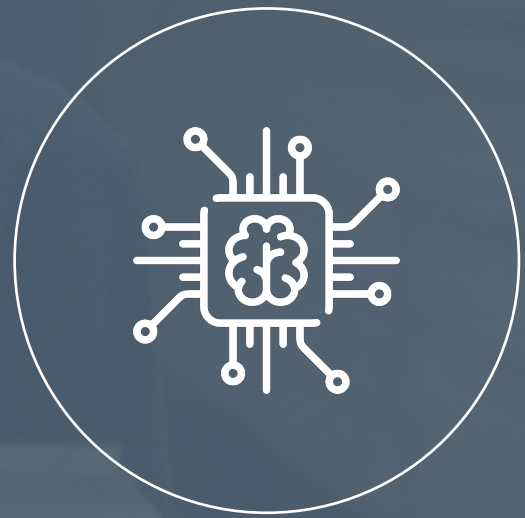
GATO

A Scalable Real-time Architecture for Learning Knowledge from Unsupervised Sensorimotor Interaction



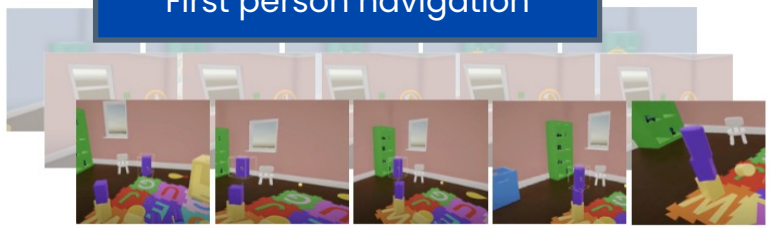
Introduction

Niklas Forsstroem
forsstroemniklas@gmail.com

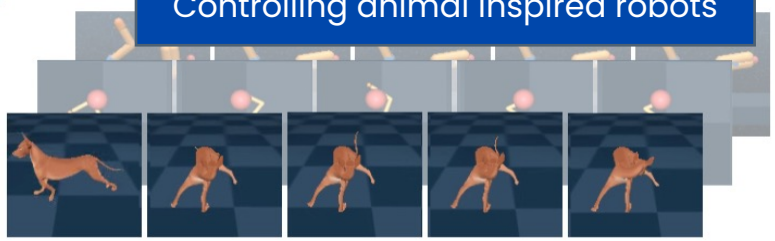
Outperform humans in Atari games

sample efficiency & generalization

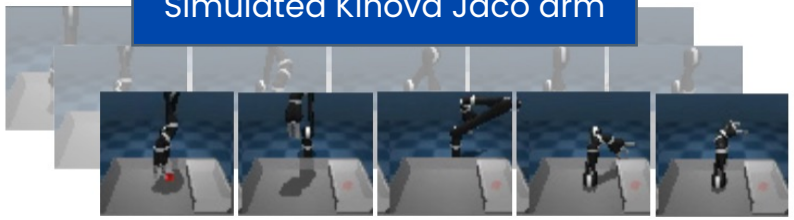
First person navigation



Controlling animal inspired robots



Simulated Kinova Jaco arm



Gato

Transformer sequence model with single set of weights

Maintaining dialogues

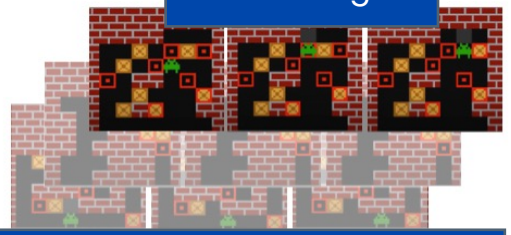
What is the capital of France?

Paris.

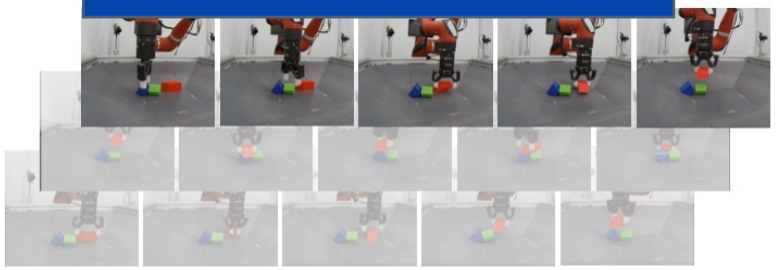
Can you write me a poem?

I don't know exactly what to write. There's just so much to answer.

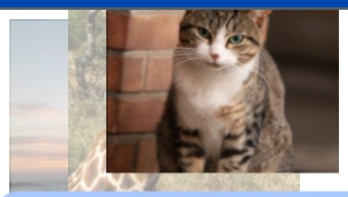
Planning



Stack blocks with real robot arm

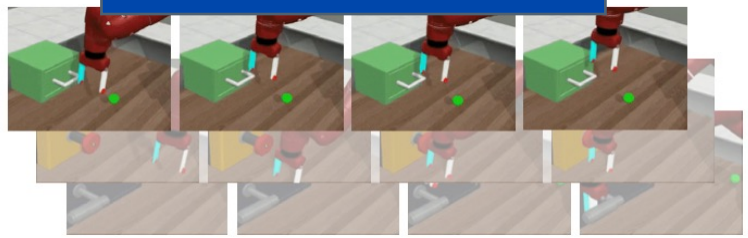


Generate image captions



A cat that is sitting next to a brick wall and green grass some bushes as the sun sets

Multi task learning



Introduction

Using single neural sequence model across tasks has many benefits

- No need to construct different models for each individual domain
- Lowers requirement to tailor training for specific phenomenon
- Performance of large models does not stagnate as easily – and may be increased as compute and dataset sizes increase
- Generic models tend to outperform more specialized domain-specific approaches over time

Tests whether training an agent which is generally capable on a large number of tasks is possible

- Test if this general agent can be adapted with little extra data to succeed at an even larger number of tasks.

For simplicity, Gato was trained offline in a purely supervised manner

- Can in principle also be trained with either offline or online reinforcement learning (RL).

Model

General approach

Design principle is to train on widest variety of relevant data:

- Images
- Text
- Proprioception
- Joint torques
- Button presses
- Other discrete and continuous observations and actions.

Network architecture has two main components:

- Parameterized embedding function which transforms tokens to token embeddings
- Sequence model which outputs a distribution over the next discrete token.

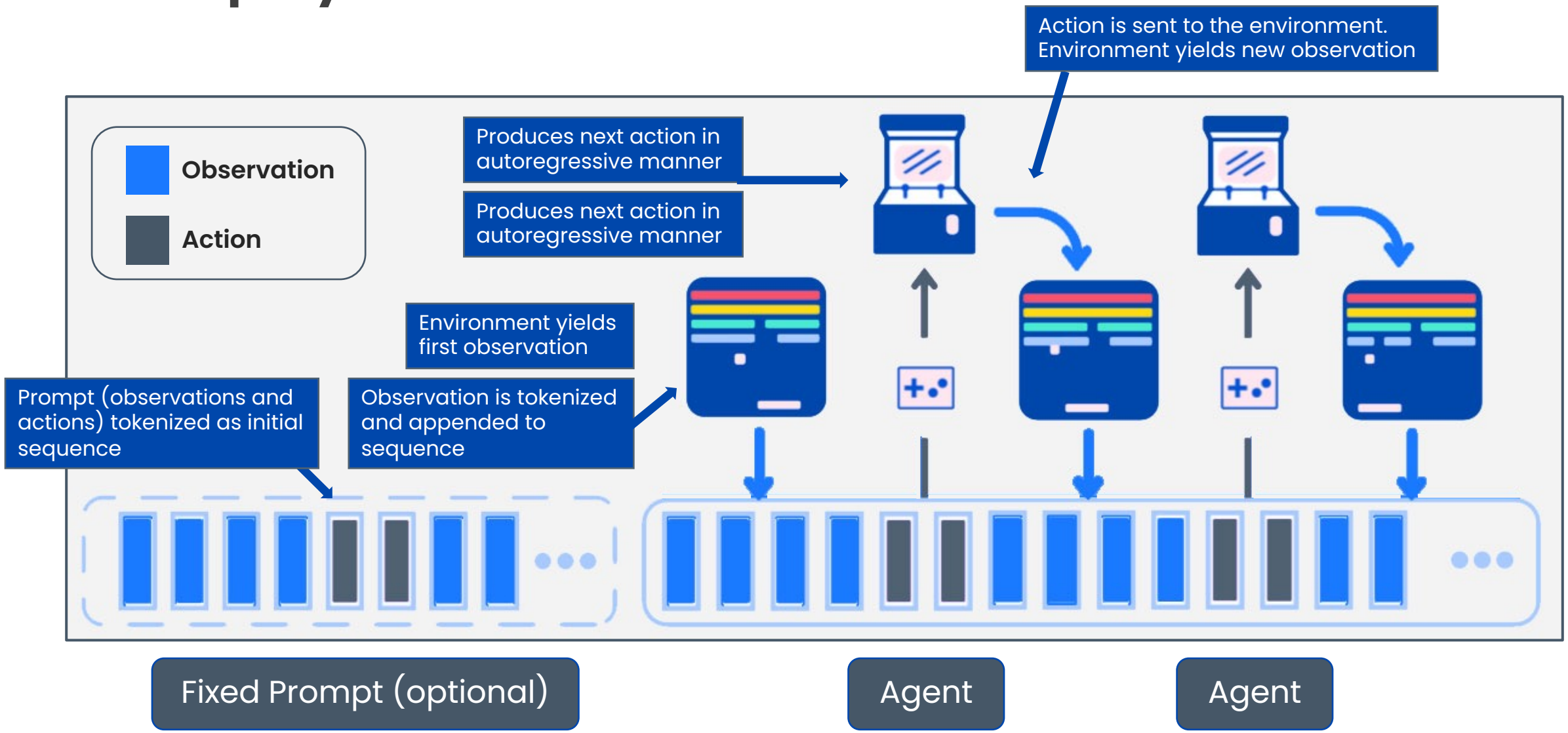
Use transformer sequence model for simplicity and scalability

- 1.2B parameter decoder-only transformer with 24 layers, an embedding size of 2048, and a post-attention feedforward hidden size of 8196.

To enable processing this multi-modal data, it is serialized into a flat sequence of tokens

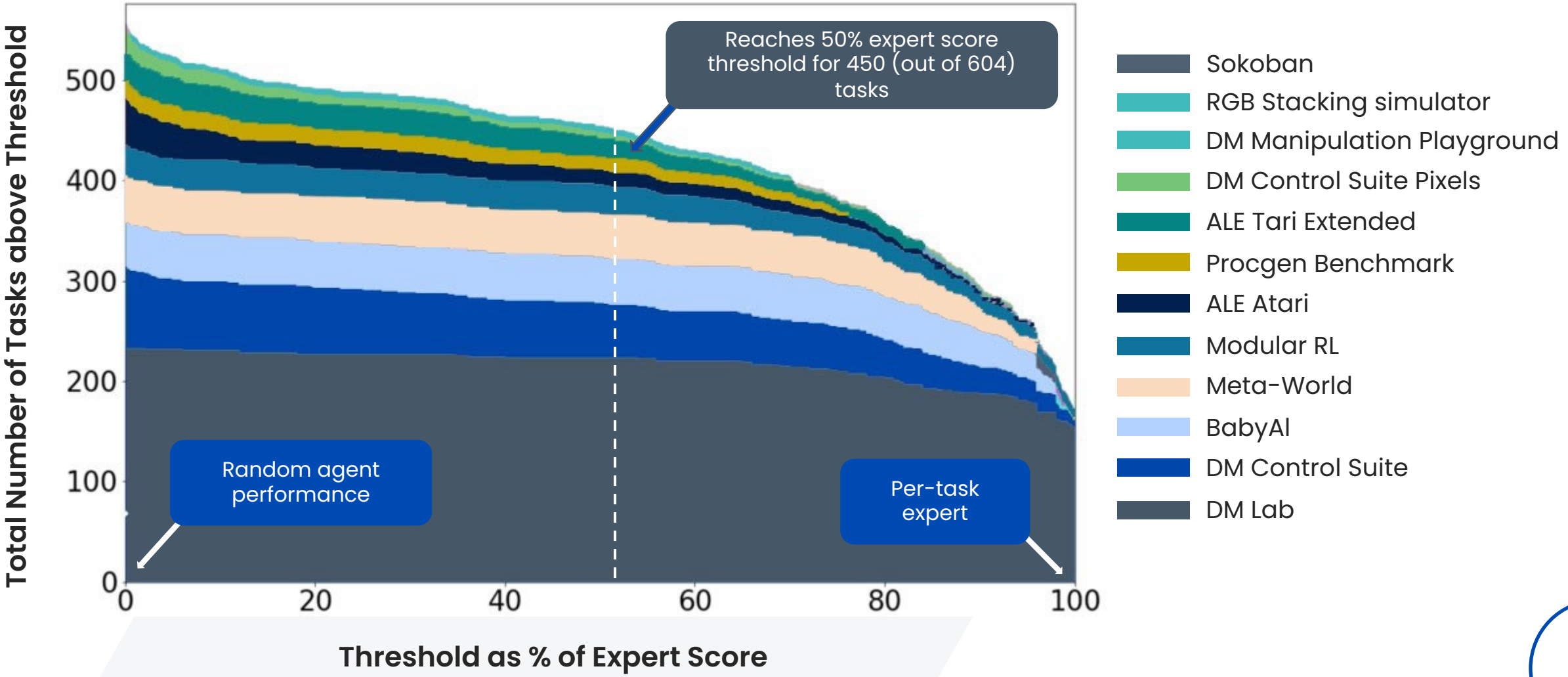
- Embeddings depend on the type of input
- Final sequence contains all embeddings with specific intrinsic order

2.4 Deployment



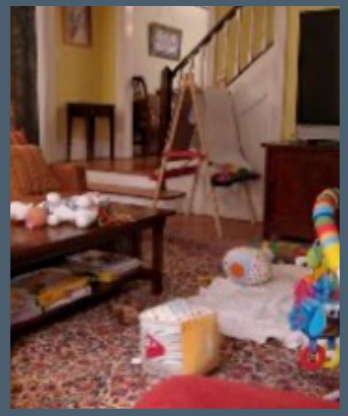
4.1 Simulated control tasks

Number of tasks completed as fun. of quality



4.3 Text samples: Image captions from Gato

Representative sample of Gato's image captioning performance.
Sampled without cherry-picking



The colorful ceramic toys are on the living room floor.

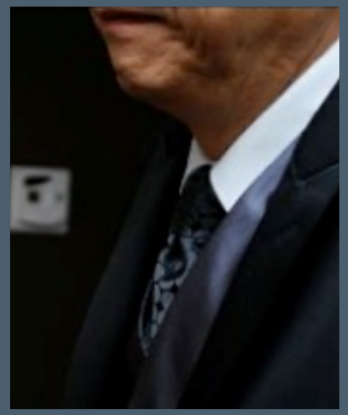
A living room with three different color deposits on the floor.

A room with a long red rug a tv and some pictures.

A bearded man is holding a plate of food.

Man holding up a banana to take a picture of it.

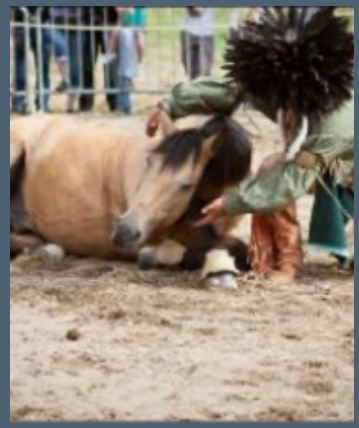
A man smiles while holding up a slice of cake.



Man standing in the street earing a suit and tie.

A man in a blue suit with a white bow tie and black shoes.

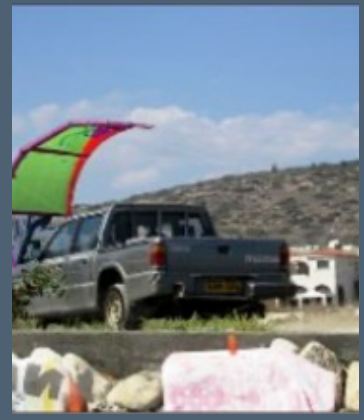
A man with a hat in his hand looking at the camera



A group of people that is next to a big horse.

A tan horse holding a piece of cloth lying on the ground.

Two horses are laying on their side of the dirt.



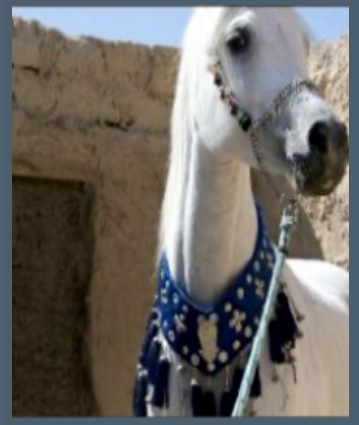
Man biting a kite while standing on a construction site.

A big truck in the middle of the road.

A truck with a kite painted on the back is parked by rocks.

4.3 Text samples: Image captions from Gato

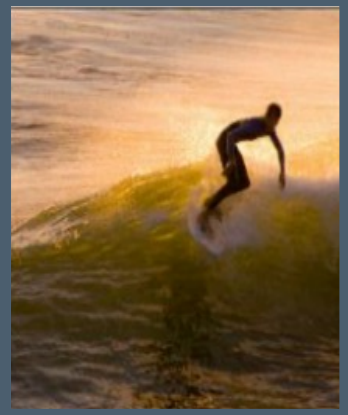
Representative sample of Gato's image captioning performance.
Sampled without cherry-picking



A white horse with a blue and silver bridle

A white horse with blue and gold chains.

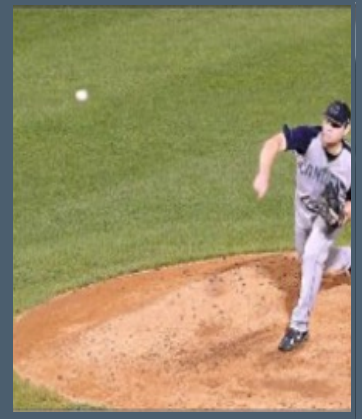
A horse is being shown behind a wall.



a couple of people are out in the ocean

A surfer riding a wave in the ocean.

A surfer with a wet suit riding a wave.



A baseball player pitching a ball on top of a baseball field.

A man throwing a baseball at a pitcher on baseball field.

A baseball player at bat and catcher in the dirt during a baseball game.

Pistachios on top of a bowl with coffee on the side.

A bowl and a glass of liquid sits on a table.

A white plate filled with a banana bread next to a cup of coffee.

A group of children eating pizza at the table.

Two boys having pizza for lunch with their friends.

The boys are eating pizza together at the table.

Fine-tuning on Robotic Stacking Tasks

Adaptation to Perceptual Variations

Evaluated agent's adaptability to perceptual variations and permutations in the objective specification.

Adding simulated demonstrations of the stack blue on green task to the fine-tuning dataset improved performance

10% was an ideal sampling ratio for this data.

Trained agent (physical robot) to stack red objects onto blue ones

- All simulated and real robotics data in pretraining set stacks red object on blue, and does not include the test set shapes
- Manually collected 500 demonstrations of "stack blue on green" with a 3D mouse for fine-tuning

60% success rate after evaluating fine-tuned Gato on the real robot

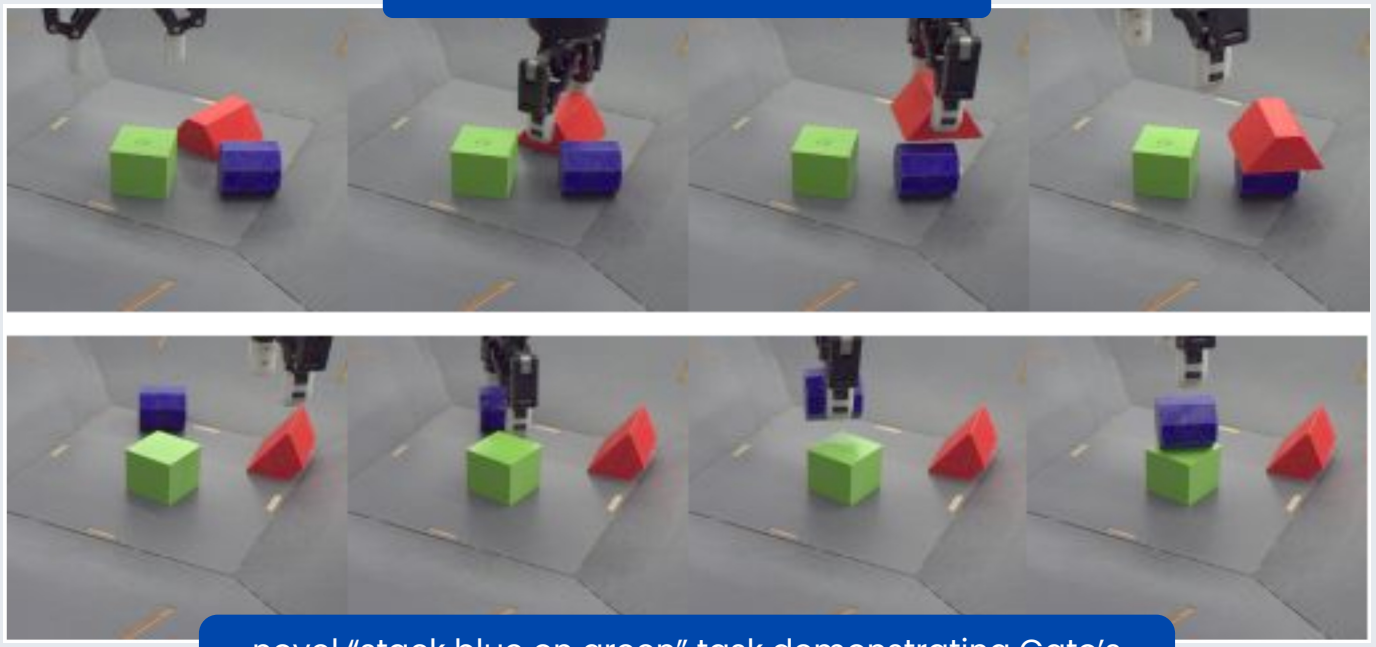
- Baseline trained on blue-on-green data achieved only 0.5% success
- Baseline would consistently move towards the blue object
  and occasionally pick it up and place it on top of the green object
- A full stable stack was almost never achieved.

Fine-tuning on Robotic Stacking Tasks

Adaptation to Perceptual Variations

Test triplet

standard “stack red on blue” task



novel “stack blue on green” task demonstrating Gato’s out of distribution adaptation to perceptual variations.

Scaling law analysis

Model size scaling laws results

- How performance changes with increased model capacity.
- To get a single mean normalized score:
  - ✓ Model performance (% of expert score) is evaluated for each task in all domains
  - ✓ Percentage scores across the tasks a given domain are averaged.
  - ✓ Percentage scores across all domains are mean-aggregated
- For three model sizes, normalized return is plotted as training progresses
- For equivalent token count, there is a significant performance improvement with increased scale.