is of a rotary dial telephone on a black background."