Transfer Learning: Multi-Task Learning and Few/Zero Shot Learning

Danna Gurari
University of Colorado Boulder
Fall 2022

https://home.cs.colorado.edu/~DrG/Courses/NeuralNetworksAndDeepLearning/AboutCourse.html

Review

- Last lecture topic:
  - Transfer learning definition
  - Overview of self-supervised learning
  - Generative-based methods
  - Generative adversarial networks
  - Context-based methods
- Assignments (Canvas)
  - Final project proposal due in one week
- Questions?

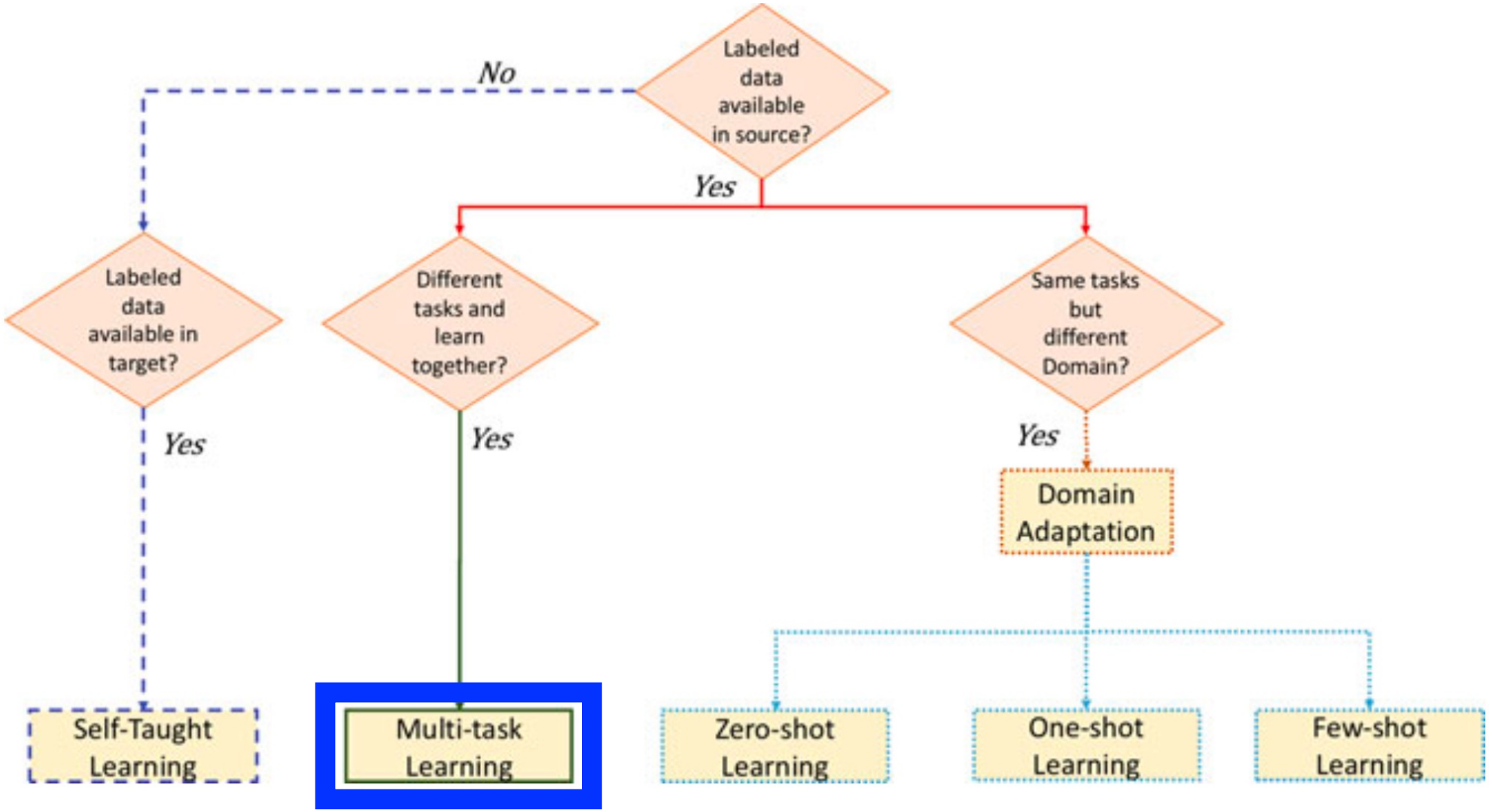
Today’s Topics

- Multi-task learning
- Few-shot learning
- Zero-shot learning
- Cloud GPU tutorial

Today’s Topics

- Multi-task learning
- Few-shot learning
- Zero-shot learning
- Cloud GPU tutorial

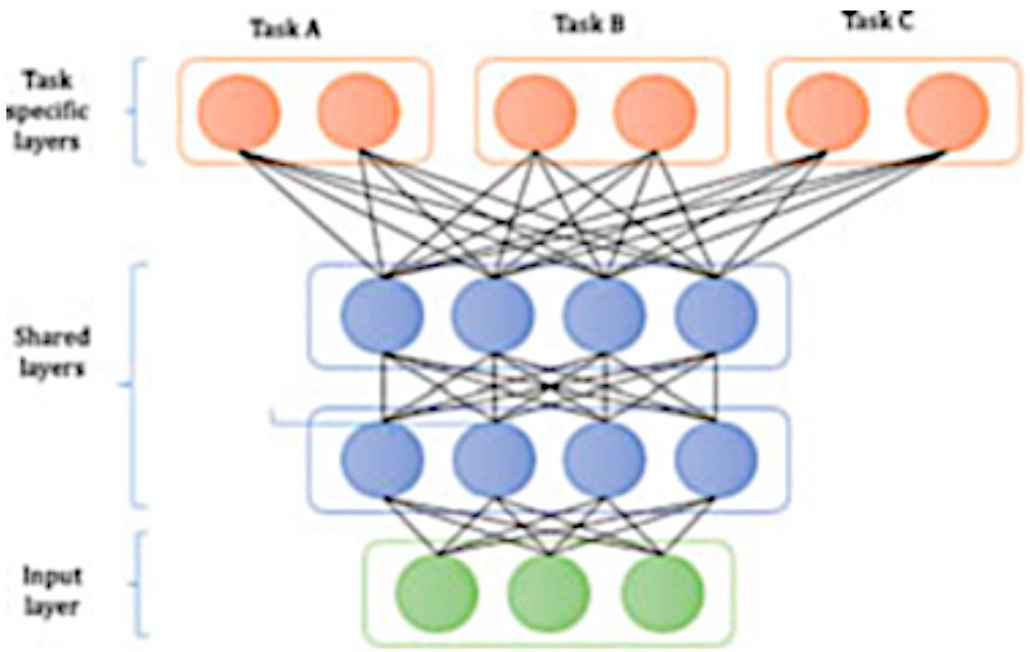
Recap: Transfer Learning Approaches



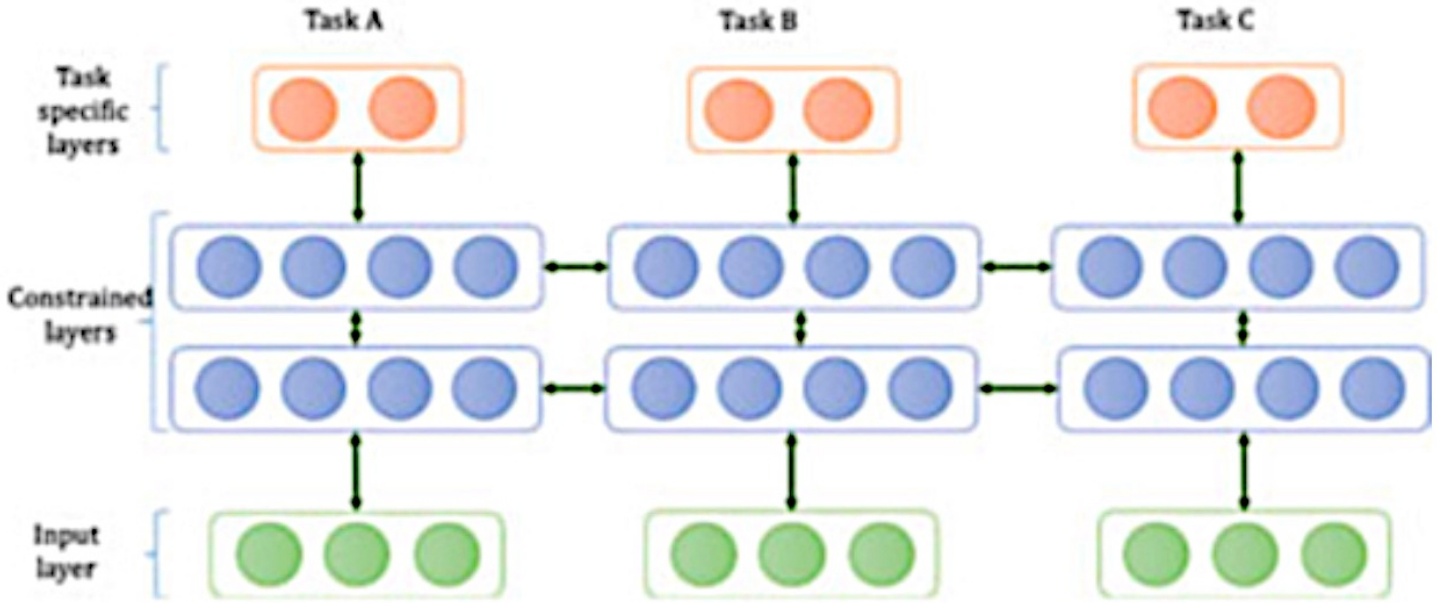
Kamath et al. Deep Learning for NLP and Speech Recognition. 2019.

Approaches

1. Hard parameter sharing



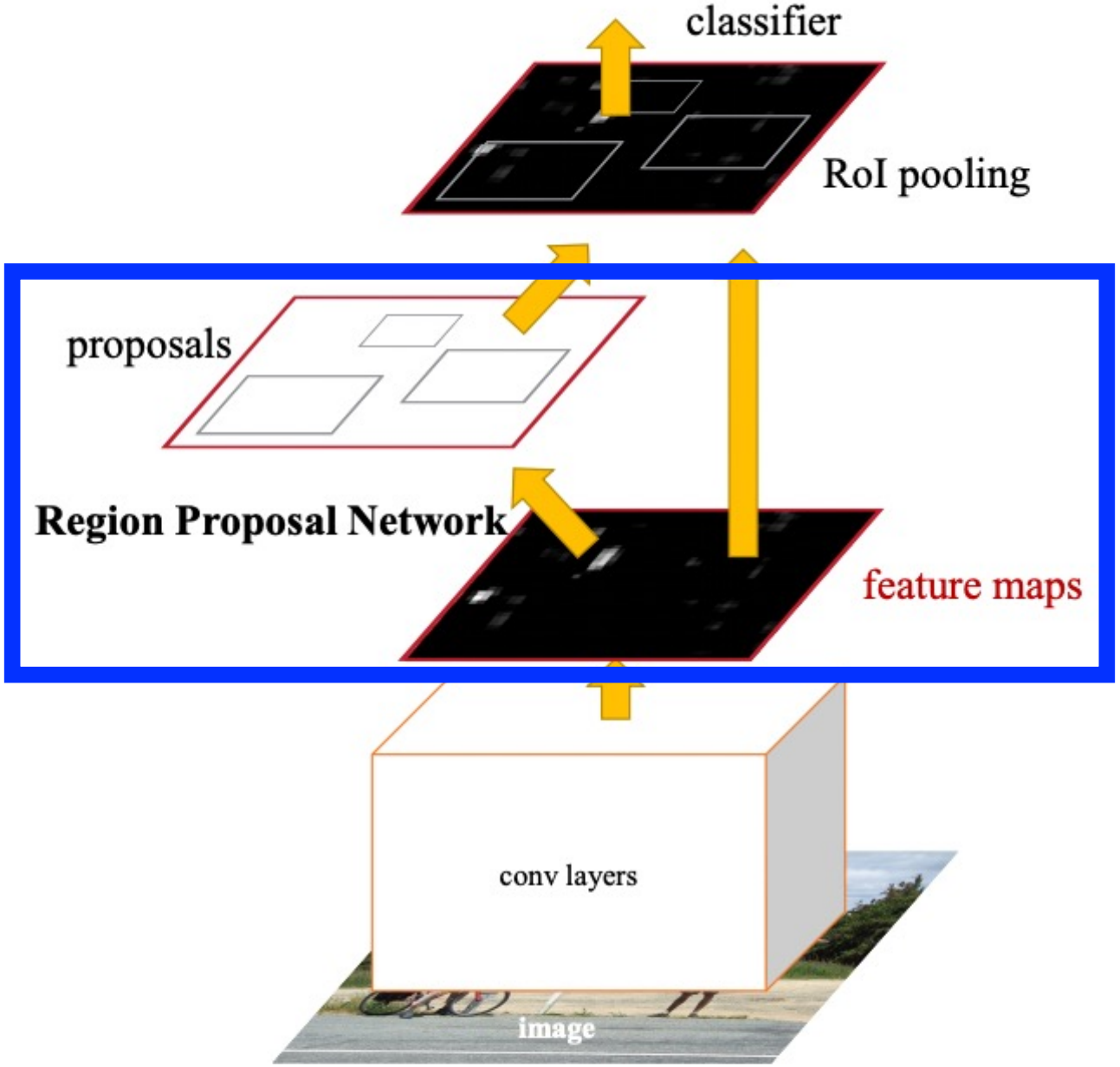
2. Soft parameter sharing



Kamath et al. Deep Learning for NLP and Speech Recognition. 2019.

Recall Faster R-CNN

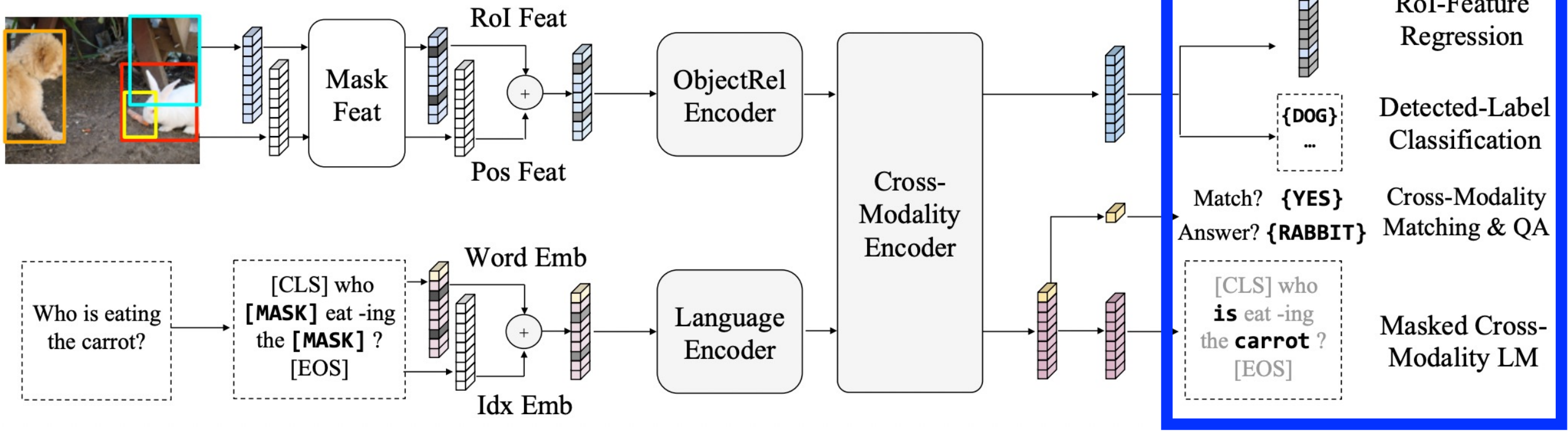
Convolutional layers are shared for region proposal and detection



Ren Shaoqing Ren et al. “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks.” Neurips 2015.

Recall LXMERT

Same architecture used for multiple tasks

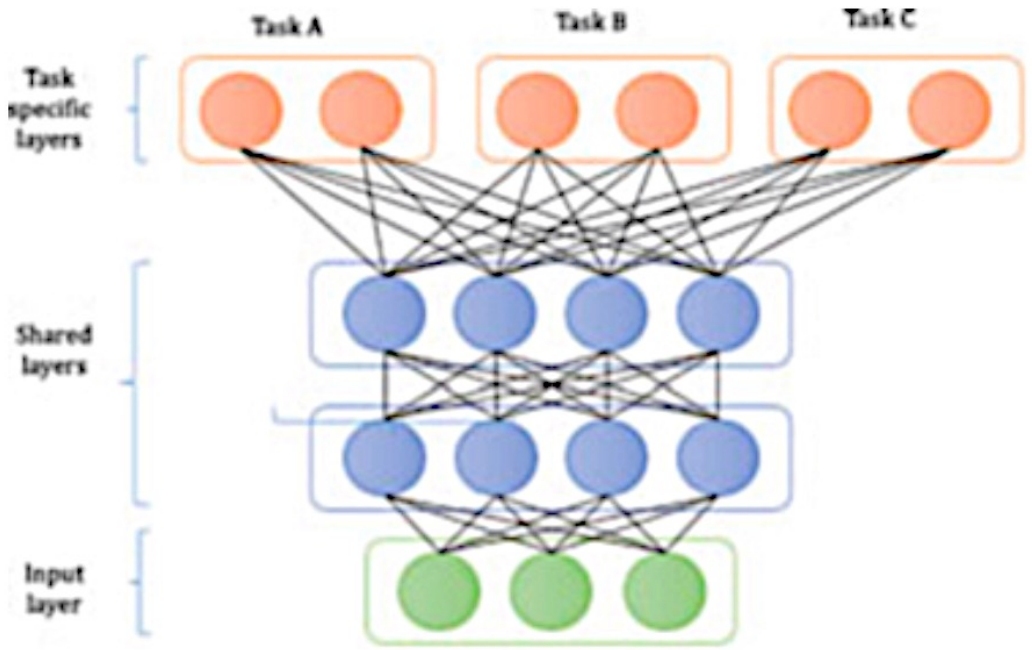


Tan and Bansal, EMNLP 2019.

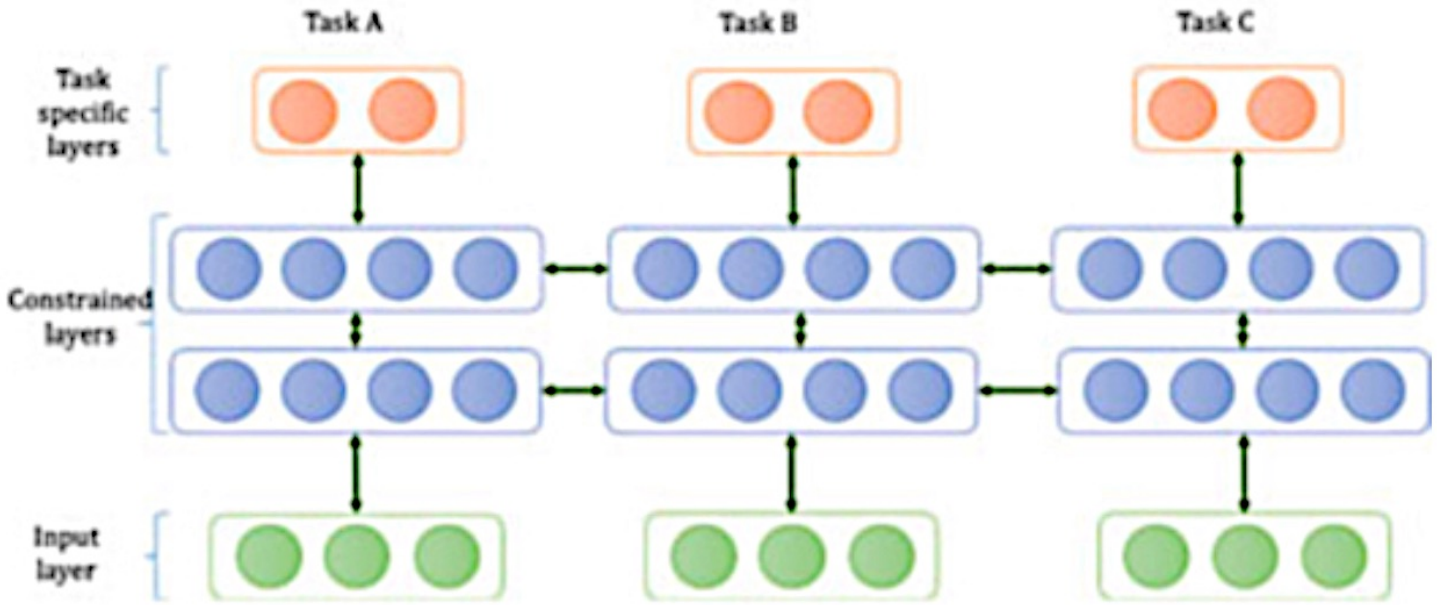
Approaches

Constraints enforce similarity of each task's parameters

1. Hard parameter sharing



2. Soft parameter sharing



Kamath et al. Deep Learning for NLP and Speech Recognition. 2019.

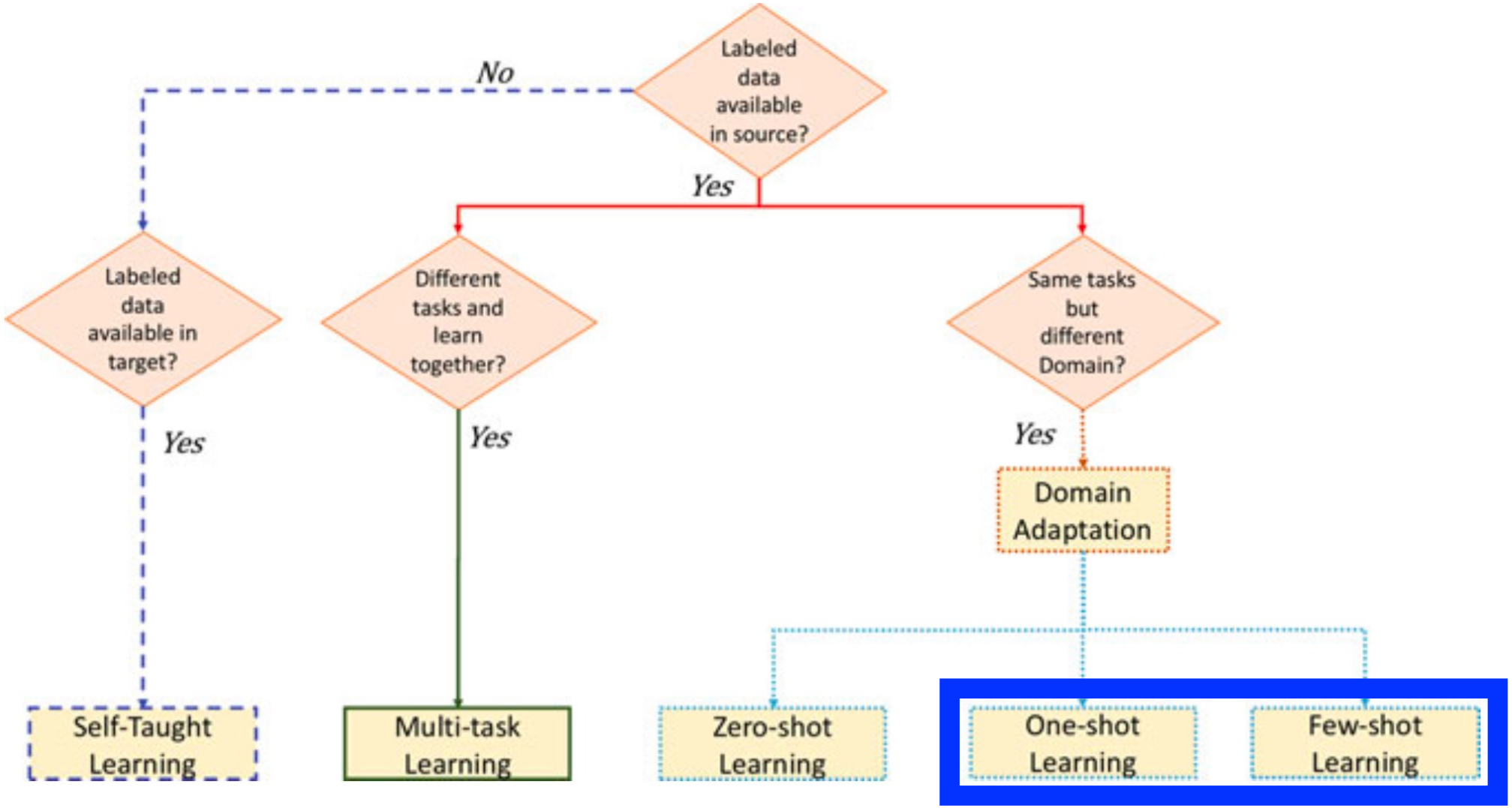
General Benefits from Parameter Sharing?

- Data augmentation
- Enables features found for one task to be available for another
- Emphasizes generalizable features that are common across tasks

Today’s Topics

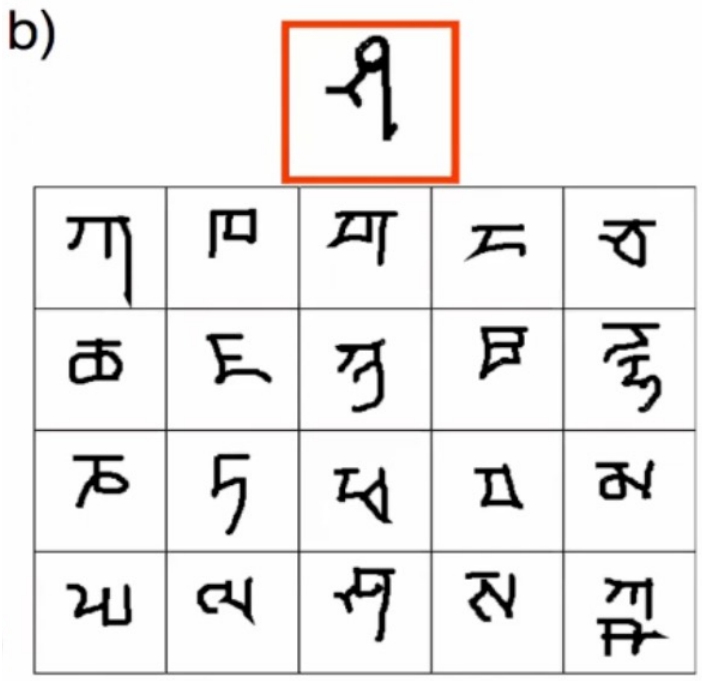
- Multi-task learning
- Few-shot learning
- Zero-shot learning
- Cloud GPU tutorial

Recap: Transfer Learning Approaches



Kamath et al. Deep Learning for NLP and Speech Recognition. 2019.

Intuition: Generalize Current Knowledge

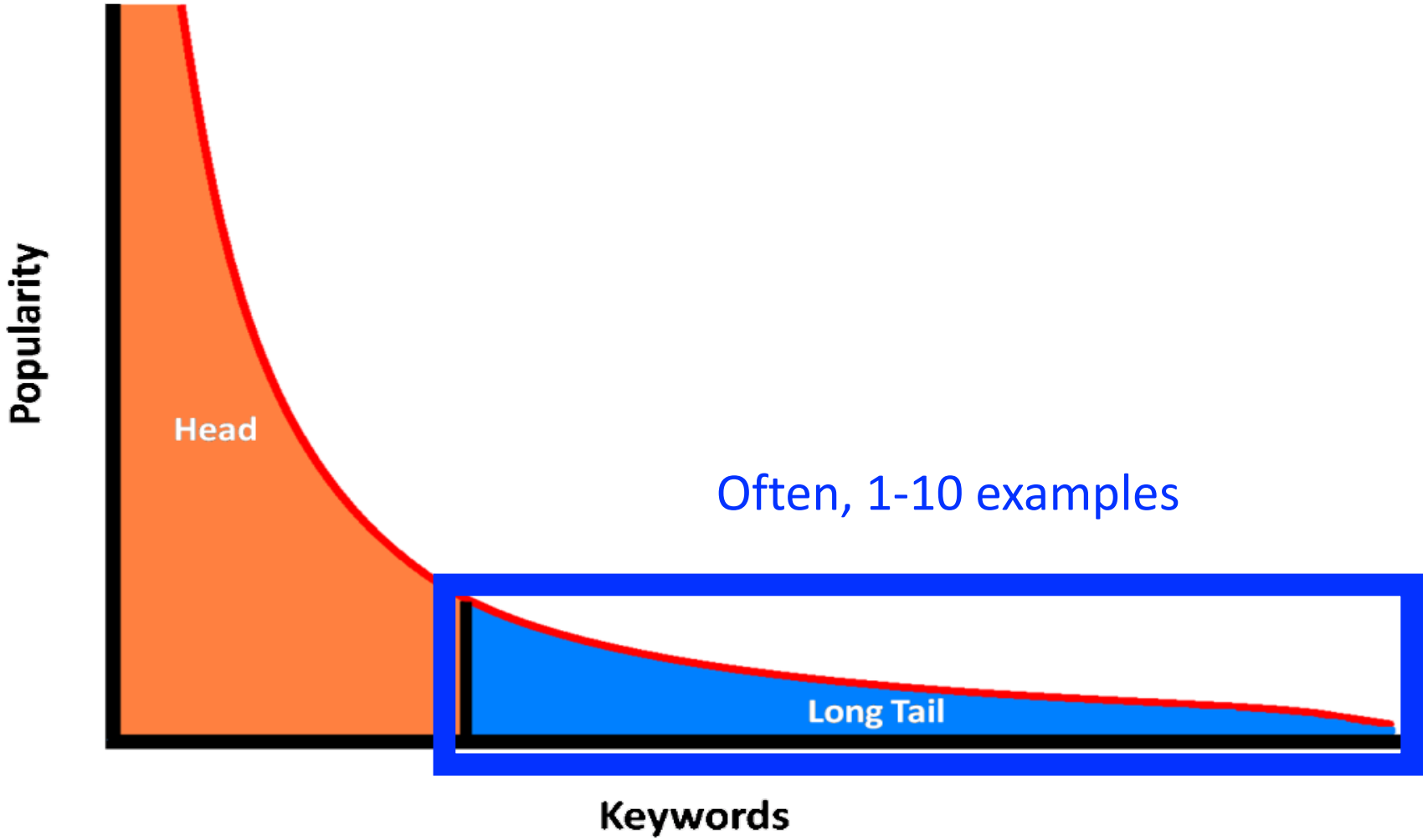


Lake et al, 2013, 2015

Given one example per category, identify the category of the query

https://www.youtube.com/watch?v=9j4iH9TPTd8

Problem Set-up: Learn from Few Examples

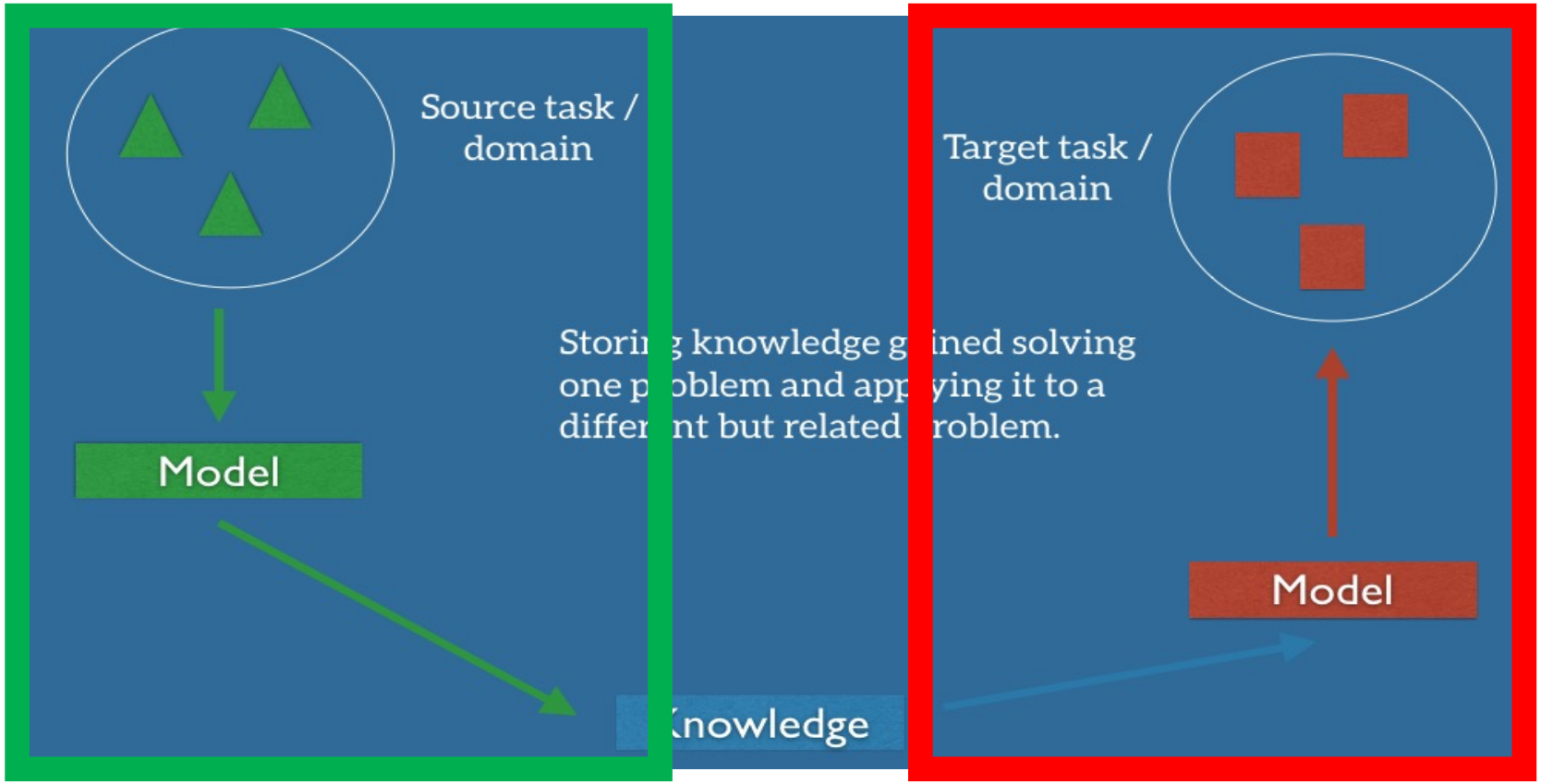


Often, 1-10 examples

https://daredevilmusicproduction.com/long-tail/

Problem Set-up: Learn from Few Examples

Sufficient Labeled Data



Insufficient Labeled Data

- Few shot learning: evaluate only for categories with few examples
- Generalized few shot learning: evaluate on all categories

https://ruder.io/transfer-learning/

Popular Approaches

- Design-time approach: fine-tuning
- Run-time approach: meta learning

Popular Approaches

- Design-time approach: fine-tuning
- Run-time approach: meta learning

Recall Fine-Tuning

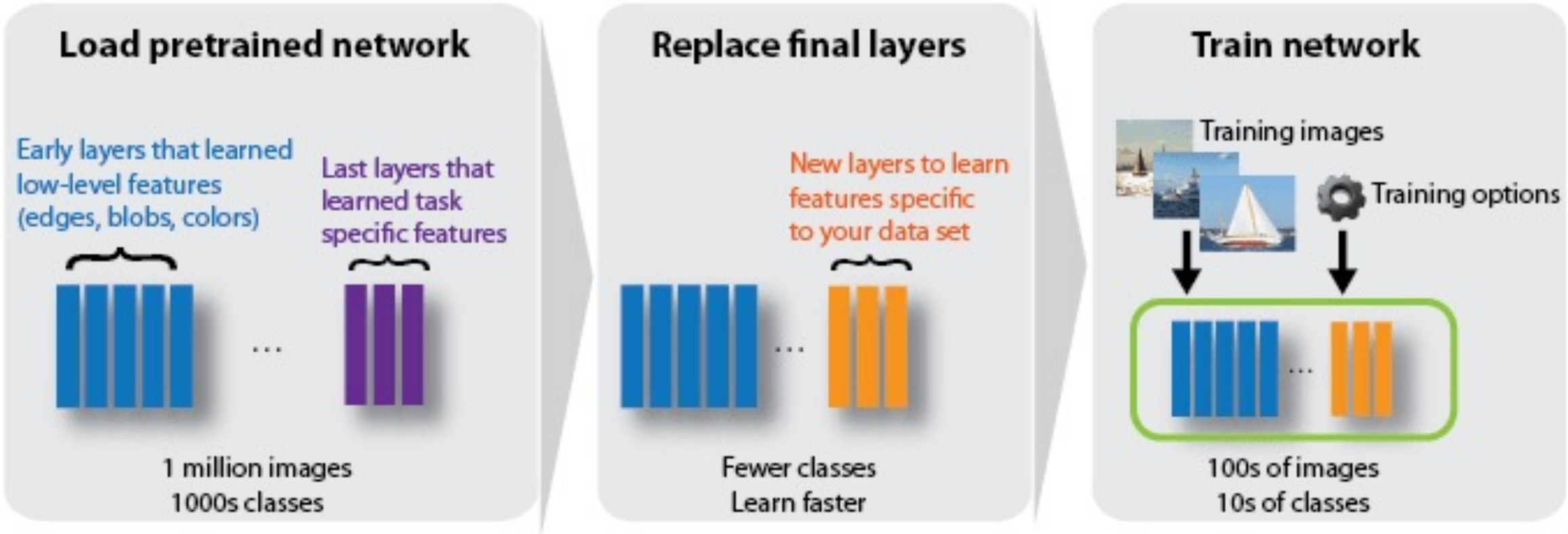
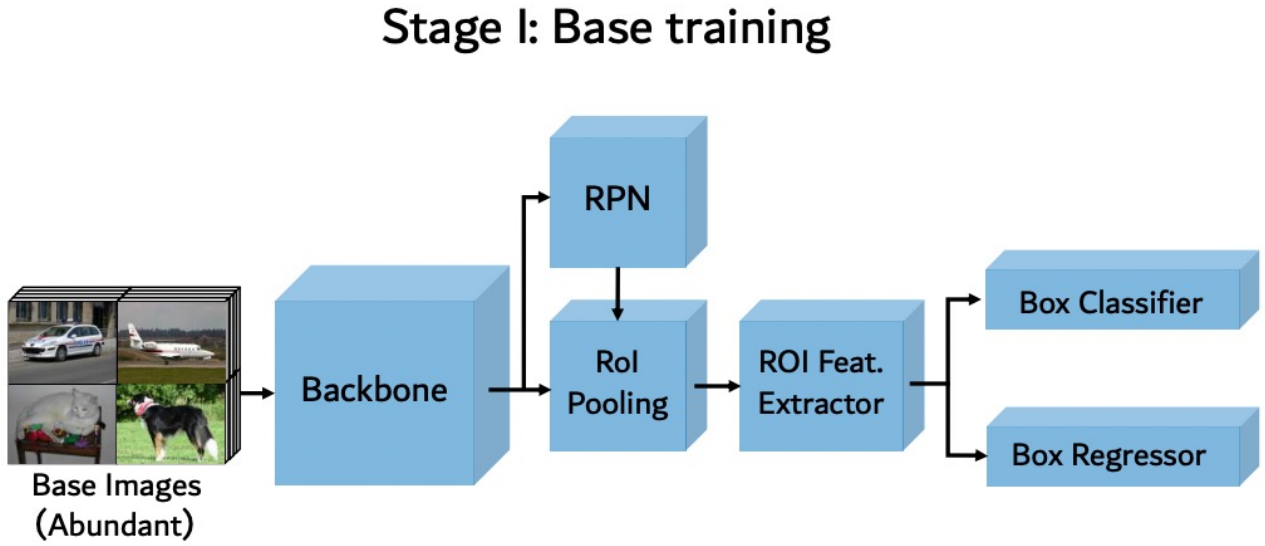
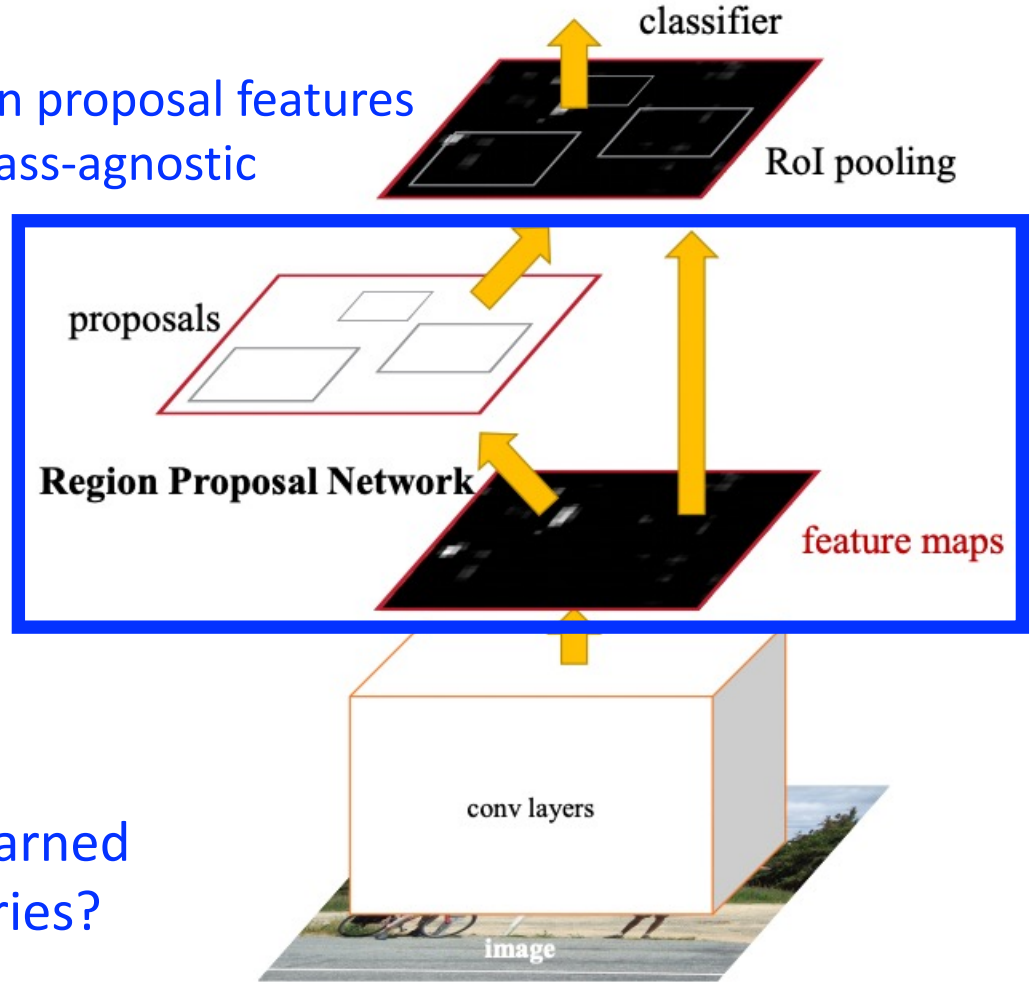


Image Source: https://www.mathworks.com/help/deeplearning/ug/transfer-learning-using-alexnet.html

e.g., Fine-Tuning for Object Detection



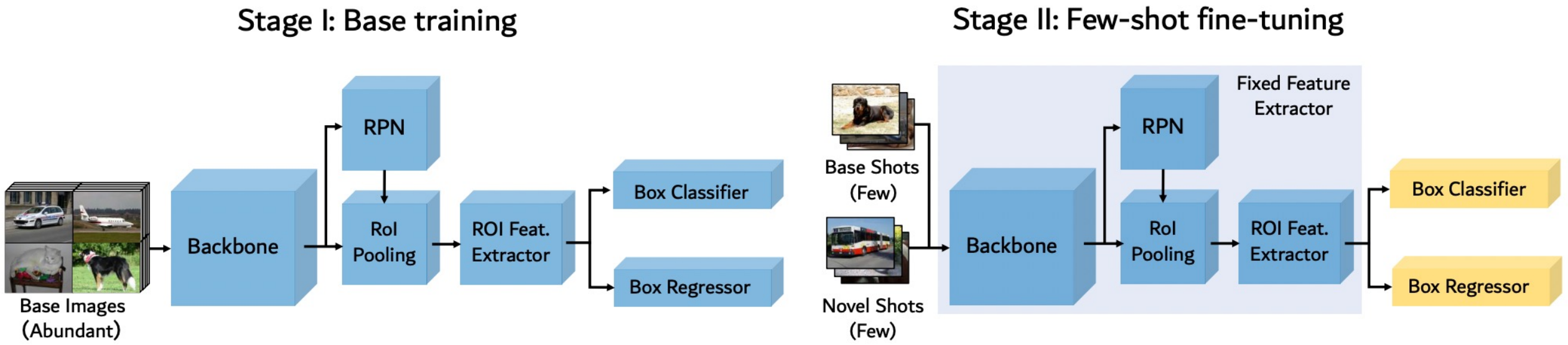
Region proposal features are class-agnostic



Faster R-CNN architecture: Why would we anticipate learned features would generalize well to locating novel categories?

Ren Shaoqing Ren et al. "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks." Neurips 2015.
Wang et al. Frustratingly simple few-shot object detection. arXiv 2020.

e.g., Fine-Tuning for Object Detection



K shots from both base and novel categories used for training

Why include shots from both base and novel categories?

Wang et al. Frustratingly simple few-shot object detection. arXiv 2020.

e.g., Fine-Tuning for Object Detection

Tested with cross validation on 3 splits from VOC

mAP scores for training with 1, 2, 3, 5, and 10 examples (shots) per category

Method / Shot	Backbone	Novel Set 1					Novel Set 2					Novel Set 3				
		1	2	3	5	10	1	2	3	5	10	1	2	3	5	10
YOLO-joint (Kang et al., 2019)	YOLOv2	0.0	0.0	1.8	1.8	1.8	0.0	0.1	0.0	1.8	0.0	1.8	1.8	1.8	3.6	3.9
YOLO-ft (Kang et al., 2019)		3.2	6.5	6.4	7.5	12.3	8.2	3.8	3.5	3.5	7.8	8.1	7.4	7.6	9.5	10.5
YOLO-ft-full (Kang et al., 2019)		6.6	10.7	12.5	24.8	38.6	12.5	4.2	11.6	16.1	33.9	13.0	15.9	15.0	32.2	38.4
FSRW (Kang et al., 2019)		14.8	15.5	26.7	33.9	47.2	15.7	15.3	22.7	30.1	40.5	21.3	25.6	28.4	42.8	45.9
MetaDet (Wang et al., 2019b)		17.1	19.1	28.9	35.0	48.8	18.2	20.6	25.9	30.6	41.5	20.1	22.3	27.9	41.9	42.9
FRCN+joint (Wang et al., 2019b)	FRCN w/VGG16	0.3	0.0	1.2	0.9	1.7	0.0	0.0	1.1	1.9	1.7	0.2	0.5	1.2	1.9	2.8
FRCN+joint-ft (Wang et al., 2019b)		9.1	10.9	13.7	25.0	39.5	10.9	13.2	17.6	19.5	36.5	15.0	15.1	18.3	33.1	35.9
MetaDet (Wang et al., 2019b)		18.9	20.6	30.2	36.8	49.6	21.8	23.1	27.8	31.7	43.0	20.6	23.9	29.4	43.9	44.1
FRCN+joint (Yan et al., 2019)	FRCN w/R-101	2.7	3.1	4.3	11.8	29.0	1.9	2.6	8.1	9.9	12.6	5.2	7.5	6.4	6.4	6.4
FRCN+ft (Yan et al., 2019)		11.9	16.4	29.0	36.9	36.9	5.9	8.5	23.4	29.1	28.8	5.0	9.6	18.1	30.8	43.4
FRCN+ft-full (Yan et al., 2019)		13.8	19.6	32.8	41.5	45.6	7.9	15.3	26.2	31.6	39.1	9.8	11.3	19.1	35.0	45.1
Meta R-CNN (Yan et al., 2019)		19.9	25.5	35.0	45.7	51.5	10.4	19.4	29.6	34.8	**45.4**	14.3	18.2	27.5	41.2	48.1
FRCN+ft-full (Our Impl.)	FRCN w/R-101	15.2	20.3	29.0	40.1	45.5	13.4	20.6	28.6	32.4	38.8	19.6	20.8	28.7	42.2	42.1
TFA w/ fc (Ours)		36.8	29.1	43.6	**55.7**	**57.0**	18.2	**29.0**	33.4	**35.5**	39.0	27.7	33.6	42.5	48.7	**50.2**
TFA w/ cos (Ours)		**39.8**	**36.1**	**44.7**	**55.7**	56.0	**23.5**	26.9	**34.1**	35.1	39.1	**30.8**	**34.8**	**42.8**	**49.5**	49.8

Consistently outperforms baselines by 2-20 points on novel categories

Wang et al. Frustratingly simple few-shot object detection. arXiv 2020.

e.g., Fine-Tuning for Object Detection

Tested with cross validation on 3 splits from VOC

mAP scores for training with 1, 2, 3, 5, and 10 examples (shots) per category

Method / Shot	Backbone	Novel Set 1					Novel Set 2					Novel Set 3				
		1	2	3	5	10	1	2	3	5	10	1	2	3	5	10
YOLO-joint (Kang et al., 2019)	YOLOv2	0.0	0.0	1.8	1.8	1.8	0.0	0.1	0.0	1.8	0.0	1.8	1.8	1.8	3.6	3.9
YOLO-ft (Kang et al., 2019)		3.2	6.5	6.4	7.5	12.3	8.2	3.8	3.5	3.5	7.8	8.1	7.4	7.6	9.5	10.5
YOLO-ft-full (Kang et al., 2019)		6.6	10.7	12.5	24.8	38.6	12.5	4.2	11.6	16.1	33.9	13.0	15.9	15.0	32.2	38.4
FSRW (Kang et al., 2019)		14.8	15.5	26.7	33.9	47.2	15.7	15.3	22.7	30.1	40.5	21.3	25.6	28.4	42.8	45.9
MetaDet (Wang et al., 2019b)		17.1	19.1	28.9	35.0	48.8	18.2	20.6	25.9	30.6	41.5	20.1	22.3	27.9	41.9	42.9
FRCN+joint (Wang et al., 2019b)	FRCN w/VGG16	0.3	0.0	1.2	0.9	1.7	0.0	0.0	1.1	1.9	1.7	0.2	0.5	1.2	1.9	2.8
FRCN+joint-ft (Wang et al., 2019b)		9.1	10.9	13.7	25.0	39.5	10.9	13.2	17.6	19.5	36.5	15.0	15.1	18.3	33.1	35.9
MetaDet (Wang et al., 2019b)		18.9	20.6	30.2	36.8	49.6	21.8	23.1	27.8	31.7	43.0	20.6	23.9	29.4	43.9	44.1
FRCN+joint (Yan et al., 2019)	FRCN w/R-101	2.7	3.1	4.3	11.8	29.0	1.9	2.6	8.1	9.9	12.6	5.2	7.5	6.4	6.4	6.4
FRCN+ft (Yan et al., 2019)		11.9	16.4	29.0	36.9	36.9	5.9	8.5	23.4	29.1	28.8	5.0	9.6	18.1	30.8	43.4
FRCN+ft-full (Yan et al., 2019)		13.8	19.6	32.8	41.5	45.6	7.9	15.3	26.2	31.6	39.1	9.8	11.3	19.1	35.0	45.1
Meta R-CNN (Yan et al., 2019)		19.9	25.5	35.0	45.7	51.5	10.4	19.4	29.6	34.8	**45.4**	14.3	18.2	27.5	41.2	48.1
FRCN+ft-full (Our Impl.)	FRCN w/R-101	15.2	20.3	29.0	40.1	45.5	13.4	20.6	28.6	32.4	38.8	19.6	20.8	28.7	42.2	42.1
TFA w/ fc (Ours)		36.8	29.1	43.6	**55.7**	**57.0**	18.2	**29.0**	33.4	**35.5**	39.0	27.7	33.6	42.5	48.7	**50.2**
TFA w/ cos (Ours)		**39.8**	**36.1**	**44.7**	**55.7**	56.0	**23.5**	26.9	**34.1**	35.1	39.1	**30.8**	**34.8**	**42.8**	**49.5**	49.8

Similar performance boosts also observed on two more datasets (COCO and LVIS)

Wang et al. Frustratingly simple few-shot object detection. arXiv 2020.

Fine-Tuning

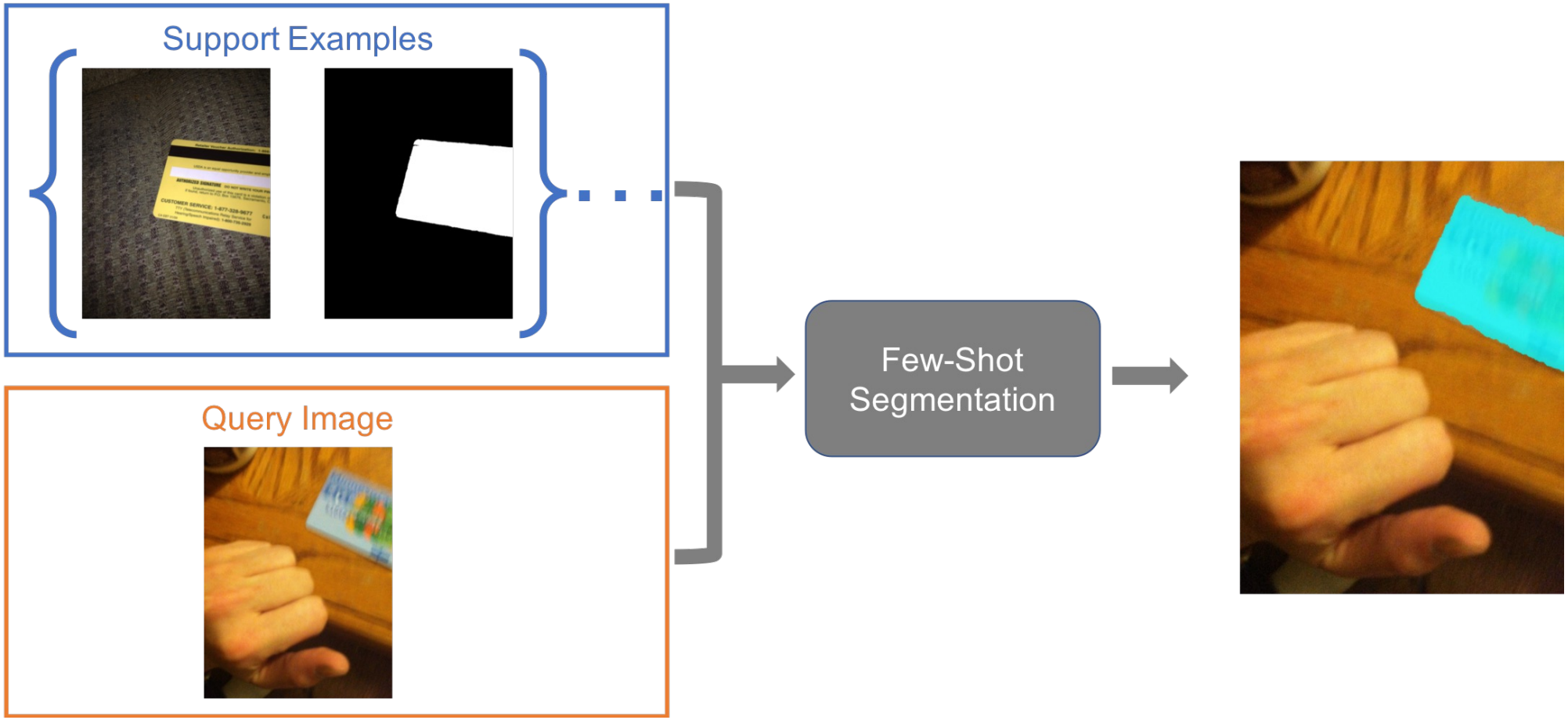
What are limitations of this approach for real-world applications?

- Must retrain algorithm to add new categories

Popular Approaches

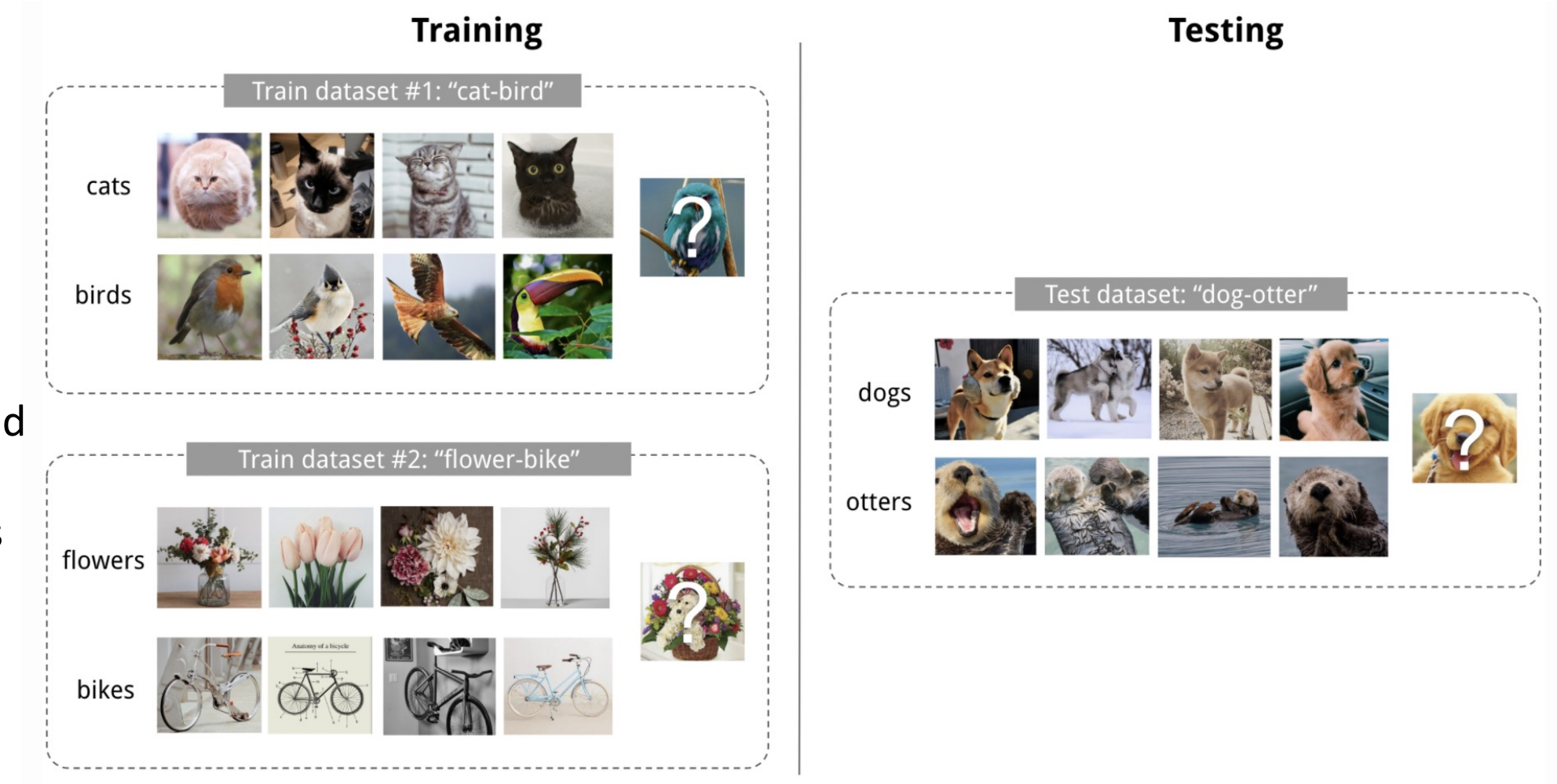
- Design-time approach: fine-tuning
- Run-time approach: meta learning

Meta Learner: Update Model with Support Set



Implementation: Trained Model Updates Itself to Generalize to Support Set Categories

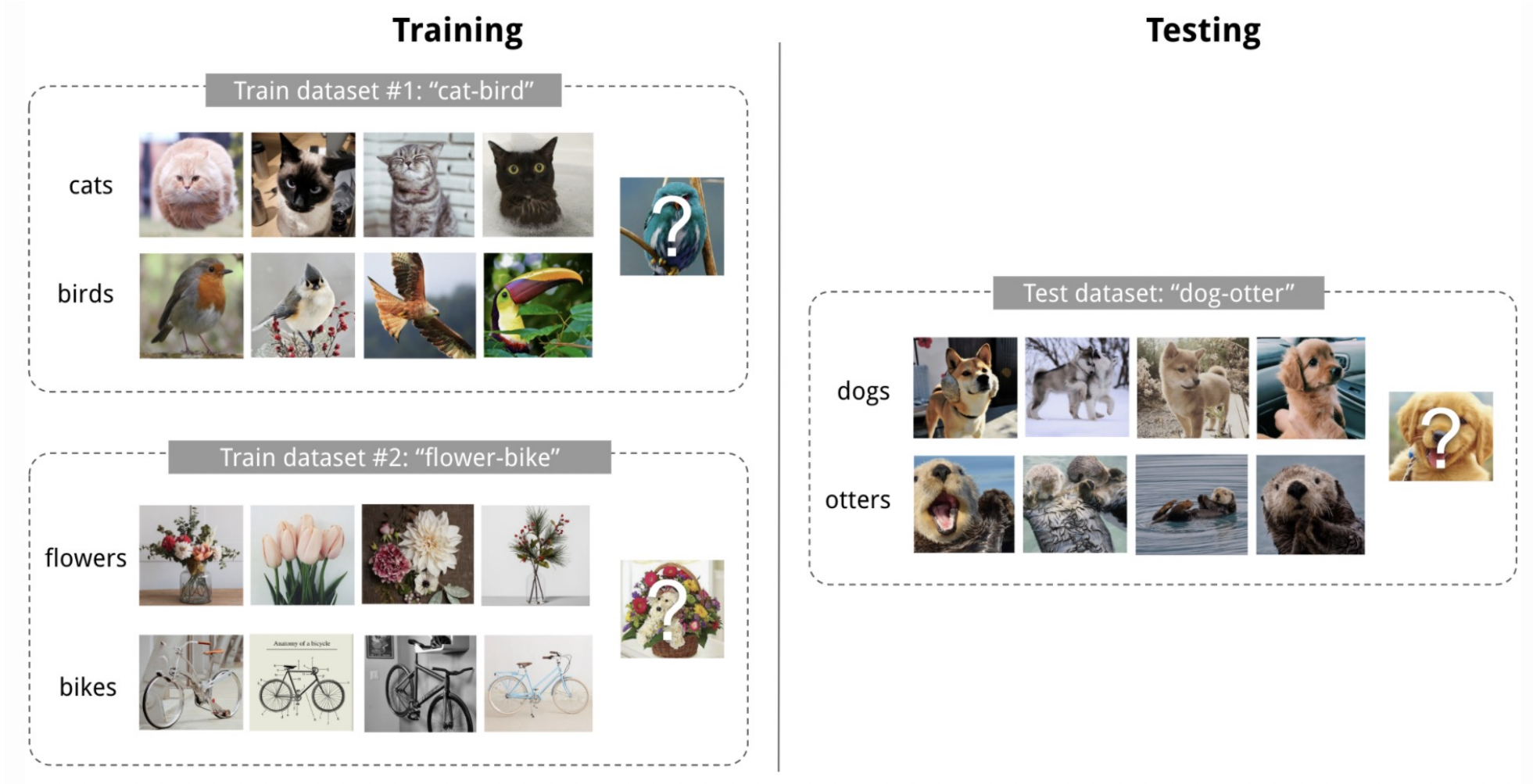
Goal: learn features during training that are class-agnostic and so can generalize to novel test categories



https://lilianweng.github.io/lil-log/2018/11/30/meta-learning.html

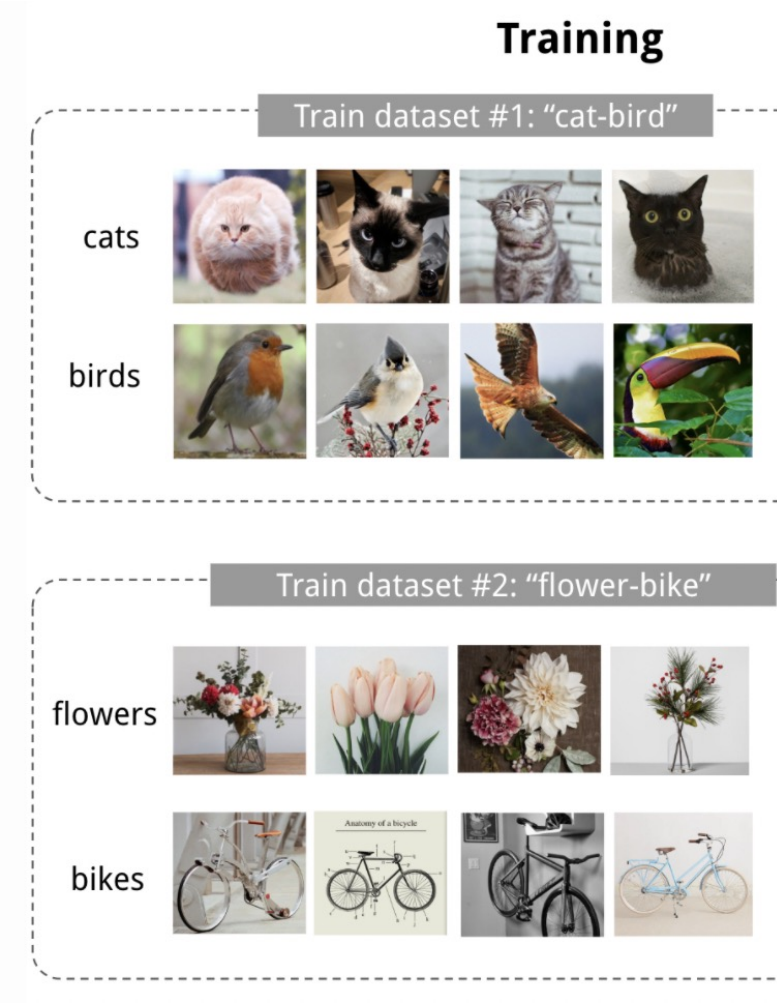
Implementation: Trained Model Updates Itself to Generalize to Support Set Categories

How many shots are observed at testing?



https://lilianweng.github.io/lil-log/2018/11/30/meta-learning.html

Implementation: Trained Model Updates Itself to Generalize to Support Set Categories

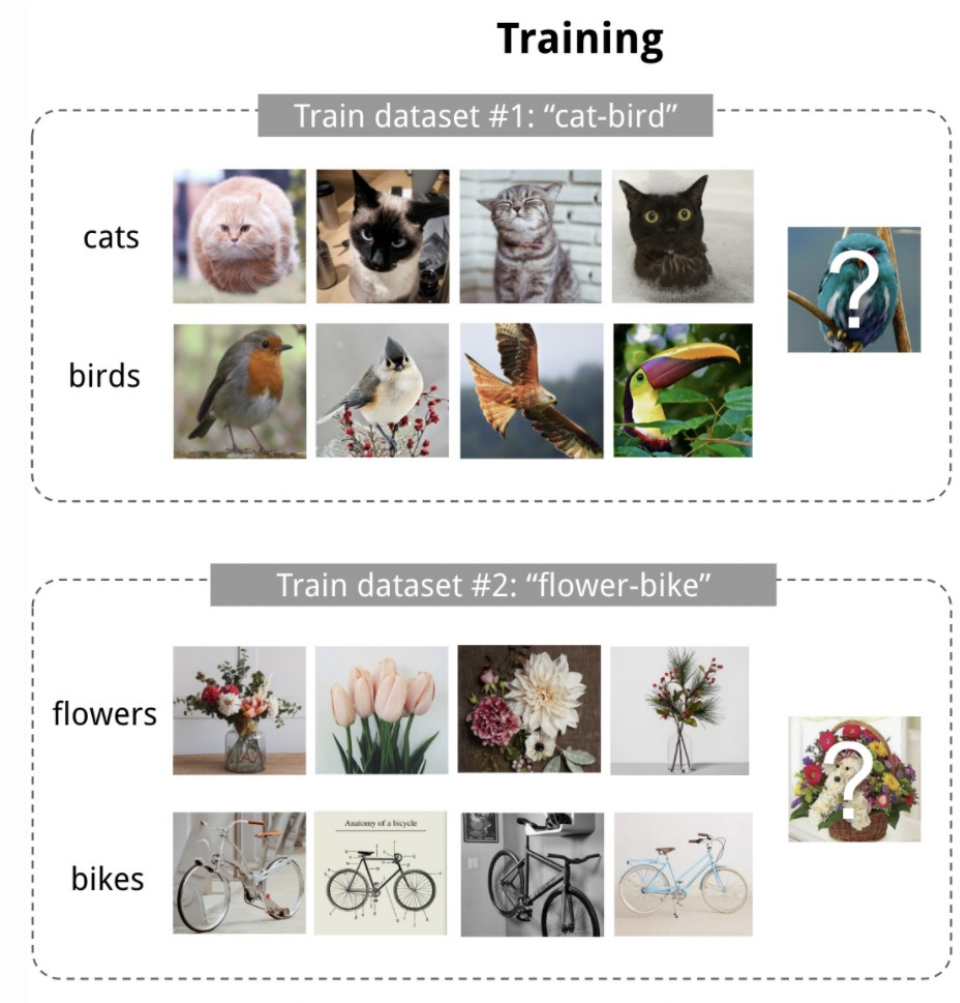


How many "shots" should be observed at each training round?

- 4 (must match test time)

https://lilianweng.github.io/lil-log/2018/11/30/meta-learning.html

Implementation: Trained Model Updates Itself to Generalize to Support Set Categories



Given support categories, detect which one the "query" matches

In the example, how many categories would a trained model support?

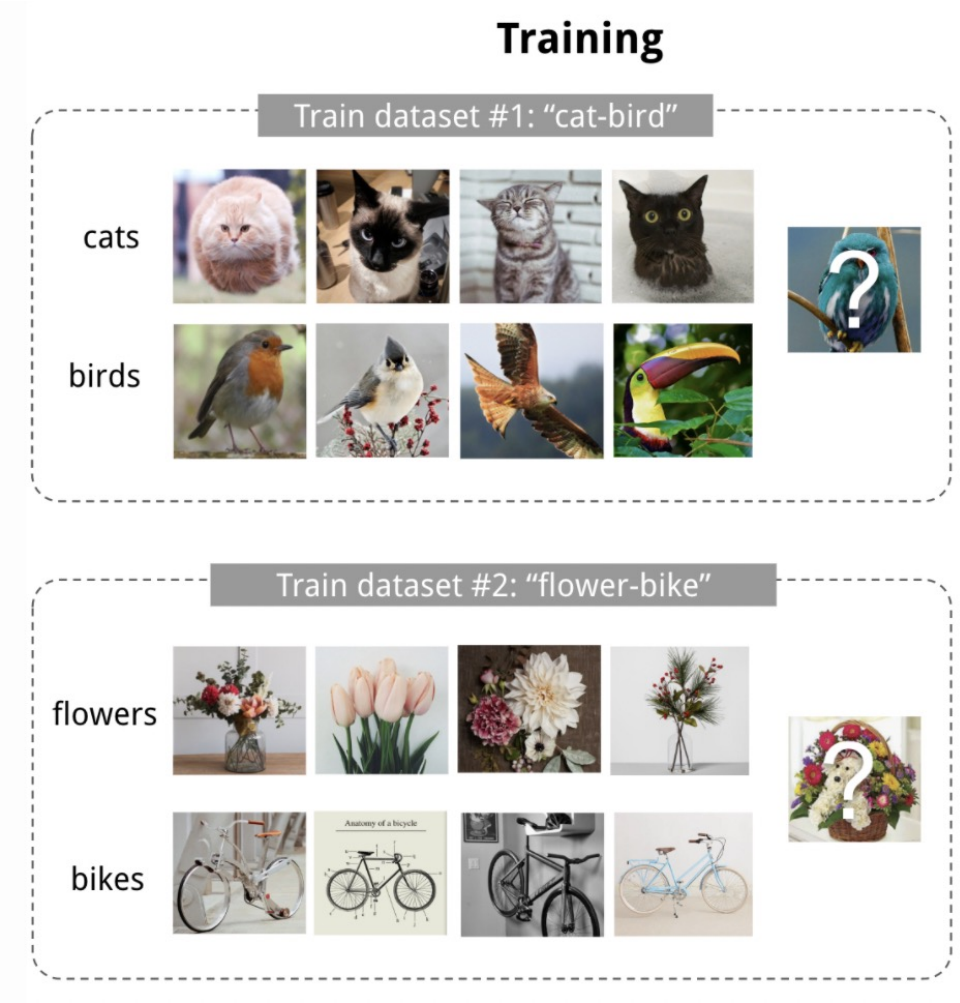
https://lilianweng.github.io/lil-log/2018/11/30/meta-learning.html

Implementation: Trained Model Updates Itself to Generalize to Support Set Categories

Recall support categories are never observed during training

https://lilianweng.github.io/lil-log/2018/11/30/meta-learning.html

Implementation: Trained Model Updates Itself to Generalize to Support Set Categories

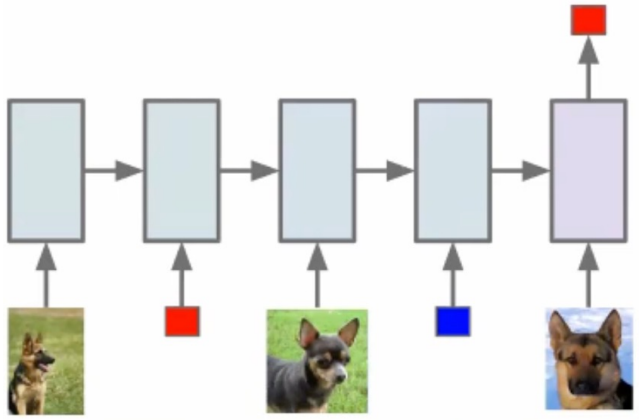


How to train a model to do this?

https://lilianweng.github.io/lil-log/2018/11/30/meta-learning.html

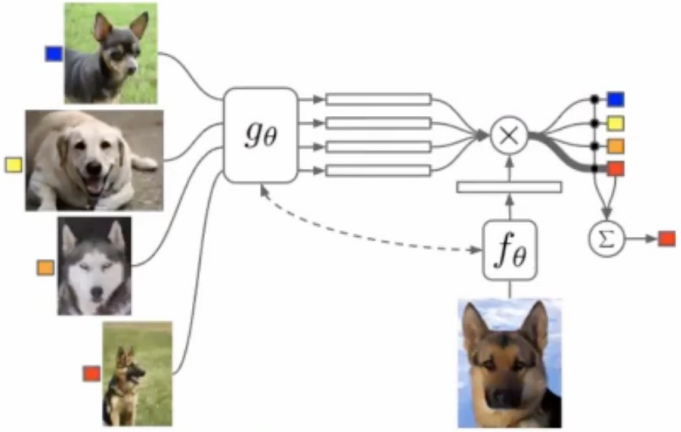
Implementation: Trained Model Updates Itself to Generalize to Support Set Categories

Model Based



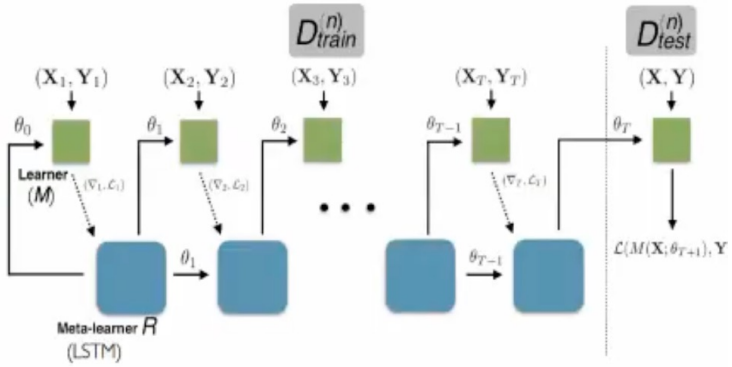
- Santoro et al. '16
- Duan et al. '17
- Wang et al. '17
- Munkhdalai & Yu '17
- Mishra et al. '17
- ...

Metric Based



- Koch '15
- Vinyals et al. '16
- Snell et al. '17
- Shyam et al. '17
- Sung et al. '17
- ...

Optimization Based

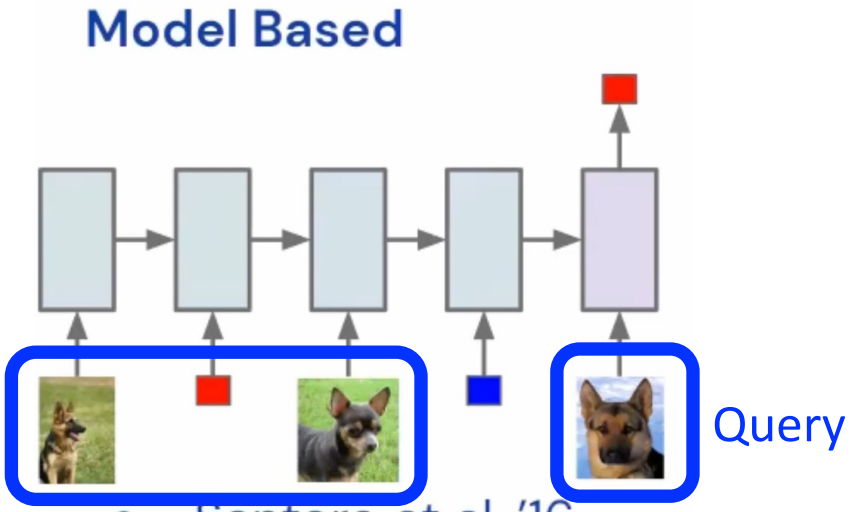


- Schmidhuber '87, '92
- Bengio et al. '90, '92
- Hochreiter et al. '01
- Li & Malik '16
- Andrychowicz et al. '16
- Ravi & Larochelle '17
- Finn et al. '17
- ...

Adapted from Finn '17

https://www.youtube.com/watch?v=9j4iH9TPTd8

Implementation: Trained Model Updates Itself to Generalize to Support Set Categories

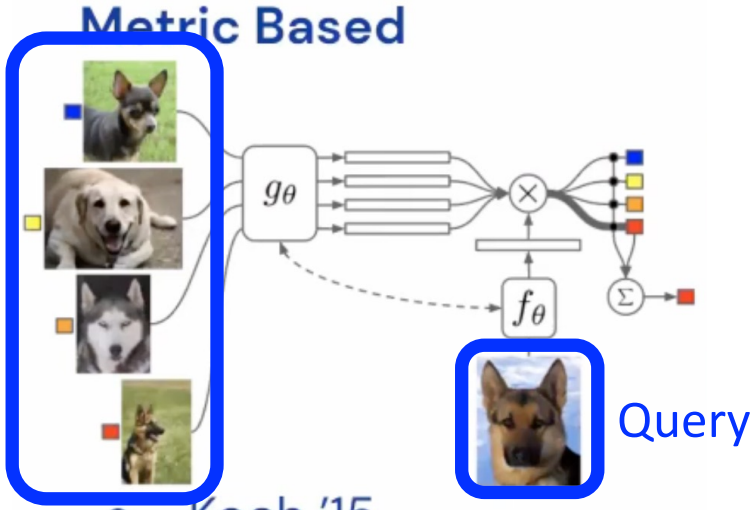


Adapted from Finn '17

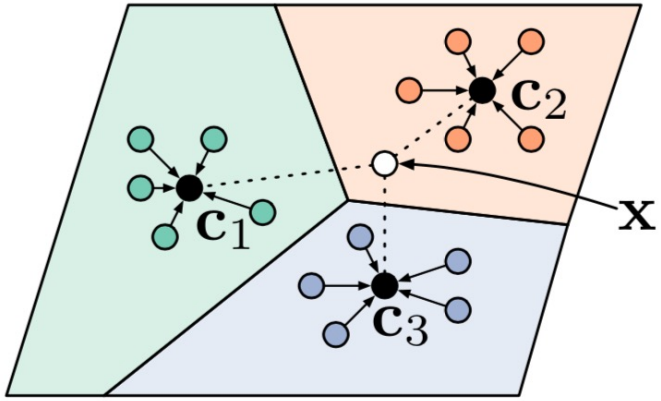
e.g., learn set-invariant neural networks, such as those that rely on attention, to locate similarity

https://www.youtube.com/watch?v=9j4iH9TPTd8

Implementation: Trained Model Updates Itself to Generalize to Support Set Categories



Compare query to each support category; e.g., establish a "prototype" for each support set

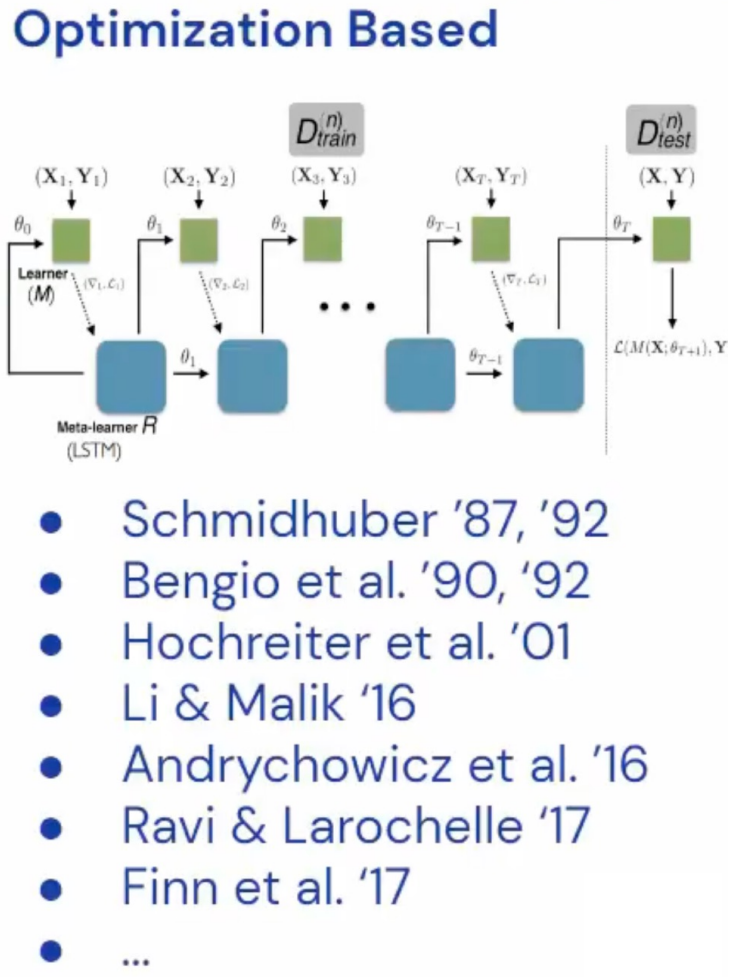


https://lilianweng.github.io/posts/2018-11-30-meta-learning/

https://www.youtube.com/watch?v=9j4iH9TPTd8

Implementation: Trained Model Updates Itself to Generalize to Support Set Categories

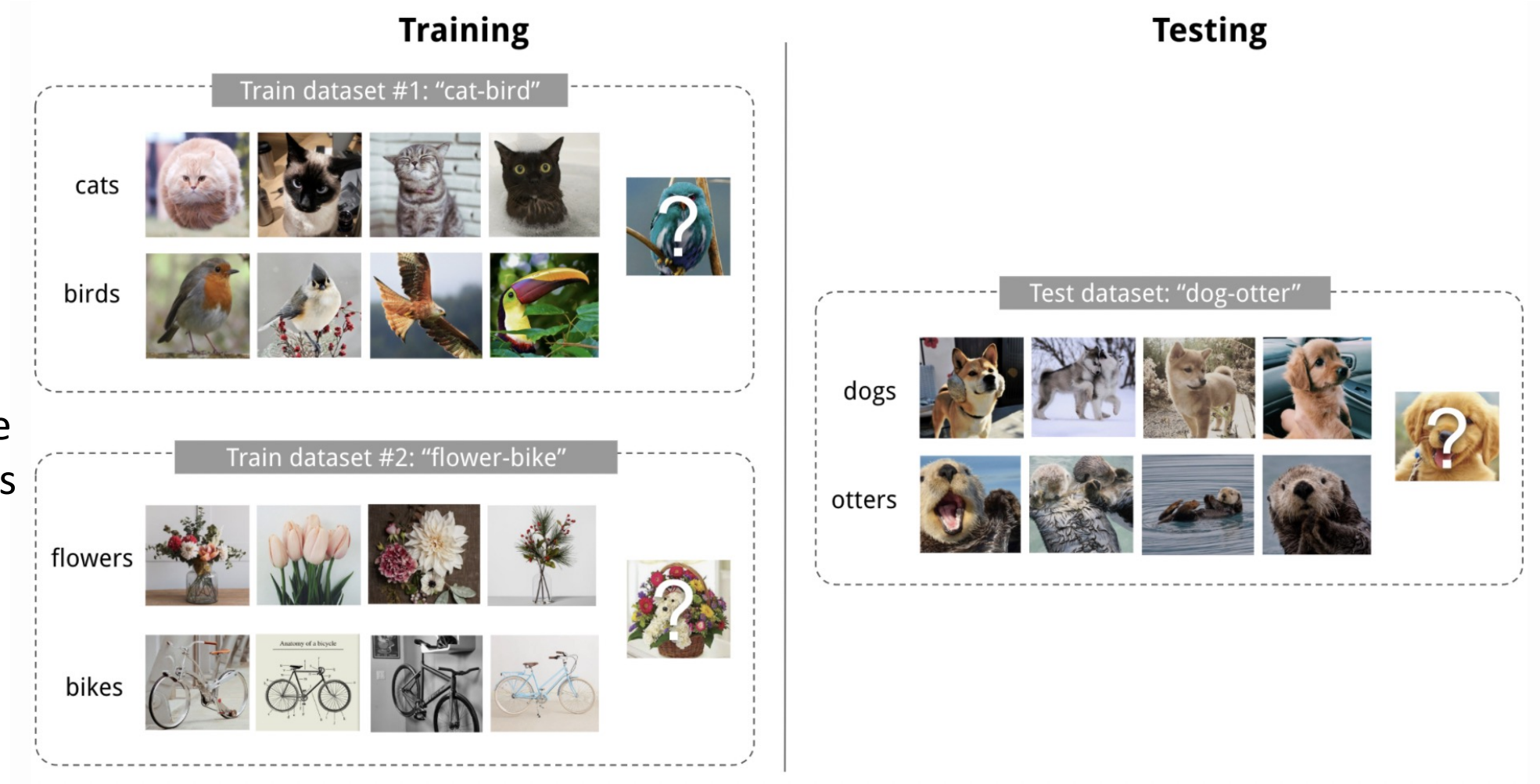
Function to optimize is conditioned on the support set; e.g., tweak “forget” gate of LSTM



https://www.youtube.com/watch?v=9j4iH9TPTd8

Implementation: Trained Model Updates Itself to Generalize to Support Set Categories

Post-training, model can be deployed in the wild on new categories



https://lilianweng.github.io/lil-log/2018/11/30/meta-learning.html

Meta Learner: Update Model with Support Set

What are limitations of this approach for real-world applications?

- Requires large amount of memory to process the support set on top of the query set

Bronskill et al. Memory Efficient Meta-Learning with Large Images. Neurips 2021.

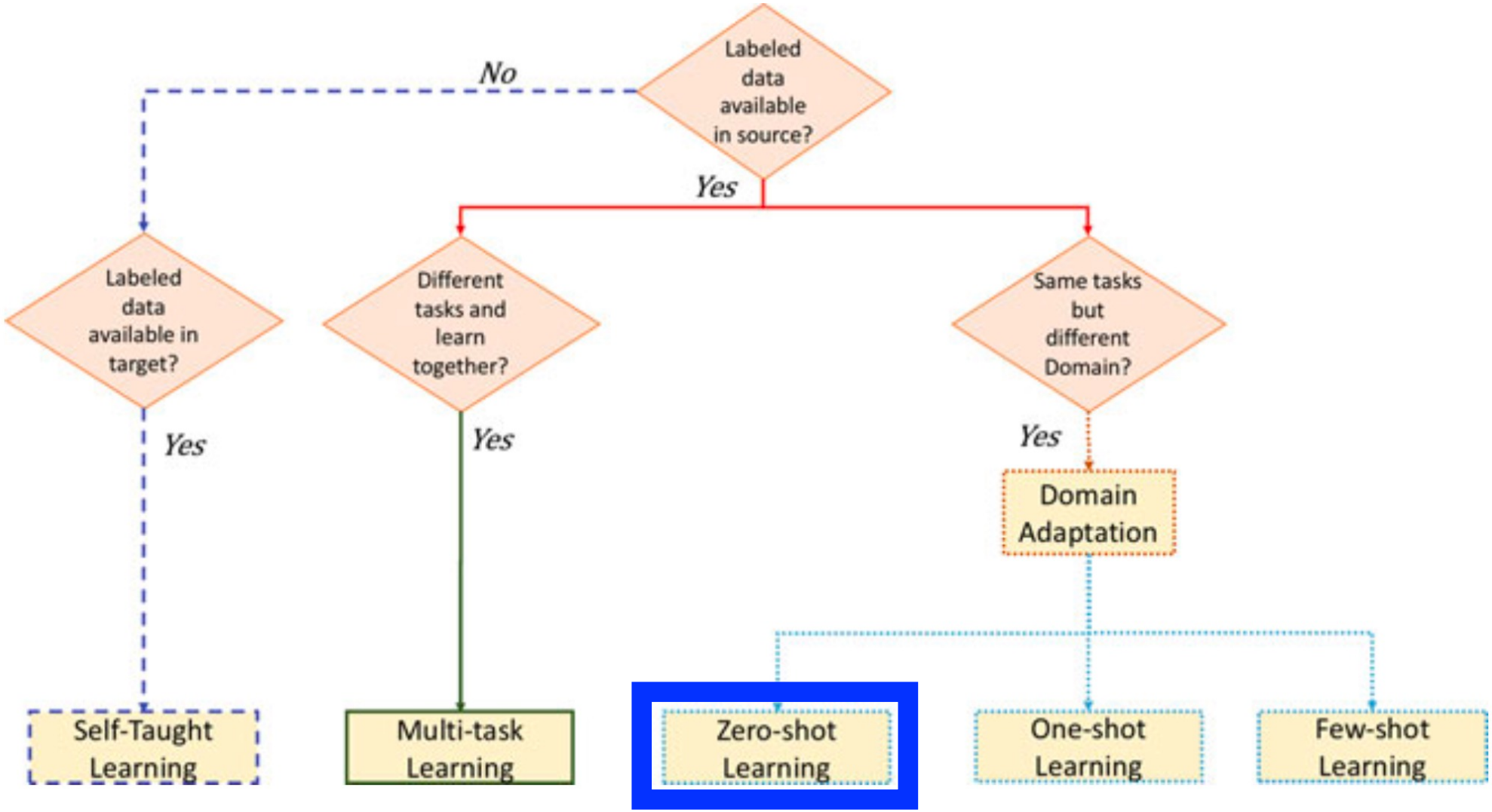
Popular Approaches

- Design-time approach: fine-tuning
- Run-time approach: meta learning

Today’s Topics

- Multi-task learning
- Few-shot learning
- Zero-shot learning
- Cloud GPU tutorial

Recap: Transfer Learning Approaches



Kamath et al. Deep Learning for NLP and Speech Recognition. 2019.

Intuition: Generalize Current Knowledge to Quickly Generalize to New Categories

What is this?

How many examples do you think you would need to see to recognize another one of these?

Intuition: Generalize Current Knowledge to Quickly Generalize to New Categories

Could see 0 examples if you knew the object fuses a person on top with a horse on the bottom

Intuition: Generalize Current Knowledge to Quickly Generalize to New Categories

Could see 0 examples of a zebra if you knew it looks like a horse with black and white stripes

Key Idea: Learn from Auxiliary Labels How to Perform a Different Task with Zero Training Examples

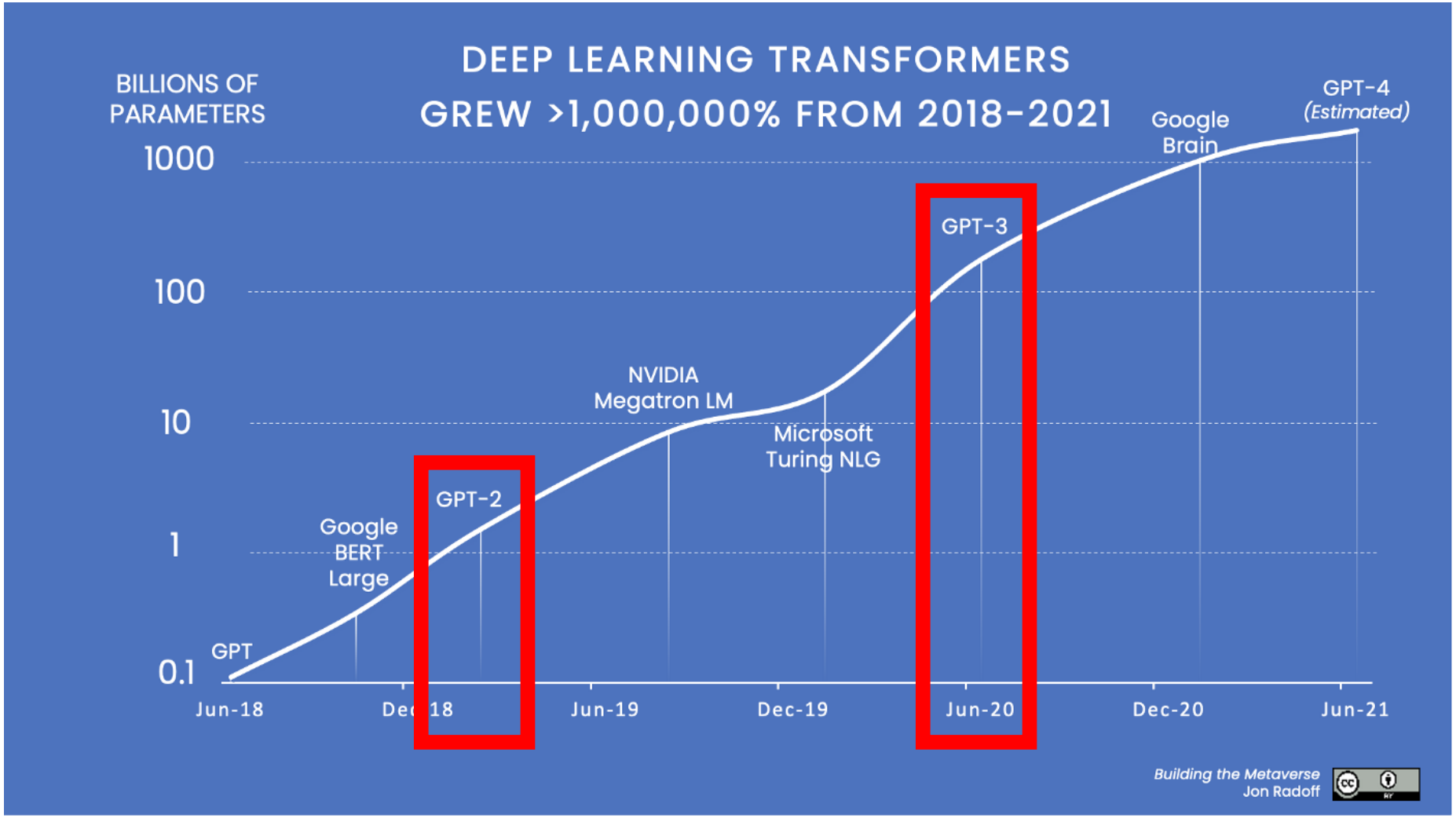
Popular Approaches

- Text representation: GPT-2 and GPT-3
- Joint image-language representation: CLIP

Popular Approaches

- Text representation: GPT-2 and GPT-3
- Joint image-language representation: CLIP

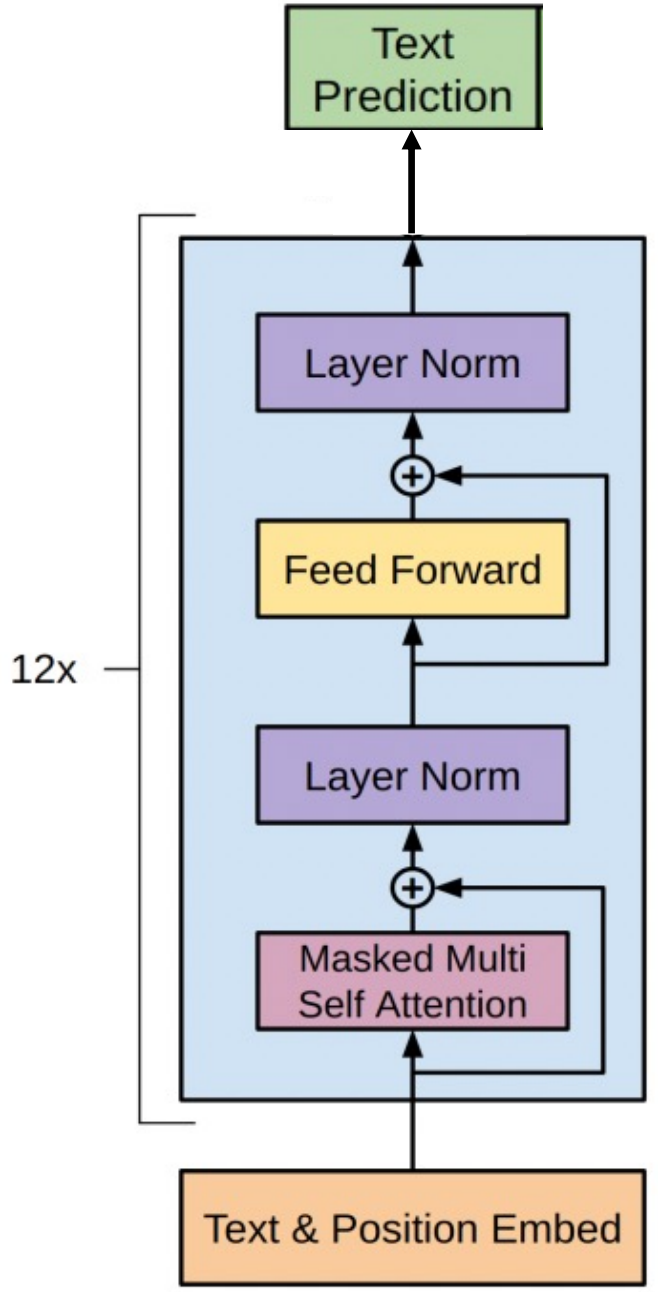
Recall: Growing Size of Transformers



https://medium.com/building-the-metaverse/the-metaverse-and-artificial-intelligence-ai-577343895411

Architecture for Every GPT-n

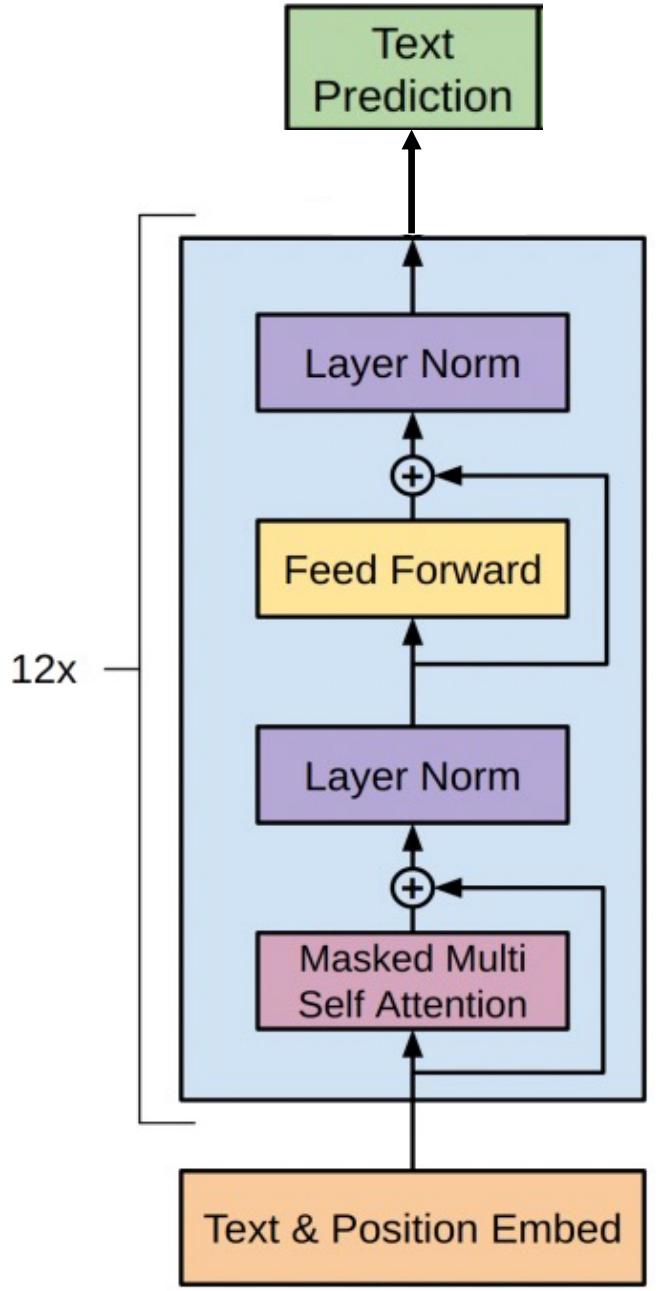
GPT-2/3 use same decoder as GPT-1, which comes from original transformer, but with MANY more parameters (e.g., 175 billion) by increasing context size and number of stacked blocks



Radford et al. Improving Language Understanding by Generative Pre-Training. Technical Report 2018.

Pretraining GPT-n

Increased amounts of training data are used and training is performed for longer durations



Radford et al. Improving Language Understanding by Generative Pre-Training. Technical Report 2018.

Pretraining Task for Every GPT-n: Predict Next Word Given Previous Ones

e.g.,

1. Background music from a _____
2. Many people danced around the _____
3. I practiced for many years to learn how to play the _____

Zero/Few-Shot Performance Evaluated on Over 24 Datasets

e.g., Translation

Zero-shot

The model predicts the answer given only a natural language description of the task. No gradient updates are performed.

```
1  Translate English to French:        ← task description
2  cheese =>                           ← prompt
```

Brown et al. Language Models are Few-Shot Learners. Neurips 2020.

e.g., News Article Generation

Article titles and subtitles provided as initialization:

Title: United Methodists Agree to Historic Split
Subtitle: Those who oppose gay marriage will form their own denomination
Article: **After two days of intense debate, the United Methodist Church has agreed to a historic split - one that is expected to end in the creation of a new denomination, one that will be "theologically and socially conservative," according to The Washington Post. The majority of delegates attending the church's annual General Conference in May voted to strengthen a ban on the ordination of LGBTQ clergy and to write new rules that will "discipline" clergy who officiate at same-sex weddings. But those who opposed these measures have a new plan: They say they will form a separate denomination by 2020, calling their church the Christian Methodist denomination.**
The Post notes that the denomination, which claims 12.5 million members, was in the early 20th century the "largest Protestant denomination in the U.S.," but that it has been shrinking in recent decades. The new split will be the second in the church's history. The first occurred in 1968, when roughly 10 percent of the denomination left to form the Evangelical United Brethren Church. The Post notes that the proposed split "comes at a critical time for the church, which has been losing members for years," which has been "pushed toward the brink of a schism over the role of LGBTQ people in the church." Gay marriage is not the only issue that has divided the church. In 2016, the denomination was split over ordination of transgender clergy, with the North Pacific regional conference voting to ban them from serving as clergy, and the South Pacific regional conference voting to allow them.

Brown et al. Language Models are Few-Shot Learners. Neurips 2020.

e.g., News Article Generation

Humans could distinguish GPT-3 generated articles from real articles for only 52% of shown articles

Title: United Methodists Agree to Historic Split
Subtitle: Those who oppose gay marriage will form their own denomination
Article: **After two days of intense debate, the United Methodist Church has agreed to a historic split - one that is expected to end in the creation of a new denomination, one that will be "theologically and socially conservative," according to The Washington Post. The majority of delegates attending the church's annual General Conference in May voted to strengthen a ban on the ordination of LGBTQ clergy and to write new rules that will "discipline" clergy who officiate at same-sex weddings. But those who opposed these measures have a new plan: They say they will form a separate denomination by 2020, calling their church the Christian Methodist denomination.**
The Post notes that the denomination, which claims 12.5 million members, was in the early 20th century the "largest Protestant denomination in the U.S.," but that it has been shrinking in recent decades. The new split will be the second in the church's history. The first occurred in 1968, when roughly 10 percent of the denomination left to form the Evangelical United Brethren Church. The Post notes that the proposed split "comes at a critical time for the church, which has been losing members for years," which has been "pushed toward the brink of a schism over the role of LGBTQ people in the church." Gay marriage is not the only issue that has divided the church. In 2016, the denomination was split over ordination of transgender clergy, with the North Pacific regional conference voting to ban them from serving as clergy, and the South Pacific regional conference voting to allow them.

Brown et al. Language Models are Few-Shot Learners. Neurips 2020.

e.g., News Article Generation

Given GPT's ability to convincingly generate misinformation, do you think the model should be publicly-available?

Title: United Methodists Agree to Historic Split
Subtitle: Those who oppose gay marriage will form their own denomination
Article: **After two days of intense debate, the United Methodist Church has agreed to a historic split - one that is expected to end in the creation of a new denomination, one that will be "theologically and socially conservative," according to The Washington Post. The majority of delegates attending the church's annual General Conference in May voted to strengthen a ban on the ordination of LGBTQ clergy and to write new rules that will "discipline" clergy who officiate at same-sex weddings. But those who opposed these measures have a new plan: They say they will form a separate denomination by 2020, calling their church the Christian Methodist denomination.**
The Post notes that the denomination, which claims 12.5 million members, was in the early 20th century the "largest Protestant denomination in the U.S.," but that it has been shrinking in recent decades. The new split will be the second in the church's history. The first occurred in 1968, when roughly 10 percent of the denomination left to form the Evangelical United Brethren Church. The Post notes that the proposed split "comes at a critical time for the church, which has been losing members for years," which has been "pushed toward the brink of a schism over the role of LGBTQ people in the church." Gay marriage is not the only issue that has divided the church. In 2016, the denomination was split over ordination of transgender clergy, with the North Pacific regional conference voting to ban them from serving as clergy, and the South Pacific regional conference voting to allow them.

Brown et al. Language Models are Few-Shot Learners. Neurips 2020.

More Generally, Prompt-Based Methods

Pre-train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing

Pengfei Liu
Carnegie Mellon University
pliu3@cs.cmu.edu

Weizhe Yuan
Carnegie Mellon University
weizhey@cs.cmu.edu

Jinlan Fu
National University of Singapore
jinlanjonna@gmail.com

Zhengbao Jiang
Carnegie Mellon University
zhengbaj@cs.cmu.edu

Hiroaki Hayashi
Carnegie Mellon University
hiroakih@cs.cmu.edu

Graham Neubig
Carnegie Mellon University
gneubig@cs.cmu.edu

More Generally, Prompt-Based Methods

Type	Task	Input ([X])	Template	Answer ([Z])
Text CLS	Sentiment	I love this movie.	[X] The movie is [Z].	great fantastic ...
	Topics	He prompted the LM.	[X] The text is about [Z].	sports science ...
	Intention	What is taxi fare to Denver?	[X] The question is about [Z].	quantity city ...
Text-span CLS	Aspect Sentiment	Poor service but good food.	[X] What about service? [Z].	Bad Terrible ...
Text-pair CLS	NLI	[X1]: An old man with ... [X2]: A man walks ...	[X1]? [Z], [X2]	Yes No ...
Tagging	NER	[X1]: Mike went to Paris. [X2]: Paris	[X1][X2] is a [Z] entity.	organization location ...
Text Generation	Summarization	Las Vegas police ...	[X] TL;DR: [Z]	The victim ... A woman ... ...
	Translation	Je vous aime.	French: [X] English: [Z]	I love you. I fancy you. ...

Liu et al. arXiv 2021.

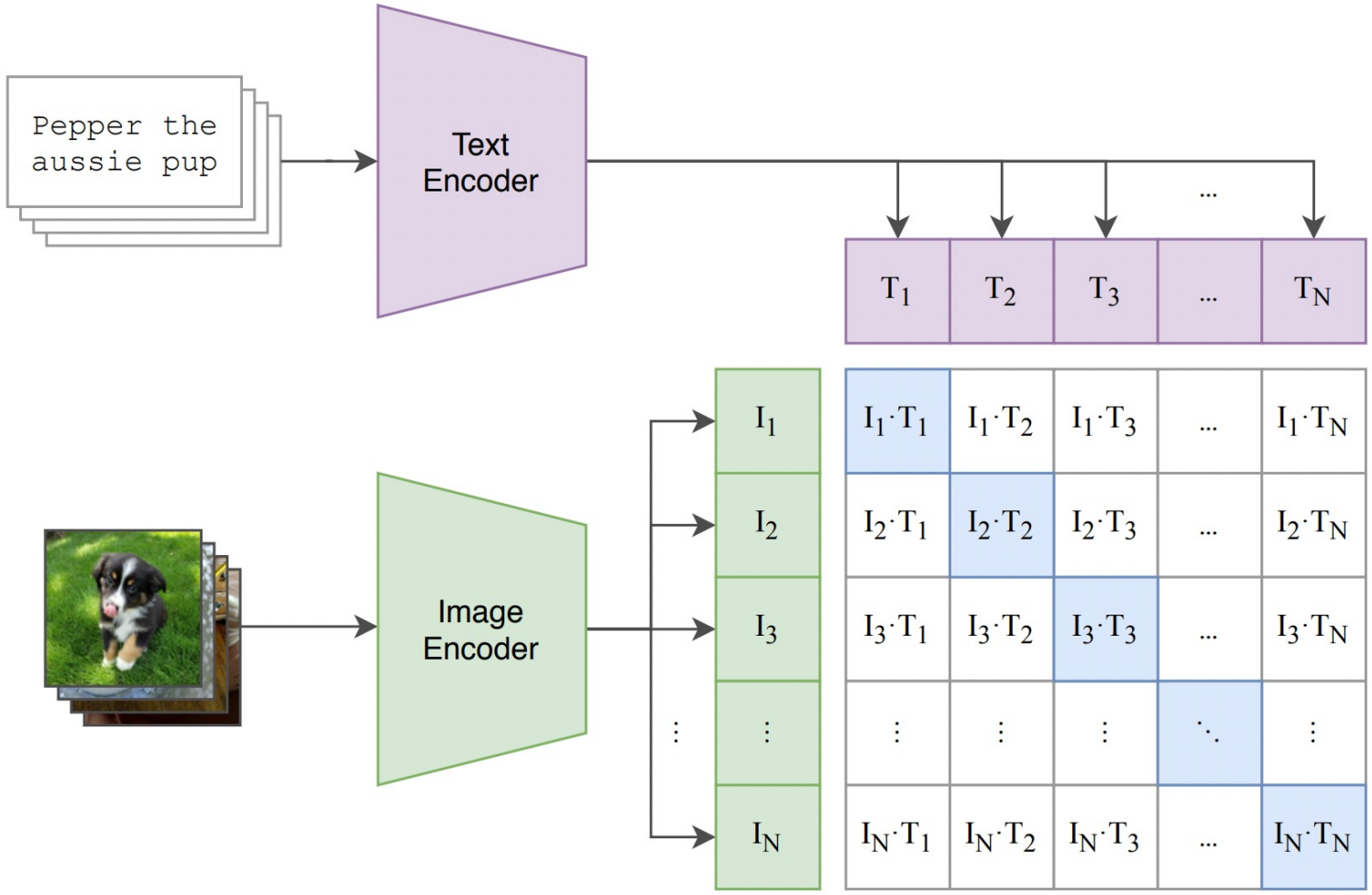
Popular Approaches

- Text representation: GPT-2 and GPT-3
- Joint image-language representation: CLIP

Contrastive Language-Image Pretraining (CLIP)

Pre-training: Learns with 400 million image-text pairs found on the Internet (i.e., unstructured auxiliary data) to represent an image encoder and a text encoder similarly for similar content (and so similar encodings for similar visual content)

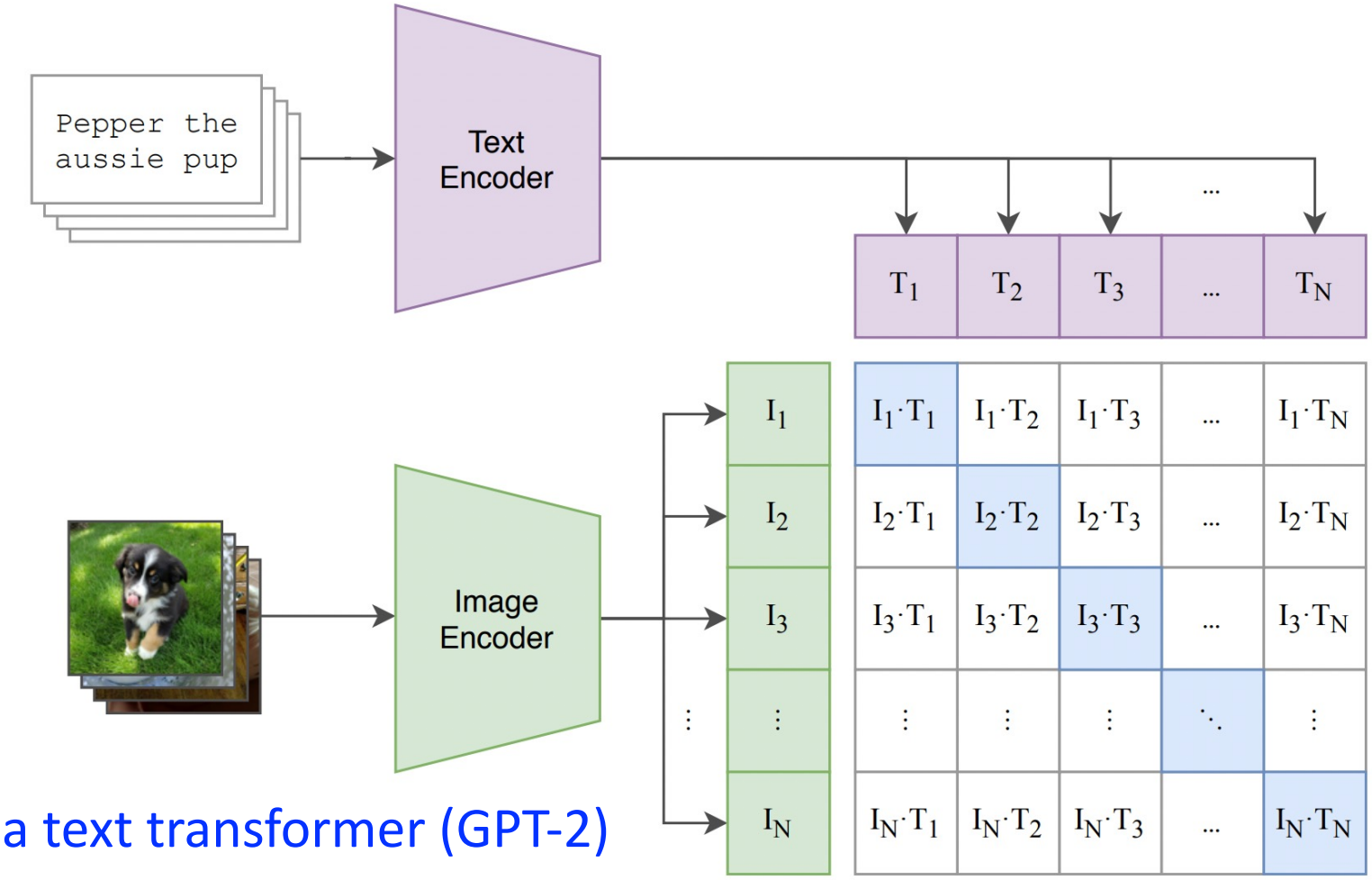
(predicts which image-text pairs match for all combinations)



Radford et al. Learning Transferable Visual Models From Natural Language Supervision. ICML 2021.

Contrastive Language-Image Pretraining (CLIP): Architecture

Tested with a visual transformer (ViT)



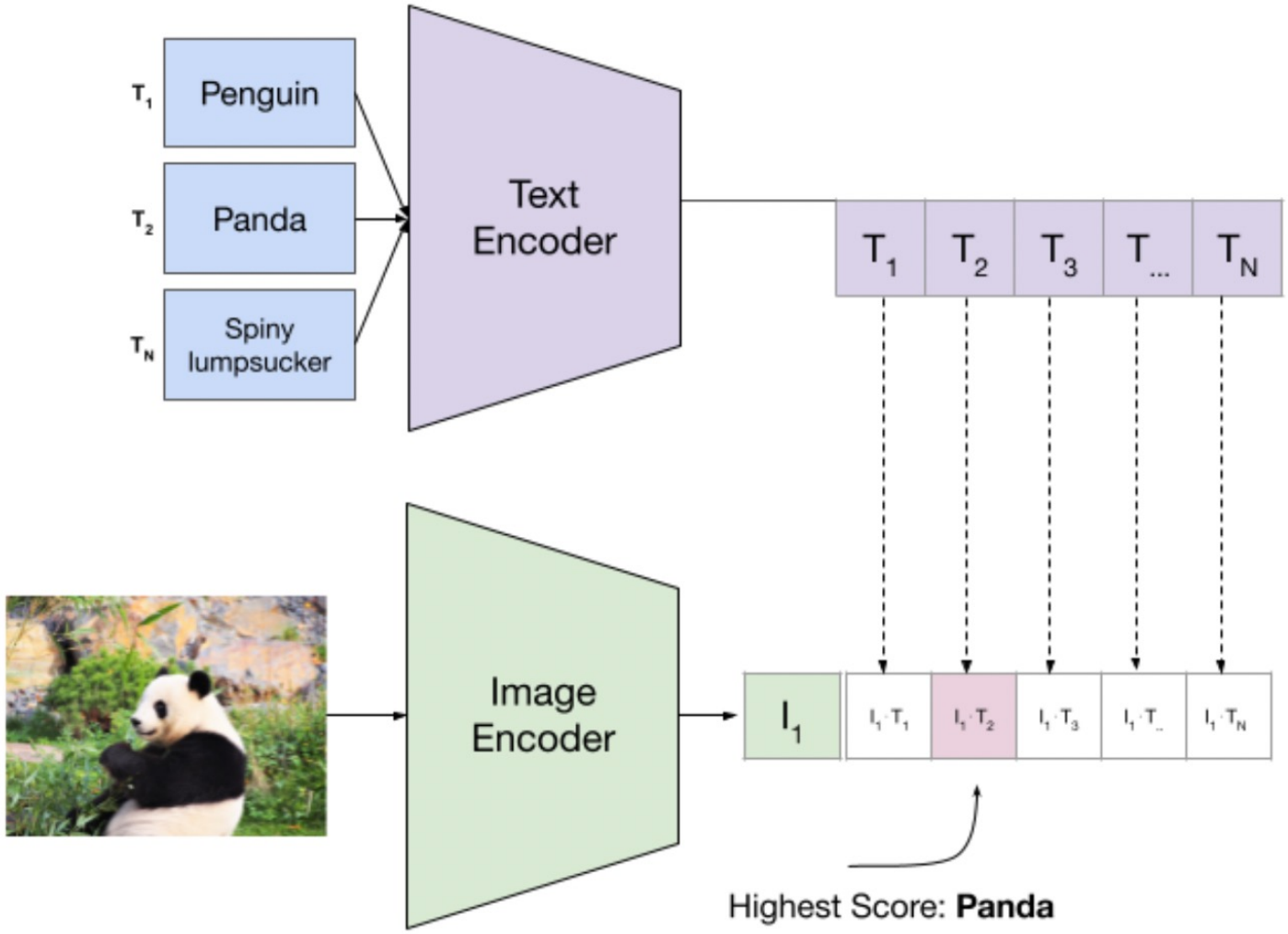
Tested with a text transformer (GPT-2)

Radford et al. Learning Transferable Visual Models From Natural Language Supervision. ICML 2021.

Zero-Shot Performance Evaluated on Over 30 Datasets

Contrastive Language-Image Pretraining (CLIP)

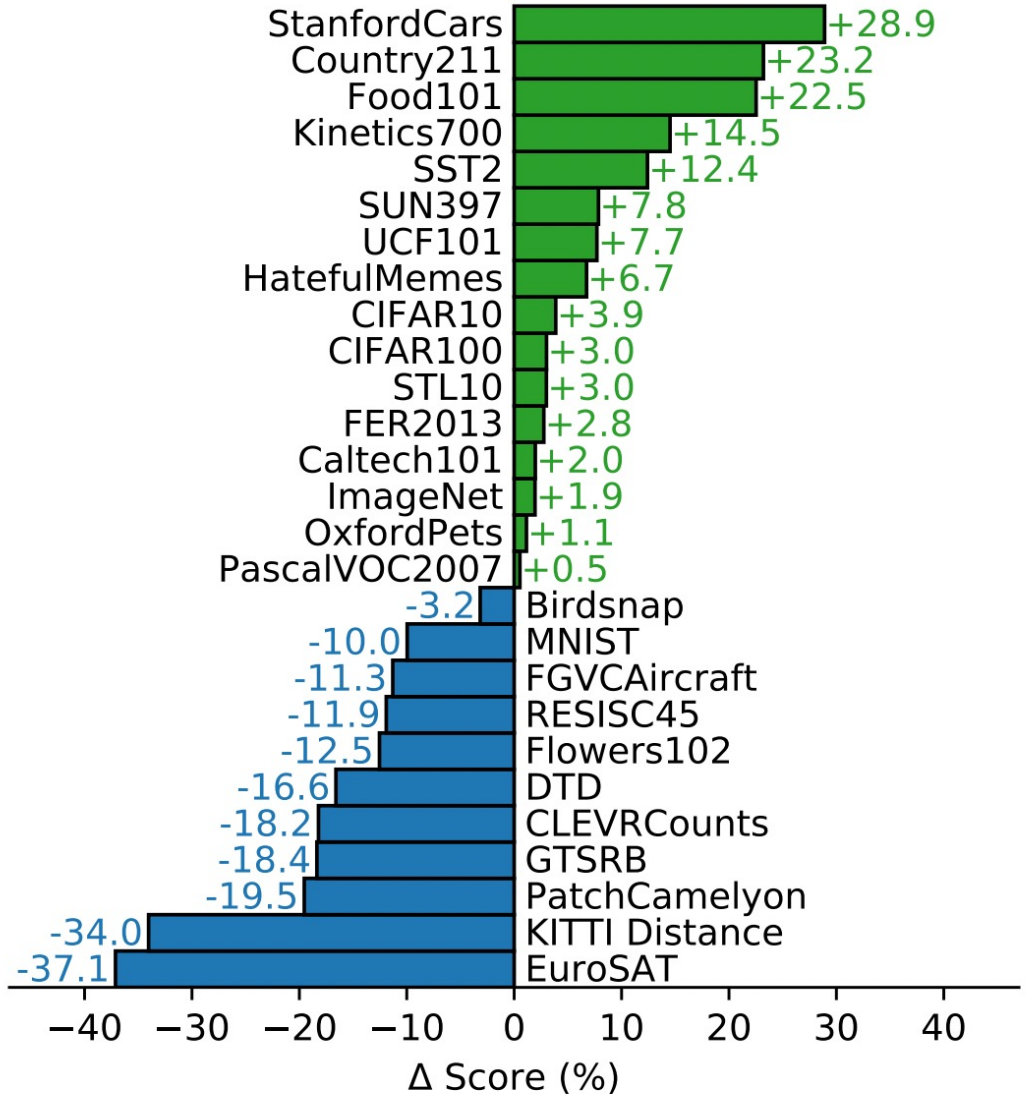
e.g., zero-shot classification: configure representations for all candidate labels (e.g., animal species) using the pretrained encoder and then predict category of image contents based on cosine similarity to category candidates



https://towardsdatascience.com/understanding-zero-shot-learning-making-ml-more-human-4653ac35ccab

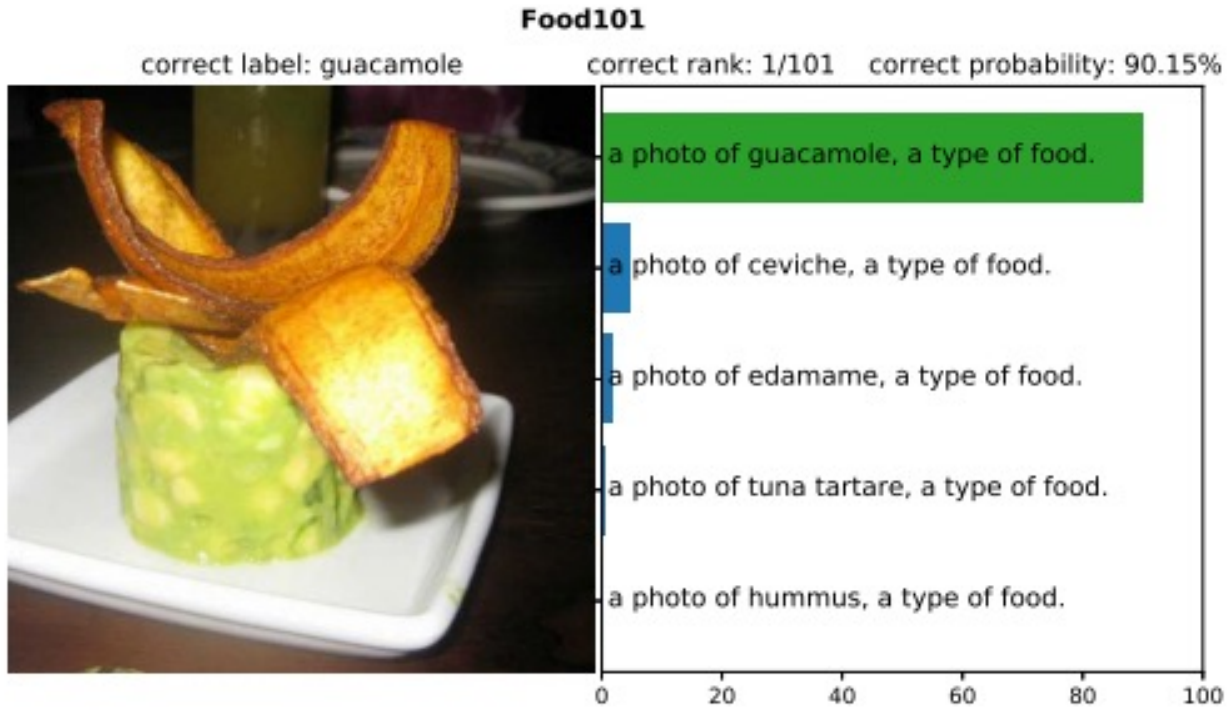
Contrastive Language-Image Pretraining (CLIP)

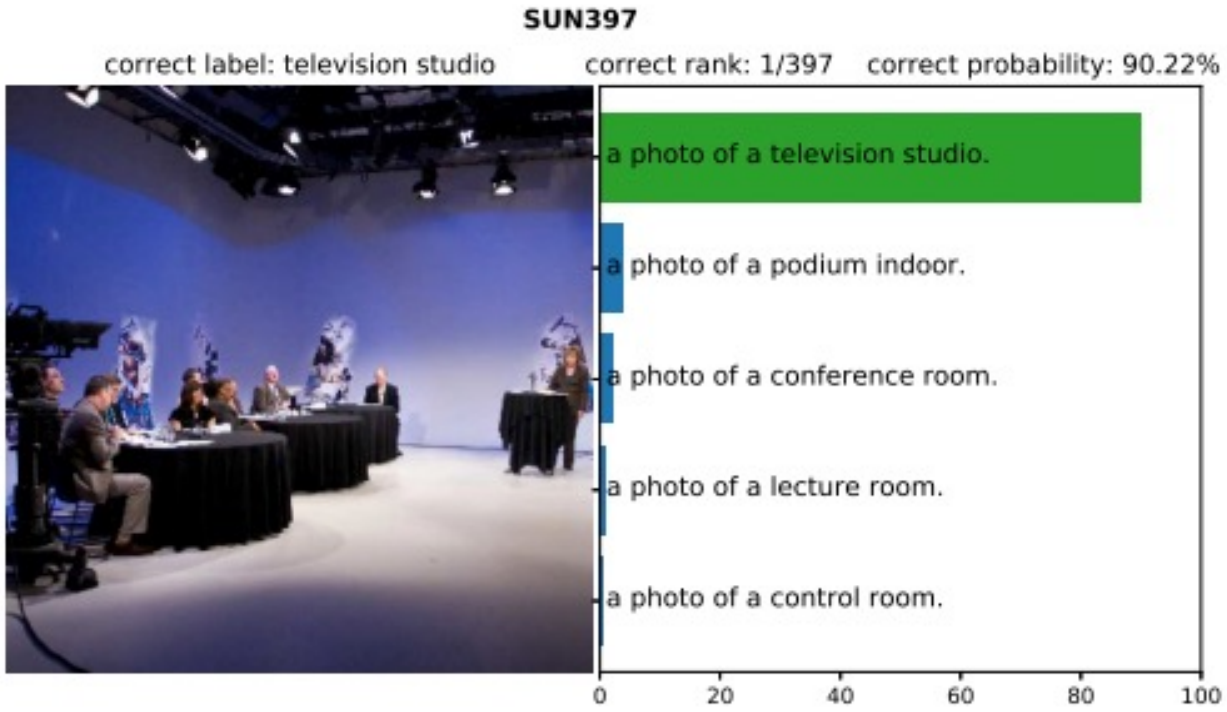
e.g., zero-shot classification:
outperforms a supervised classifier
based on ResNet for 16 of 27 datasets



Radford et al. Learning Transferable Visual Models From Natural Language Supervision. ICML 2021.

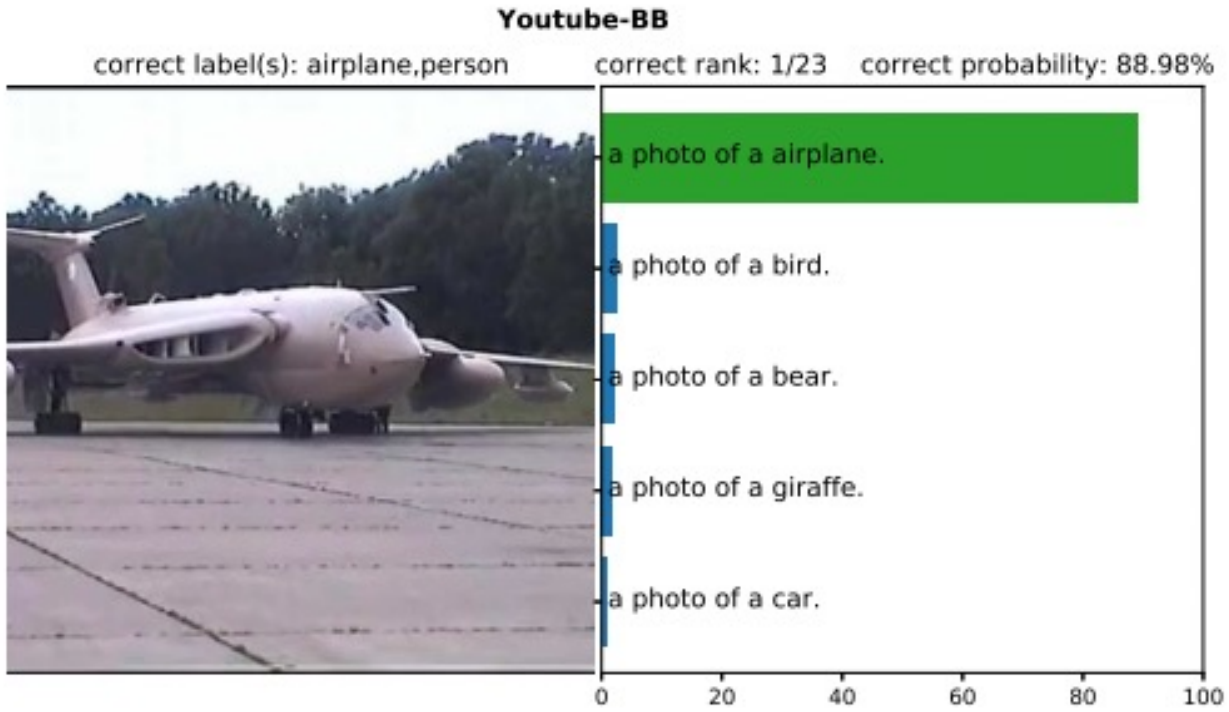
Contrastive Language-Image Pretraining (CLIP): Qualitative Results

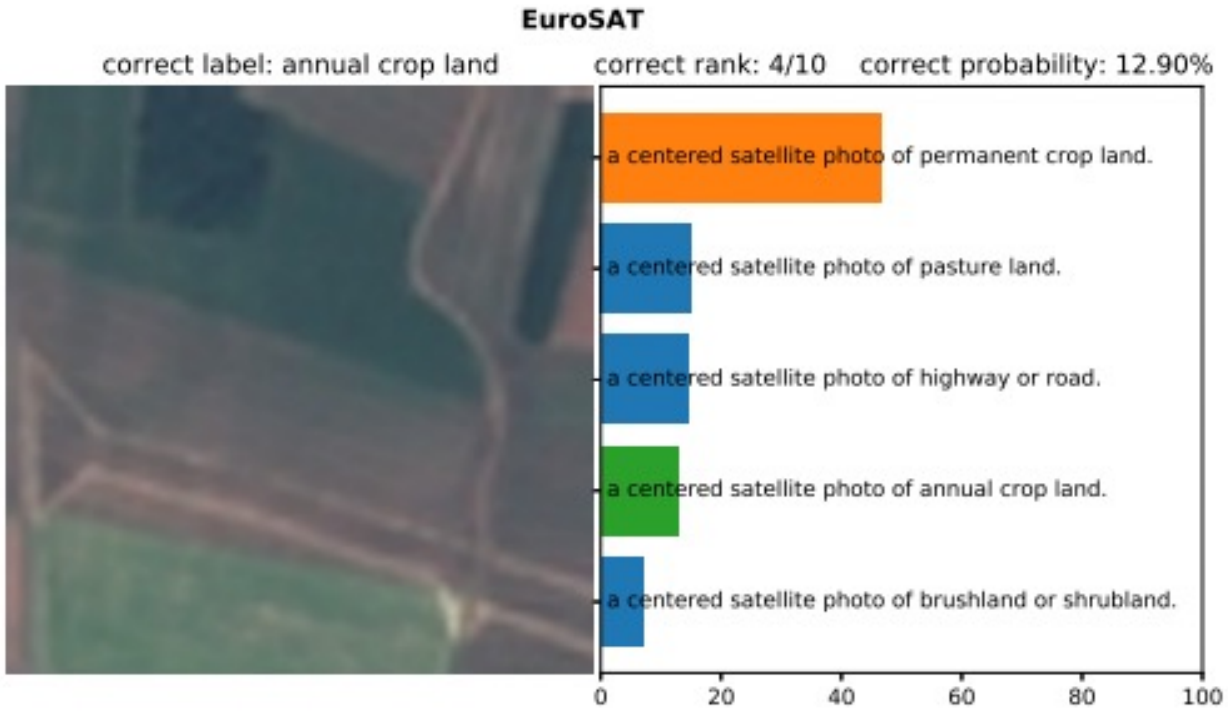




Radford et al. Learning Transferable Visual Models From Natural Language Supervision. ICML 2021.

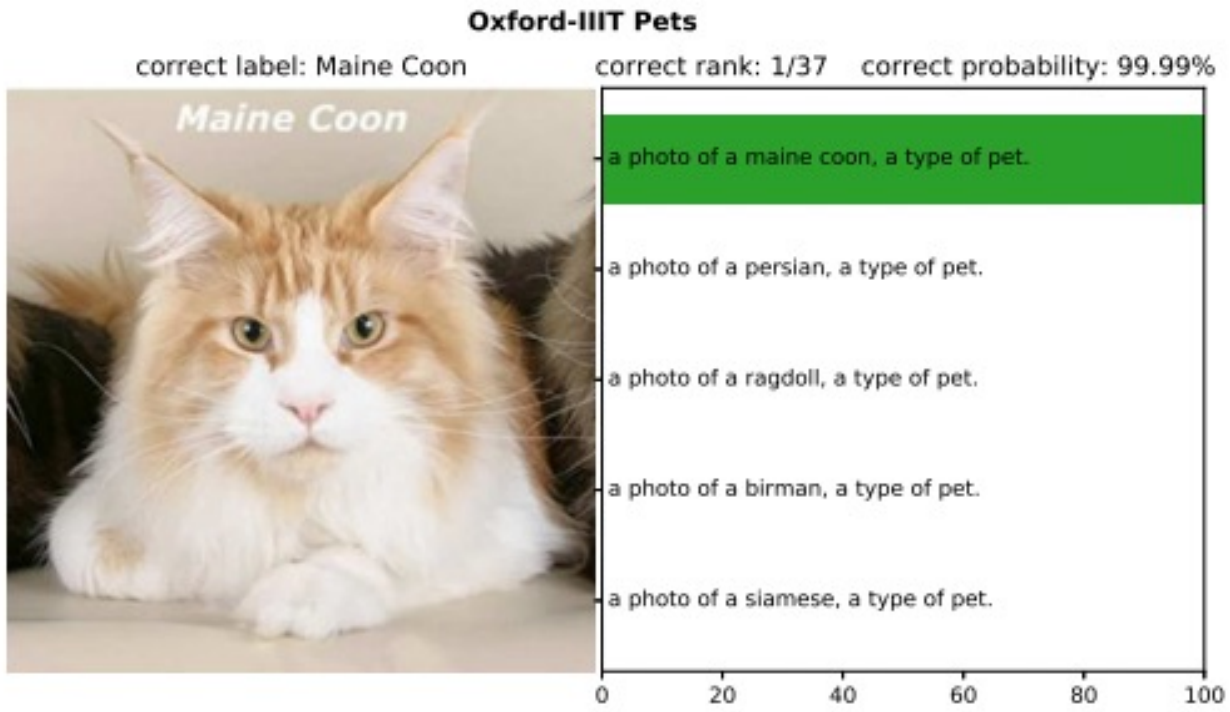
Contrastive Language-Image Pretraining (CLIP): Qualitative Results

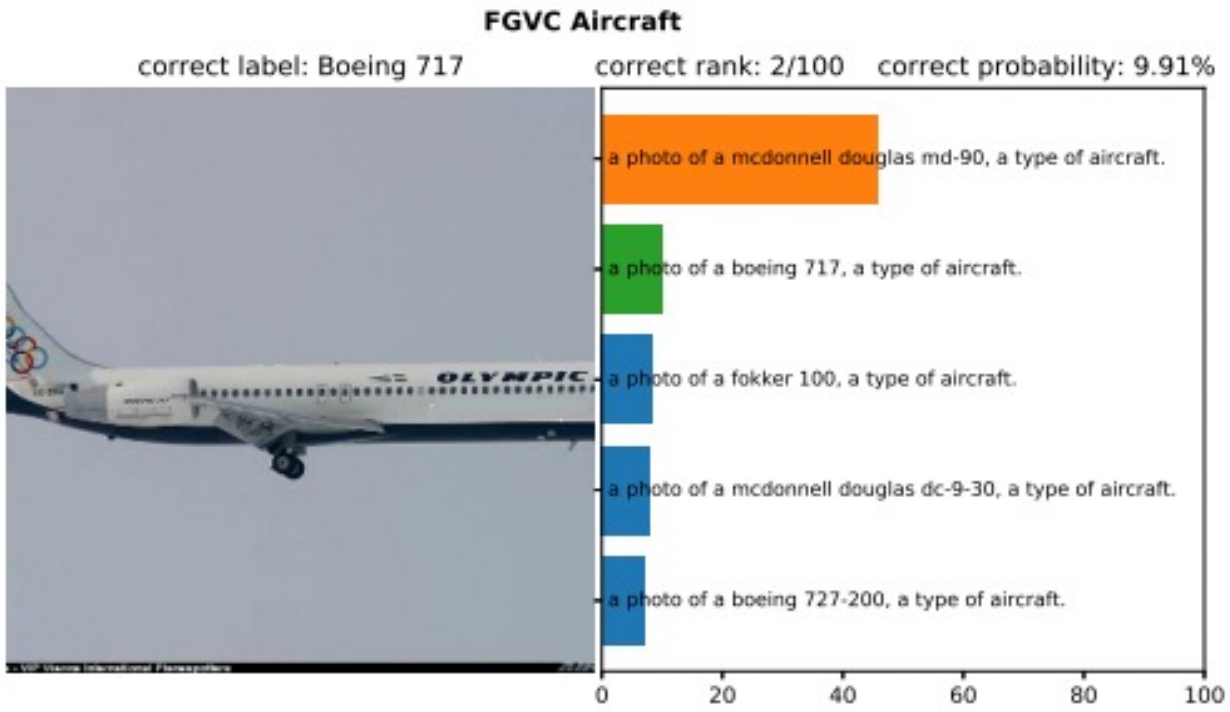




Radford et al. Learning Transferable Visual Models From Natural Language Supervision. ICML 2021.

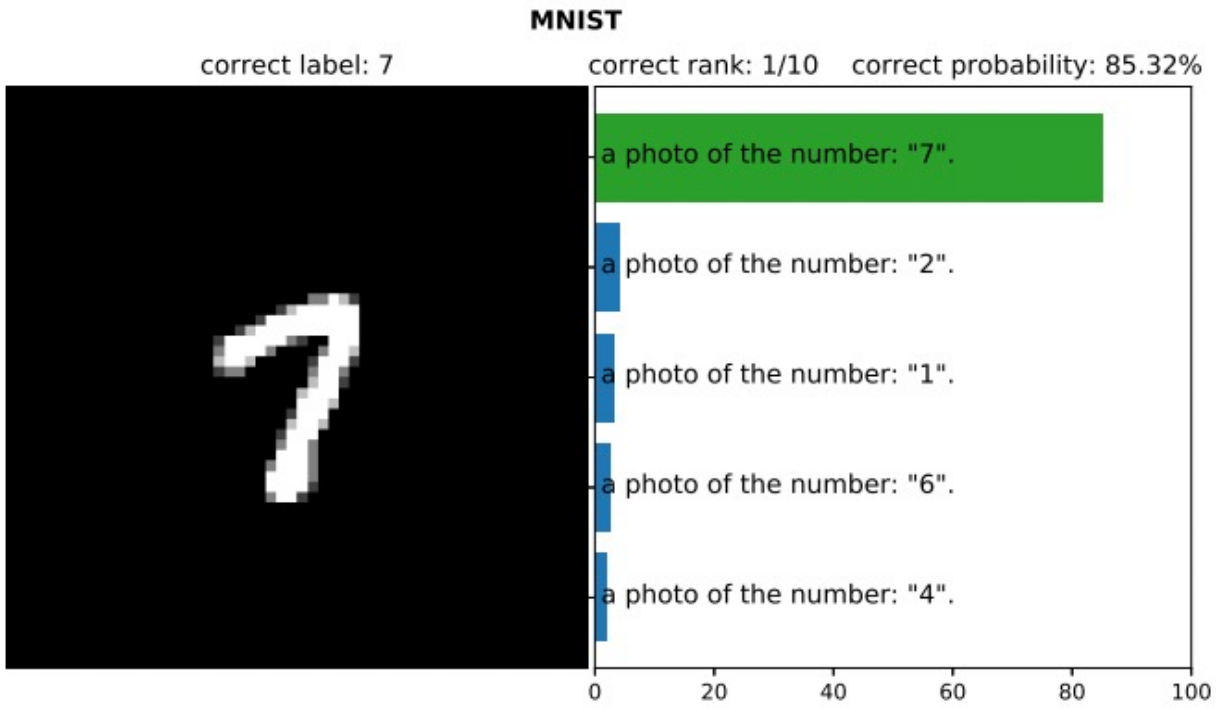
Contrastive Language-Image Pretraining (CLIP): Qualitative Results

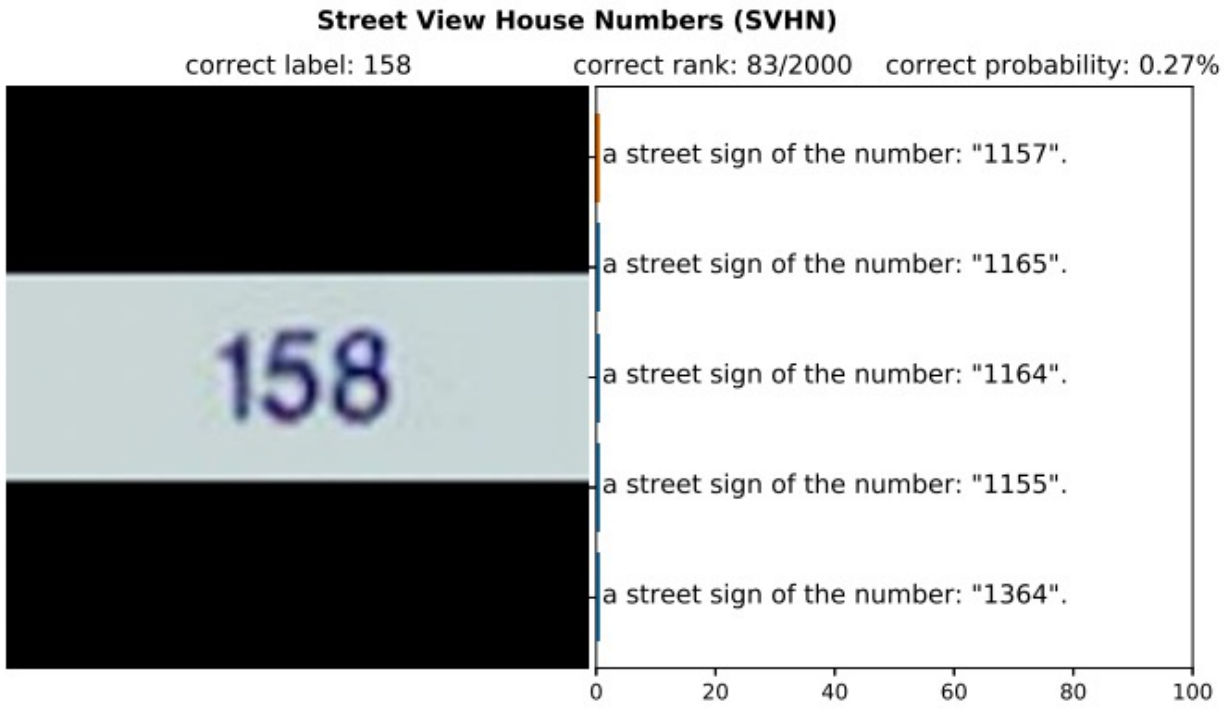




Radford et al. Learning Transferable Visual Models From Natural Language Supervision. ICML 2021.

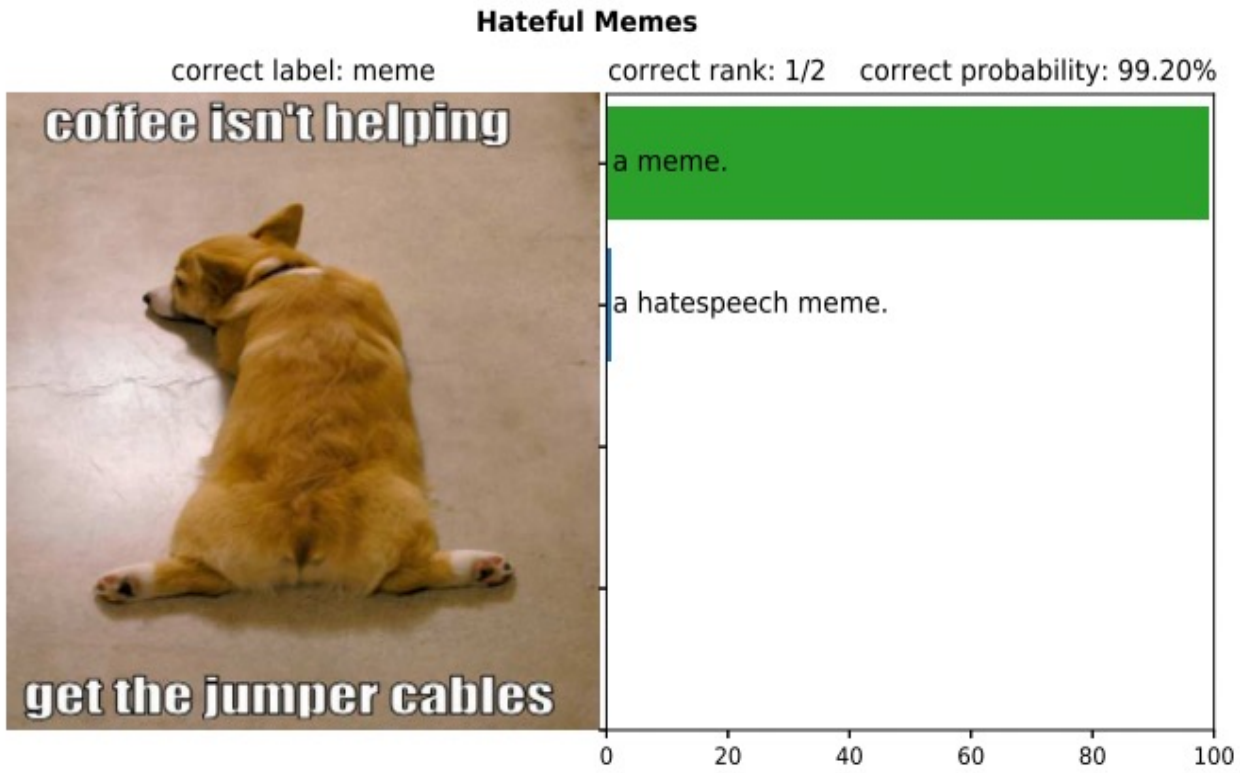
Contrastive Language-Image Pretraining (CLIP): Qualitative Results

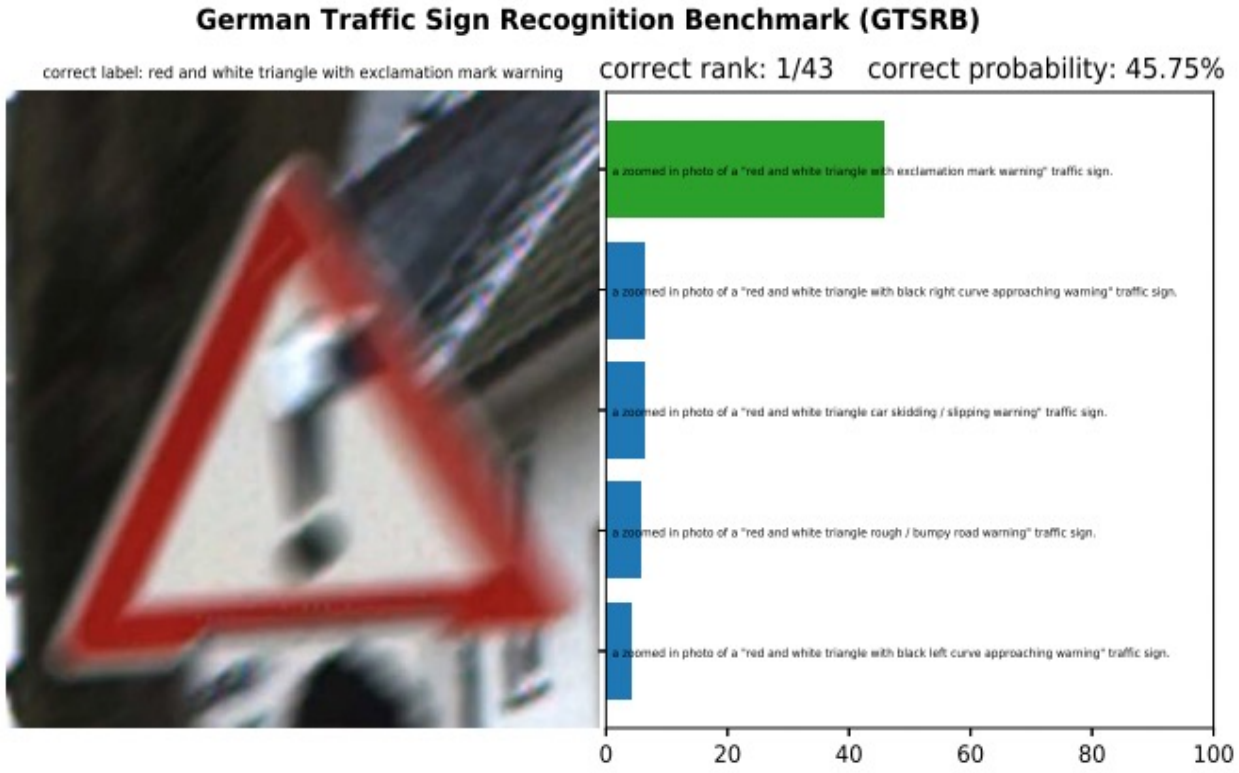




Radford et al. Learning Transferable Visual Models From Natural Language Supervision. ICML 2021.

Contrastive Language-Image Pretraining (CLIP): Qualitative Results





Radford et al. Learning Transferable Visual Models From Natural Language Supervision. ICML 2021.

Popular Approaches

- Text representation: GPT-2 and GPT-3
- Joint image-language representation: CLIP

Today’s Topics

- Multi-task learning
- Few-shot learning
- Zero-shot learning
- Cloud GPU tutorial

Today’s Topics

- Multi-task learning
- Few-shot learning
- Zero-shot learning
- Cloud GPU tutorial


The End