Poster Presentation Tips

Administrative

- Grades for Assignment 3, 4 out
- Grades for Quiz 4 out
- Grades for Milestone out
- A5 due 2 days ago

The rest of the course:

- Quiz 5 today!
- Poster Session on Thursday
- Grade reports posted to gradescope friday/saturday
    - Time to check over your late days/ any extensions etc

Administrative

- Grades for Assignment 3, 4 out
- Grades for Quiz 4 out
- Grades for Milestone out
- A5 due 2 days ago

The rest of the course:

- Quiz 5 today!
- Poster Session on Thursday
- Grade reports posted to gradescope friday/saturday
    - Time to check over your late days/ any extensions etc

Logistics

1. Arrive in lobby of allen building
2. Pick up poster from us (send by tomorrow at noon) or bring your printed poster
3. Set up poster board and poster stand
4. We will come around and assign you group 1 or group 2
    a. Group 1 presents poster in first half and looks at other poster second
    b. Group 2 looks at posters in first half and presents second
5. In either the first or second half of the session, you will stand by your poster and explain your project to us/ other students

How to make a good poster/presentation

Start with a good presentation, make the poster to match

How to make a good poster/presentation

Start with a good presentation, make the poster to match

How to make a good presentation?

30 - 60 second overview

How to make a good presentation?

30 - 60 second overview

1. **<u>High level intro</u>** (10 seconds)
    1. What is the motivation for your work? What is not done yet that you are solving? Make sure you set it up for why your project is necessary.
    2. Explain your approach at a high level

How to make a good presentation?

30 - 60 second overview

1. **<u>High level intro</u>** (10 seconds)
    1. What is the motivation for your work? What is not done yet that you are solving? Make sure you set it up for why your project is necessary.
    2. Explain your approach at a high level
2. **<u>Details of your method</u>** (15 - 25 seconds)
    1. What are you actually doing specifically?

How to make a good presentation?

30 - 60 second overview

1. **<u>High level intro</u>** (10 seconds)
    1. What is the motivation for your work? What is not done yet that you are solving? Make sure you set it up for why your project is necessary.
    2. Explain your approach at a high level
2. **<u>Details of your method</u>** (15 - 25 seconds)
    1. What are you actually doing specifically?
3. **<u>Results</u>** (15 - 25 seconds)
    1. What experiments did you run? What did you find?
    2. How does this show that you solved the problem that you set up?

How to make a good poster/presentation

Start with a good presentation, make the poster to match

How to make a good poster?

Your poster is a tool to help you with your presentation

You don’t want just all the text of everything you are planning on saying aloud

Go through your presentation script and see which of the points could use a visual component

- A diagram of your model to reference as you are explaining it?
- The graph of your results?

How to make a good poster?

Flow of visual elements in poster matches flow of presentation

1. **<u>High level intro</u>** (10 seconds)
   1. What is the motivation for your work? What is not done yet that you are solving? Make sure you set it up for why your project is necessary.
   2. Explain your approach at a high level
2. **<u>Details of your method</u>** (15 - 25 seconds)
   1. What are you actually doing specifically?
3. **<u>Results</u>** (15 - 25 seconds)
   1. What experiments did you run? What did you find?
   2. How does this show that you solved the problem that you set up?

WHAT DOES A PLATYPUS LOOK LIKE?

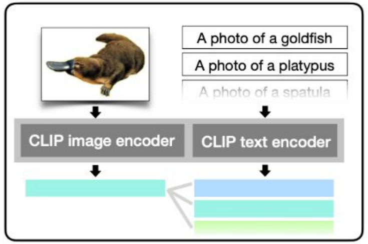
Generating **customized prompts** for **zero-shot image classification**

Sarah Pratt[1] Ian Covert[1] Rosanne Liu[2,3] Ali Farhadi[1]

1 University of Washington
2 Google DeepMind
3 ML Collective

"A platypus looks like a beaver with a ducks bill"

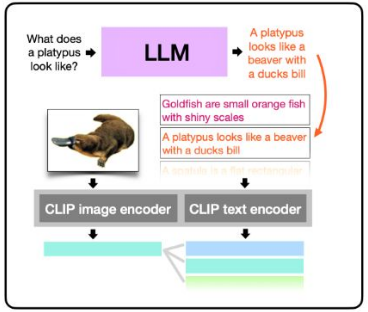
LLM



Standard CLIP Prompts :

Handwrite custom templates ("a photo of a ____") for each dataset and fill them in with each category.

- ❌ Templates contain no specific visual information
- ❌ Requires many hand-written templates (80 for ImageNet)
- ❌ Contains information about data distribution ("a black and white photo of _____")



CuPL Prompts (ours):

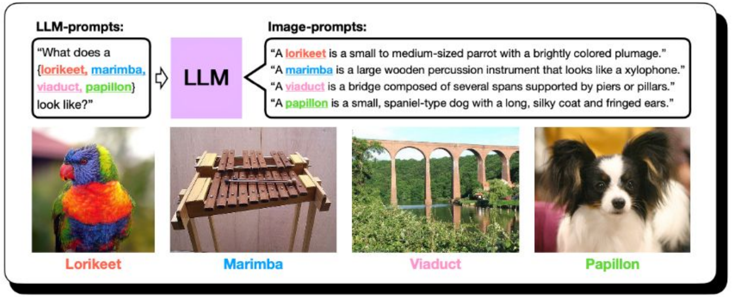
Customized Prompts via Language models

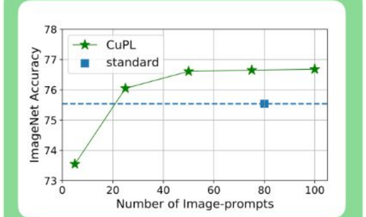
Use an LLM to generate many different descriptions of each category in the dataset.

- ★ **Higher accuracy** as prompts now can contain many important visual features (color, environment, subparts ...)
- ★ **Little to no prompt engineering** for new datasets
- ★ Doesn't use prior knowledge of data distribution

CuPL (Full) **Highest Accuracy:** Custom questions to query LLM for each dataset.

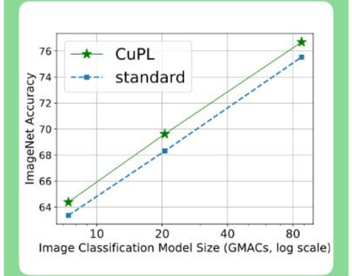
CuPL (Base) **No prompt engineering:** Uses the same 3 LLM queries for every dataset





More descriptions lead to higher accuracy and LLMs can generate many descriptions without additional human effort

Accuracy gains remain consistent as the CLIP model is scaled up



ImageNet

	Standard	CuPL (Full)	CuPL (Base)
Accuracy	75.54	**76.69**	76.19
# Hand written	80	5	**3**

Average across 15 benchmarks

	Standard	CuPL (Full)	CuPL (Base)
Avg Accuracy	73.43	**74.8**	74.15
Total # Hand written	175	45	**3**

Title + Authors

WHAT DOES A PLATYPUS LOOK LIKE?

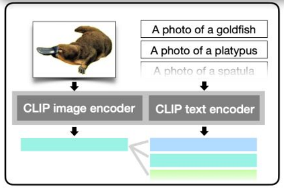
Generating **customized prompts** for **zero-shot image classification**

Sarah Pratt[1] Ian Covert[1] Rosanne Liu[2,3] Ali Farhadi[1]

1 UNIVERSITY of WASHINGTON
2 Google DeepMind
3 ML Collective

"A platypus looks like a beaver with a ducks bill"

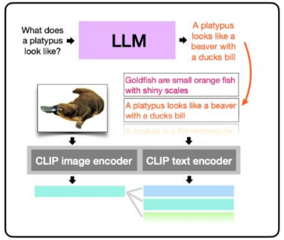
LLM



Standard CLIP Prompts :

Handwrite custom templates ("a photo of a ____") for each dataset and fill them in with each category.

- ✗ Templates contain no specific visual information
- ✗ Requires many hand-written templates (80 for ImageNet)
- ✗ Contains information about data distribution ("a black and white photo of _____")



CuPL Prompts (ours):

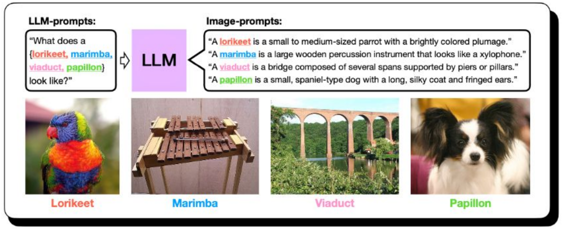
Customized Prompts via Language models

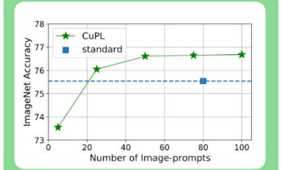
Use an LLM to generate many different descriptions of each category in the dataset.

- ★ **Higher accuracy** as prompts now can contain many important visual features (color, environment, subparts …)
- ★ **Little to no prompt engineering** for new datasets
- ★ Doesn't use prior knowledge of data distribution

CuPL (Full) **Highest Accuracy:** Custom questions to query LLM for each dataset.

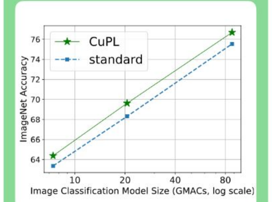
CuPL (Base) **No prompt engineering:** Uses the same 3 LLM queries for every dataset





More descriptions lead to higher accuracy and LLMs can generate many descriptions without additional human effort

Accuracy gains remain consistent as the CLIP model is scaled up



ImageNet

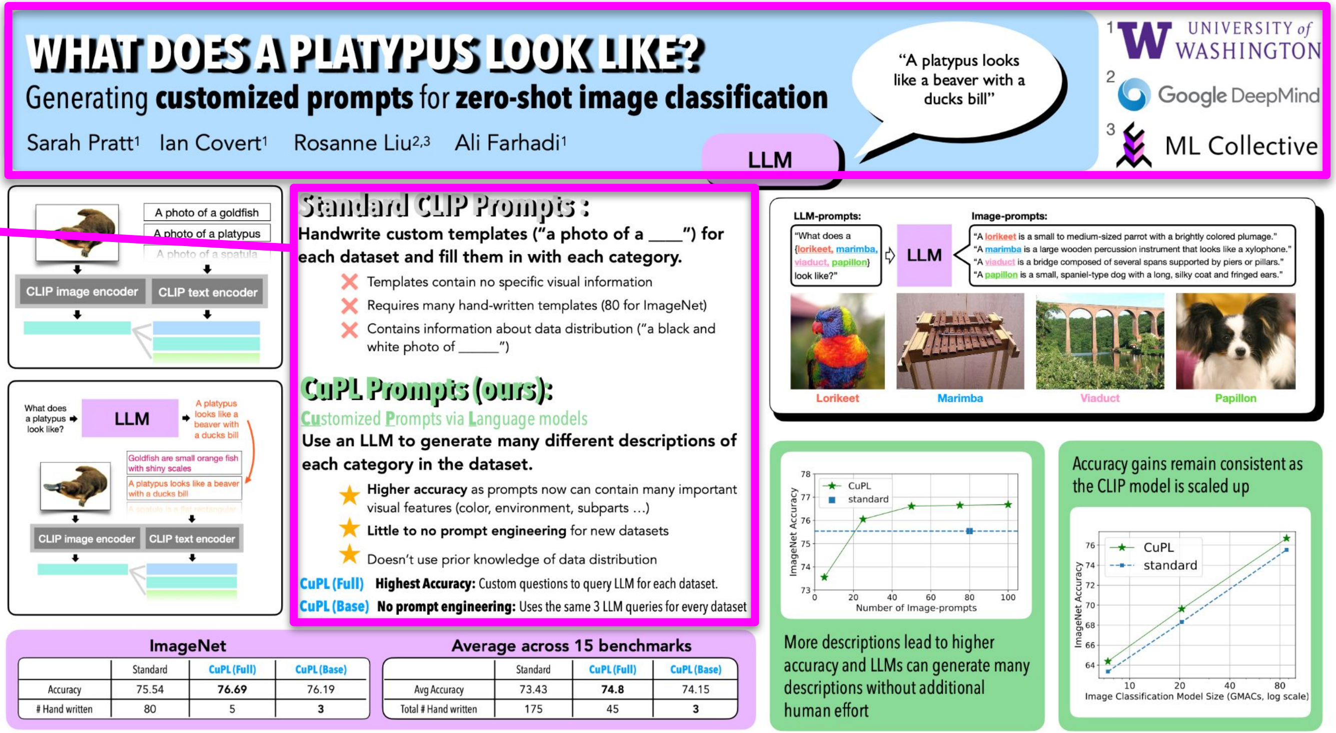
	Standard	CuPL (Full)	CuPL (Base)
Accuracy	75.54	**76.69**	76.19
# Hand written	80	5	**3**

Average across 15 benchmarks

	Standard	CuPL (Full)	CuPL (Base)
Avg Accuracy	73.43	**74.8**	74.15
Total # Hand written	175	45	**3**

Title + Authors

Motivation + method



WHAT DOES A PLATYPUS LOOK LIKE?

Generating **customized prompts** for **zero-shot image classification**

Sarah Pratt[1] Ian Covert[1] Rosanne Liu[2,3] Ali Farhadi[1]

1 UNIVERSITY of WASHINGTON
2 Google DeepMind
3 ML Collective

"A platypus looks like a beaver with a ducks bill"

LLM

A photo of a goldfish
A photo of a platypus
A photo of a spatula
CLIP image encoder
CLIP text encoder

What does a platypus look like?
LLM
A platypus looks like a beaver with a ducks bill
Goldfish are small orange fish with shiny scales
A platypus looks like a beaver with a ducks bill
CLIP image encoder
CLIP text encoder

Standard CLIP Prompts :

Handwrite custom templates ("a photo of a ____") for each dataset and fill them in with each category.

- ✗ Templates contain no specific visual information
- ✗ Requires many hand-written templates (80 for ImageNet)
- ✗ Contains information about data distribution ("a black and white photo of ______")

CuPL Prompts (ours):

Customized Prompts via Language models

Use an LLM to generate many different descriptions of each category in the dataset.

- ★ **Higher accuracy** as prompts now can contain many important visual features (color, environment, subparts …)
- ★ **Little to no prompt engineering** for new datasets
- ★ Doesn't use prior knowledge of data distribution

CuPL (Full) **Highest Accuracy:** Custom questions to query LLM for each dataset.

CuPL (Base) **No prompt engineering:** Uses the same 3 LLM queries for every dataset

LLM-prompts:

"What does a {lorikeet, marimba, viaduct, papillon} look like?"

LLM

Image-prompts:

"A lorikeet is a small to medium-sized parrot with a brightly colored plumage."
"A marimba is a large wooden percussion instrument that looks like a xylophone."
"A viaduct is a bridge composed of several spans supported by piers or pillars."
"A papillon is a small, spaniel-type dog with a long, silky coat and fringed ears."

Lorikeet Marimba Viaduct Papillon

ImageNet

	Standard	CuPL (Full)	CuPL (Base)
Accuracy	75.54	**76.69**	76.19
# Hand written	80	5	**3**

Average across 15 benchmarks

	Standard	CuPL (Full)	CuPL (Base)
Avg Accuracy	73.43	**74.8**	74.15
Total # Hand written	175	45	**3**

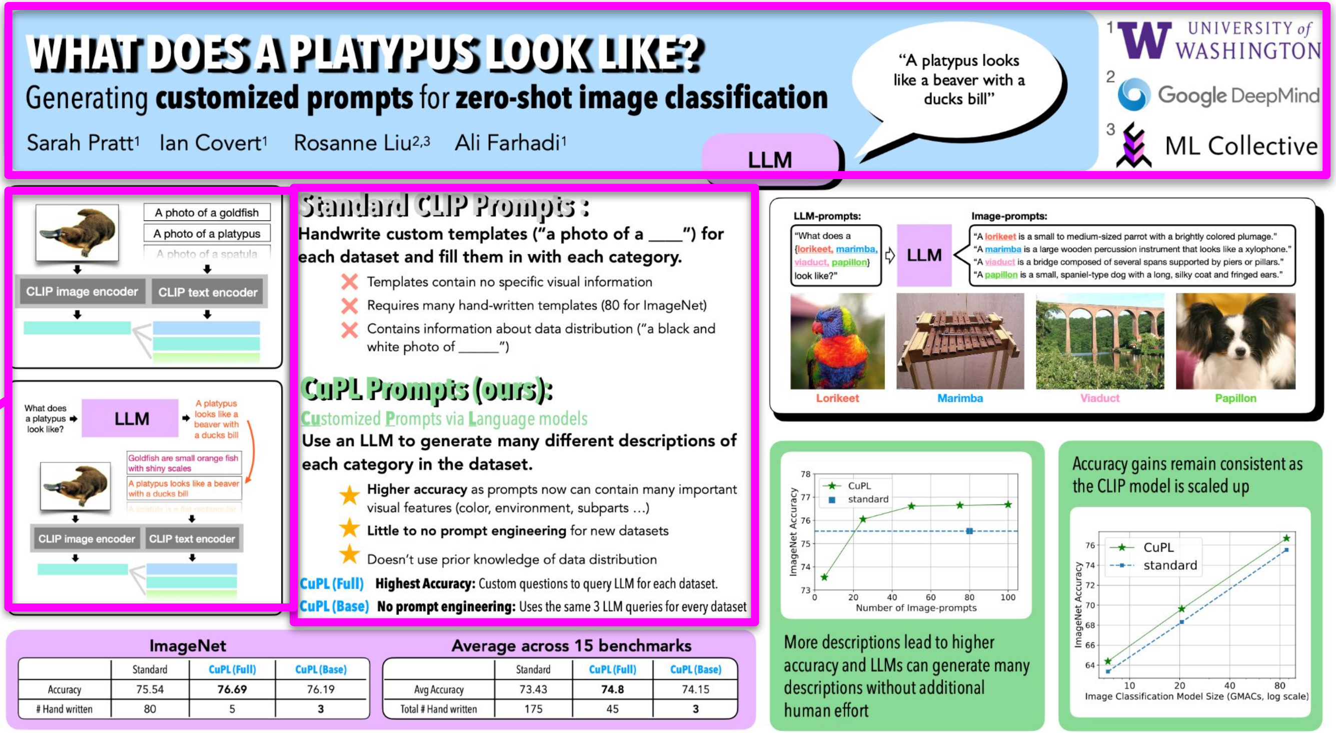
More descriptions lead to higher accuracy and LLMs can generate many descriptions without additional human effort

Accuracy gains remain consistent as the CLIP model is scaled up

Title + Authors

Motivation + method

Method Details



Title + Authors

Motivation + method

Method Details

Qualitative Results

WHAT DOES A PLATYPUS LOOK LIKE?

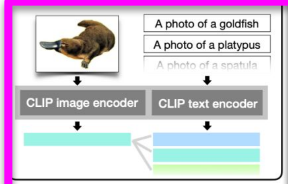
Generating **customized prompts** for **zero-shot image classification**

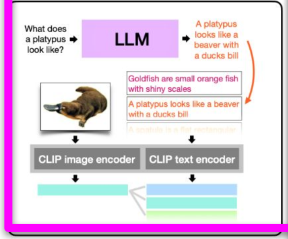
Sarah Pratt[1] Ian Covert[1] Rosanne Liu[2,3] Ali Farhadi[1]

1 UNIVERSITY of WASHINGTON
2 Google DeepMind
3 ML Collective

"A platypus looks like a beaver with a ducks bill"

LLM





Standard CLIP Prompts:

Handwrite custom templates ("a photo of a ____") for each dataset and fill them in with each category.

- ✗ Templates contain no specific visual information
- ✗ Requires many hand-written templates (80 for ImageNet)
- ✗ Contains information about data distribution ("a black and white photo of ____")

CuPL Prompts (ours):

Customized Prompts via Language models

Use an LLM to generate many different descriptions of each category in the dataset.

- ★ **Higher accuracy** as prompts now can contain many important visual features (color, environment, subparts …)
- ★ **Little to no prompt engineering** for new datasets
- ★ Doesn't use prior knowledge of data distribution

CuPL (Full) **Highest Accuracy:** Custom questions to query LLM for each dataset.

CuPL (Base) **No prompt engineering:** Uses the same 3 LLM queries for every dataset

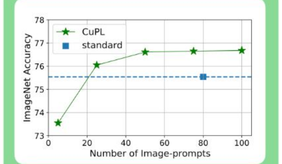
LLM-prompts:

"What does a {lorikeet, marimba, viaduct, papillon} look like?"

LLM

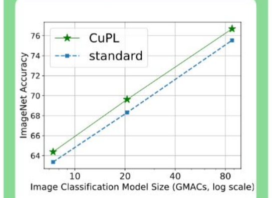
Image-prompts:

"A lorikeet is a small to medium-sized parrot with a brightly colored plumage."
"A marimba is a large wooden percussion instrument that looks like a xylophone."
"A viaduct is a bridge composed of several spans supported by piers or pillars."
"A papillon is a small, spaniel-type dog with a long, silky coat and fringed ears."



More descriptions lead to higher accuracy and LLMs can generate many descriptions without additional human effort

Accuracy gains remain consistent as the CLIP model is scaled up



ImageNet

	Standard	CuPL (Full)	CuPL (Base)
Accuracy	75.54	**76.69**	76.19
# Hand written	80	5	**3**

Average across 15 benchmarks

	Standard	CuPL (Full)	CuPL (Base)
Avg Accuracy	73.43	**74.8**	74.15
Total # Hand written	175	45	**3**

Title + Authors

WHAT DOES A PLATYPUS LOOK LIKE?

Generating **customized prompts** for **zero-shot image classification**

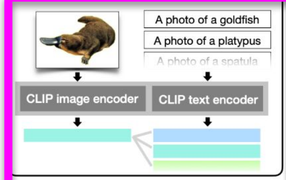
Sarah Pratt[1] Ian Covert[1] Rosanne Liu[2,3] Ali Farhadi[1]

1 UNIVERSITY of WASHINGTON
2 Google DeepMind
3 ML Collective

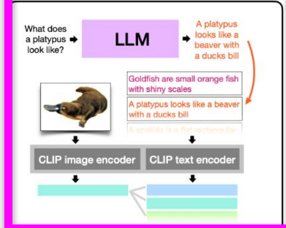
"A platypus looks like a beaver with a ducks bill"

LLM

Motivation + method



Method Details



Standard CLIP Prompts :

Handwrite custom templates ("a photo of a ____") for each dataset and fill them in with each category.

- ✗ Templates contain no specific visual information
- ✗ Requires many hand-written templates (80 for ImageNet)
- ✗ Contains information about data distribution ("a black and white photo of _____")

CuPL Prompts (ours):

Customized Prompts via Language models

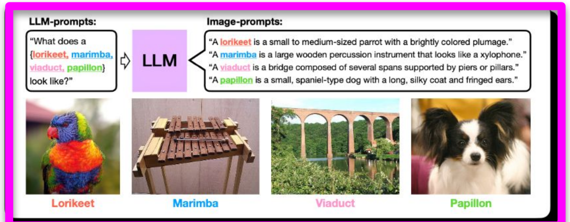
Use an LLM to generate many different descriptions of each category in the dataset.

- ★ **Higher accuracy** as prompts now can contain many important visual features (color, environment, subparts …)
- ★ **Little to no prompt engineering** for new datasets
- ★ Doesn't use prior knowledge of data distribution

CuPL (Full) **Highest Accuracy:** Custom questions to query LLM for each dataset.

CuPL (Base) **No prompt engineering:** Uses the same 3 LLM queries for every dataset

Qualitative Results

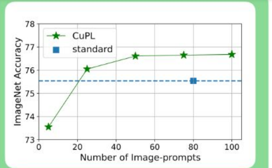


ImageNet

	Standard	CuPL (Full)	CuPL (Base)
Accuracy	75.54	**76.69**	76.19
# Hand written	80	5	**3**

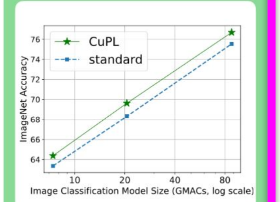
Average across 15 benchmarks

	Standard	CuPL (Full)	CuPL (Base)
Avg Accuracy	73.43	**74.8**	74.15
Total # Hand written	175	45	**3**



More descriptions lead to higher accuracy and LLMs can generate many descriptions without additional human effort

Accuracy gains remain consistent as the CLIP model is scaled up



Quantitative Results