Active One-shot Learning

Mark Woodward[1] and Chelsea Finn[2]

[1]Independent Researcher; [2]University of California, Berkeley

Introduction

Goal: Online active-learning from few examples

Approach:

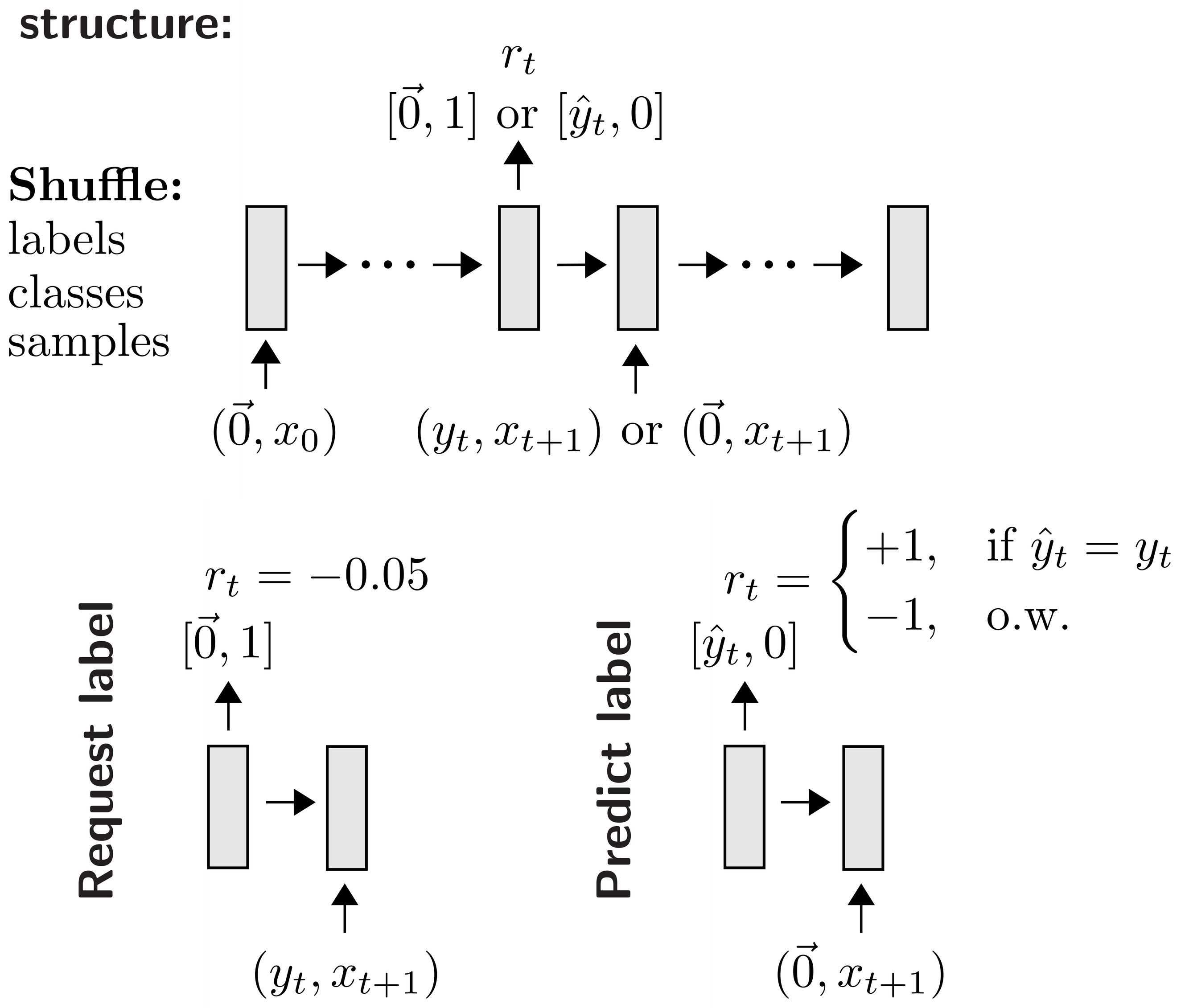
- Frame as an RL problem
- Train on a modified one-shot learning task

Key Insights:

- Train on short randomized episodes
- Train on a dataset with a large number of classes

Task Methodology

Task structure:



Rewards:

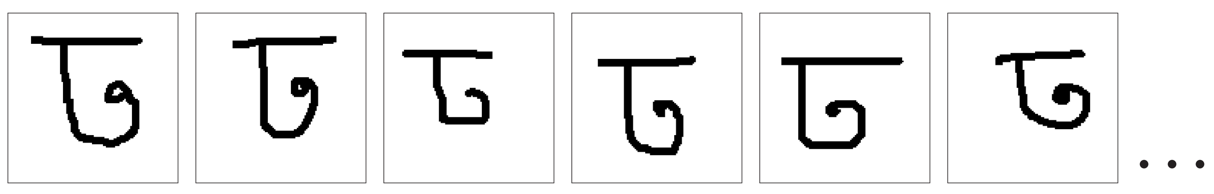
$$r_t = \begin{cases} R_{req}, & \text{if a label is requested} \\ R_{cor}, & \text{if predicting and } \hat{y}_t = y_t \\ R_{inc}, & \text{if predicting and } \hat{y}_t \neq y_t \end{cases}$$

Loss:

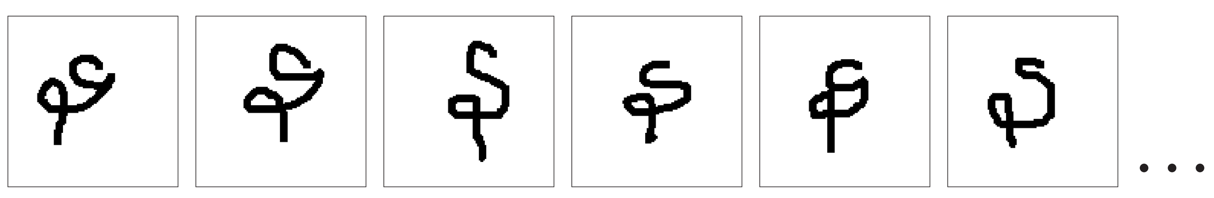
$$\mathcal{L}(\Theta) := \sum_t [Q_\Theta(o_t, a_t) - (r_t + \gamma \max_{a_{t+1}} Q_\Theta(o_{t+1}, a_{t+1}))]^2$$

Omniglot Dataset

Character 0156:



Character 0790:



- 20 hand drawn images for each character (1,623 characters)
- Lake et al., Human-level concept learning through probabilistic program induction. *Science*, 2015

Experimental Setup

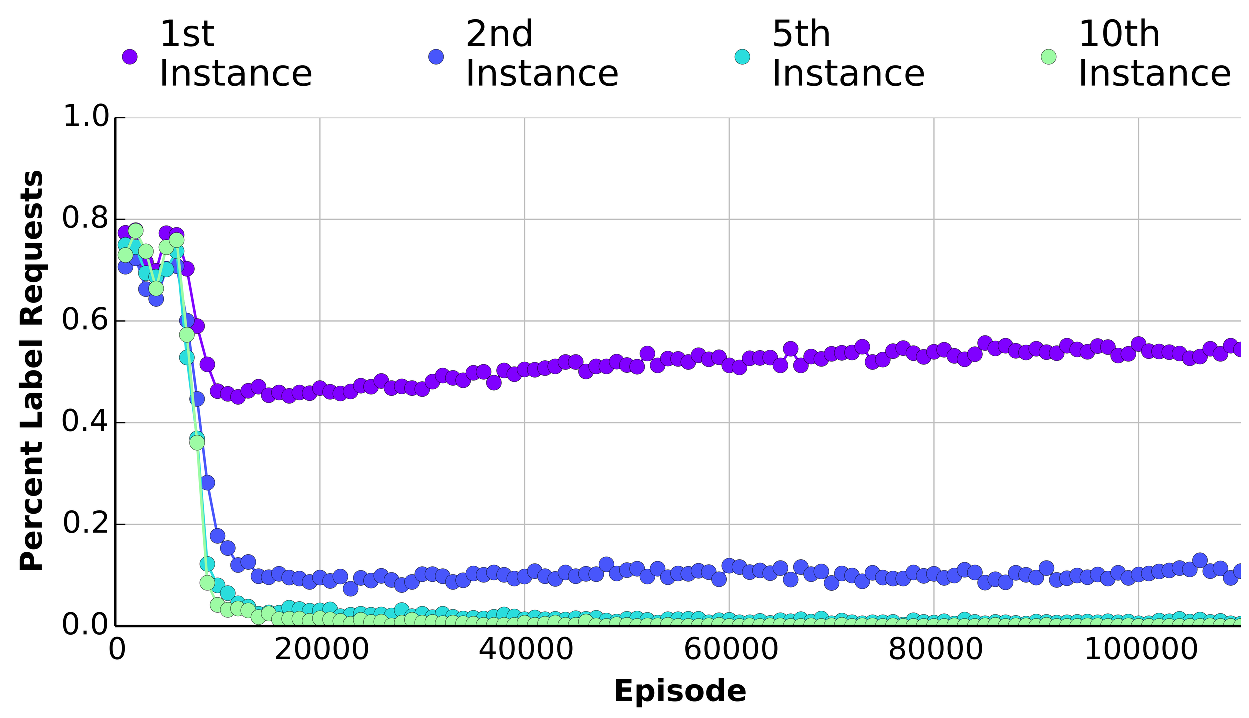
- $Q(o_t, a_t)$ is a 200 unit, single layer LSTM
- Q-learning of $Q(o_t, a_t)$
- ϵ-greedy exploration (0.05)
- 30 images per episode
- 3 classes per episode
- 50 episodes per training batch

Conclusions

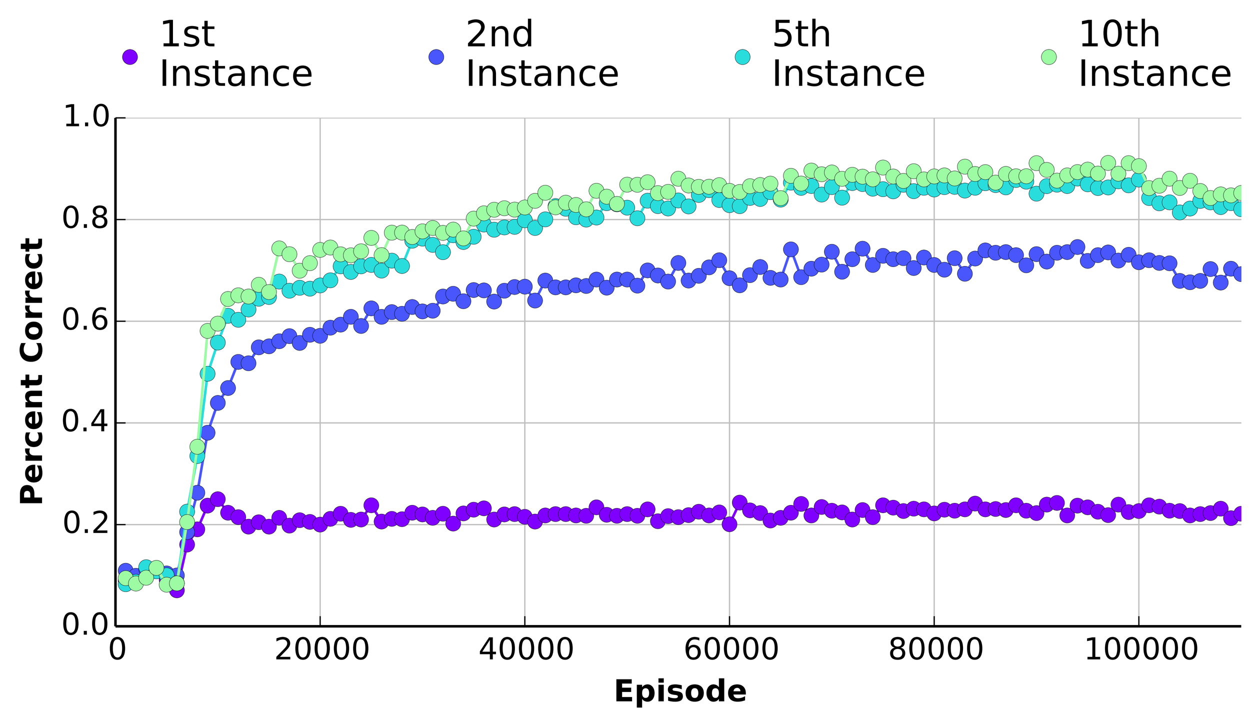
- Online active one-shot learning is possible
- The choice of rewards can trade off accuracy for requests

Results: Learning to Request Labels

Requests:



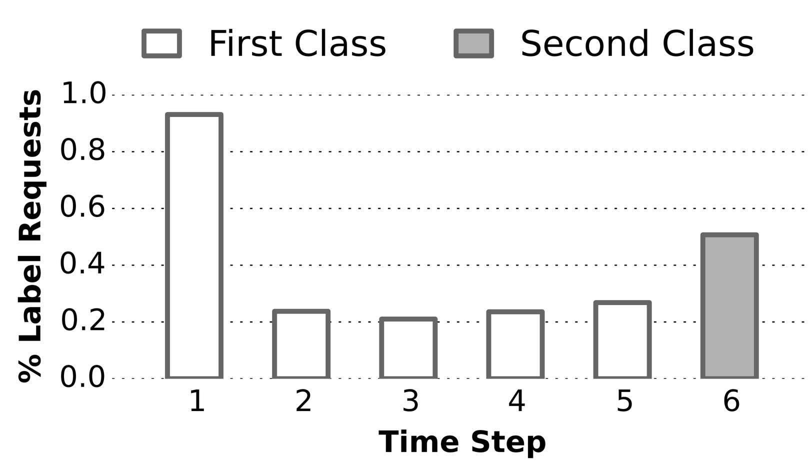
Accuracy:



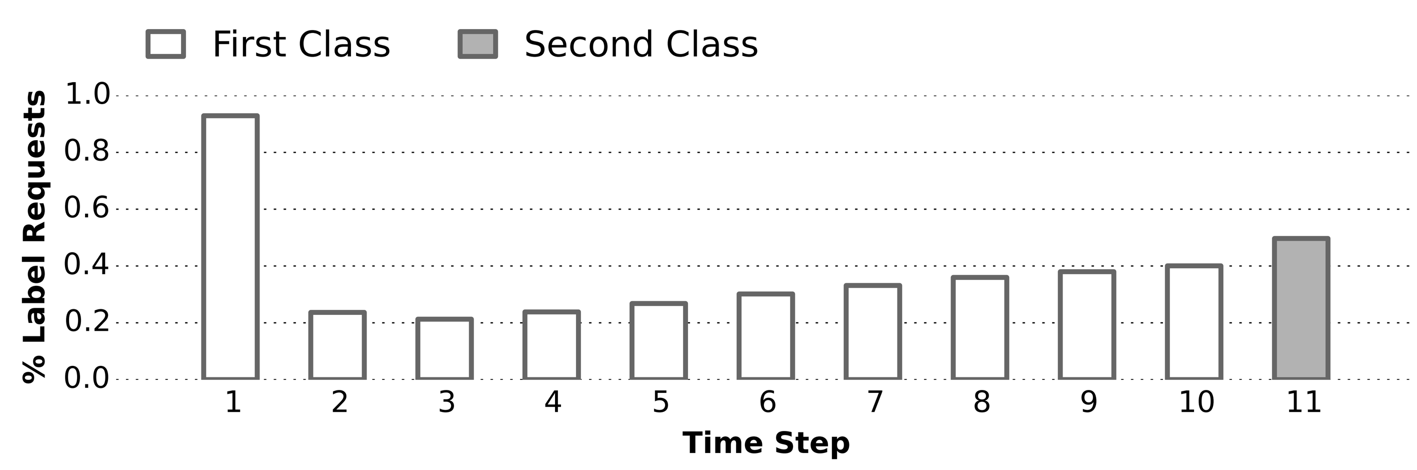
- Fewer requests and higher accuracy on later instances of a class
- At 100,000 episodes, training stops and data switches to test set

Results: Considering Uncertainty

Switch classes on step 6:



Switch classes on step 11:



- Note: This is a different task from the rest of the paper
- Shows that label requests are based on comparisons (difference in step 6's)
- Disproves that a simple policy was learned

Results: Trading Accuracy for Requests

	Accuracy (%)	Requests (%)
Supervised	91.0	100.0
RL	75.9	7.2
RL Prediction	81.8	**7.2**
RL Prediction ($R_{inc} = \mathbf{-5}$)	86.4	31.8
RL Prediction ($R_{inc} = \mathbf{-10}$)	89.3	45.6
RL Prediction ($R_{inc} = \mathbf{-20}$)	**92.8**	60.6

- % of steps that are correct and % of steps where requests are made
- Increasing the penalty for an incorrect label increases accuracy at the cost of more label requests
- "Supervised" is task from Santoro et al., One-shot Learning with MANNs, ICML 2016