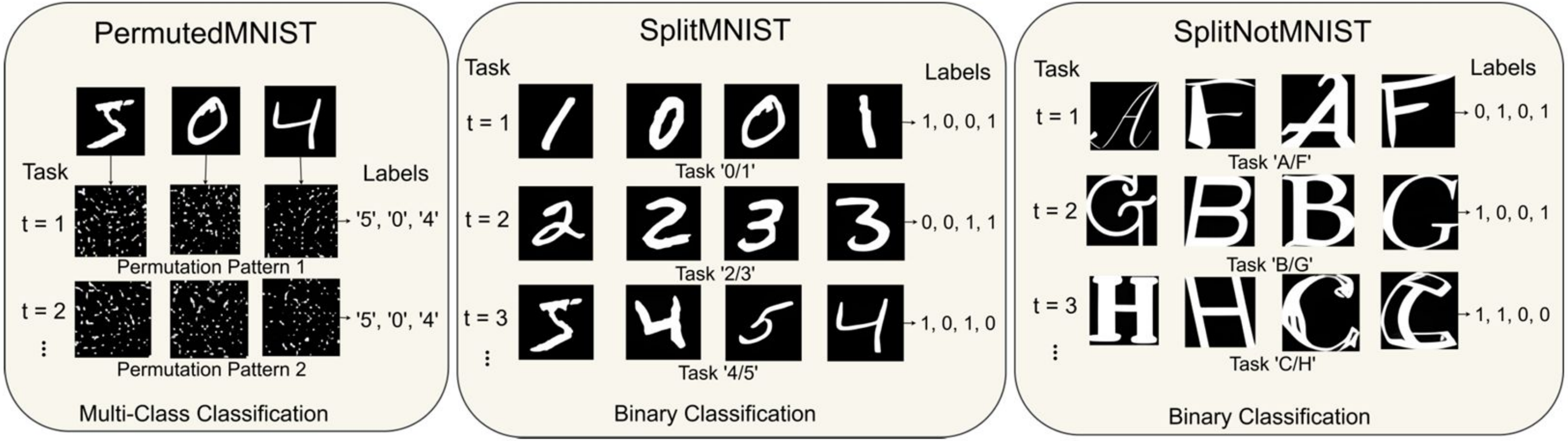
Temporal-Difference Variational Continual Learning

Luckeciano Melo, Alessandro Abate, Yarin Gal

University of Oxford

Problem Setting

- (Supervised) Continual Learning
  - Assume a task distribution $\mathcal{T}$
  - Each task $t \sim \mathcal{T}$ is represented by a dataset of pairs $\{(\boldsymbol{x}_t, y_t)\}^{N_t}$
  - At each timestep t, the model receives a training stream $\mathcal{D}_t$
  - Evaluation considers held-out sets from all previously observed tasks



Variational Continual Learning

- Bayesian Framework
  - Assume a prior over parameters $p(\boldsymbol{\theta})$
  - Given evidence of T tasks, learn posterior $p(\boldsymbol{\theta} \mid \mathcal{D}_{1:T})$
  - Assuming tasks are i.i.d, we find the following recursion:

$$p(\boldsymbol{\theta} \mid \mathcal{D}_{1:T}) \propto p(\boldsymbol{\theta})p(\mathcal{D}_{1:T} \mid \boldsymbol{\theta}) \overset{\text{i.i.d}}{=} p(\boldsymbol{\theta}) \prod_{t=1}^{T} p(\mathcal{D}_t \mid \boldsymbol{\theta}) \propto p(\boldsymbol{\theta} \mid \mathcal{D}_{1:T-1})p(\mathcal{D}_T \mid \boldsymbol{\theta})$$

- Variational CL
  - Approximate via Variational Inference

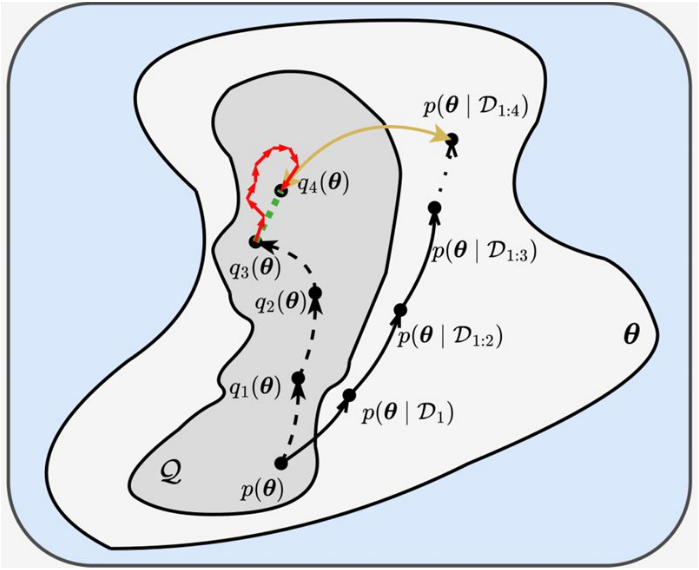
$$q_t(\boldsymbol{\theta}) = \arg\min_{q \in \mathcal{Q}} \mathscr{D}_{KL}(q(\boldsymbol{\theta}) \mid\mid \frac{1}{Z_t} q_{t-1}(\boldsymbol{\theta}) p(\mathcal{D}_t \mid \boldsymbol{\theta}))$$

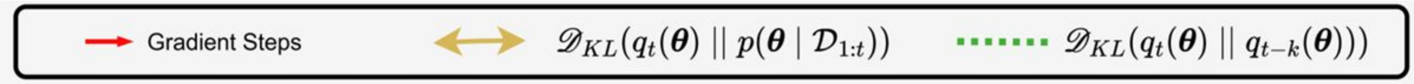
  - This is equivalent to maximizing the ELBO:

$$\mathcal{L}^t_{VCL}(\boldsymbol{\theta}) = \mathbb{E}_{\boldsymbol{\theta} \sim q_t(\boldsymbol{\theta})}[\log p(\mathcal{D}_t \mid \boldsymbol{\theta})] - \mathscr{D}_{KL}(q_t(\boldsymbol{\theta}) \mid\mid q_{t-1}(\boldsymbol{\theta}))$$

Problem

$$\mathcal{L}^t_{VCL}(\boldsymbol{\theta}) = \mathbb{E}_{\boldsymbol{\theta}\sim q_t(\boldsymbol{\theta})}[\log p(\mathcal{D}_t \mid \boldsymbol{\theta})] - \mathscr{D}_{KL}(q_t(\boldsymbol{\theta}) \,||\, q_{t-1}(\boldsymbol{\theta}))$$





- Several approximation errors may bias the posterior
  - Likelihood term estimation
  - Optimization errors
  - Function class may not be expressive enough to represent the true posterior

- As it relies on a single previous posterior, errors **compound** over successive updates

Problem

$$q_t(\boldsymbol{\theta}) = \arg\min_{q \in \mathcal{Q}} \mathscr{D}_{KL}(q(\boldsymbol{\theta}) \,||\, \frac{1}{Z_t} q_{t-1}(\boldsymbol{\theta}) p(\mathcal{D}_t \mid \boldsymbol{\theta}))$$

Proposition 4.1. *The standard KL minimization objective in Variational Continual Learning (Equation 2) is equivalently represented as the following objective, where $n \in \mathbb{N}_0$ is a hyperparameter:*

$$q_t(\boldsymbol{\theta}) = \arg\max_{q \in \mathcal{Q}} \mathbb{E}_{\boldsymbol{\theta} \sim q_t(\boldsymbol{\theta})} \Big[\sum_{i=0}^{n-1} \frac{(n-i)}{n} \log p(\mathcal{D}_{t-i} \mid \boldsymbol{\theta}) \Big] - \sum_{i=0}^{n-1} \frac{1}{n} \mathscr{D}_{KL}(q_t(\boldsymbol{\theta}) \,||\, q_{t-i-1}(\boldsymbol{\theta})). \quad (4)$$

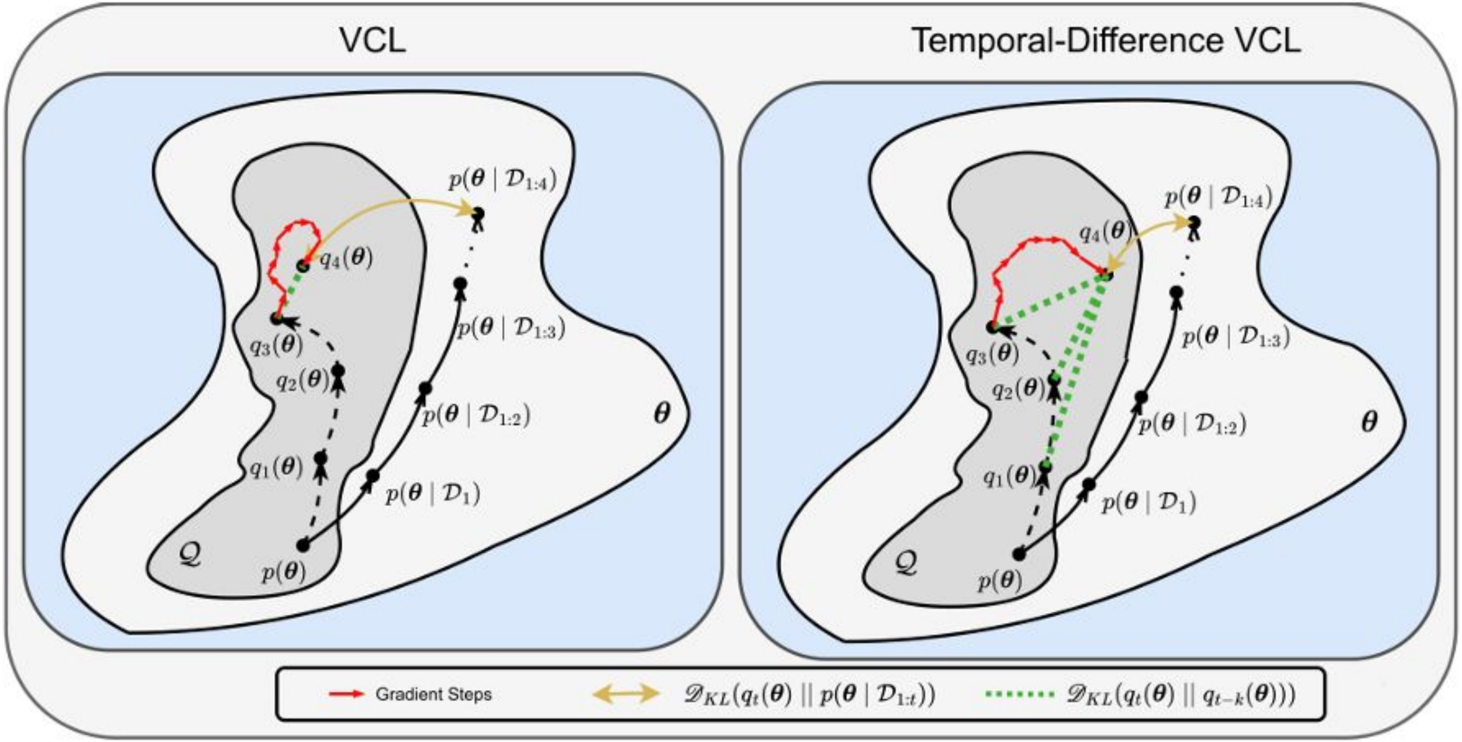
Problem

$$q_t(\boldsymbol{\theta}) = \arg\min_{q \in \mathcal{Q}} \mathscr{D}_{KL}(q(\boldsymbol{\theta}) \mid\mid \frac{1}{Z_t} q_{t-1}(\boldsymbol{\theta}) p(\mathcal{D}_t \mid \boldsymbol{\theta}))$$

Proposition 4.2. *The standard KL minimization objective in VCL (Equation 2) is equivalently represented as the following objective, with $n \in \mathbb{N}_0$, and $\lambda \in [0, 1)$ hyperparameters:*

$$q_t(\boldsymbol{\theta}) = \arg\max_{q \in \mathcal{Q}} \mathbb{E}_{\boldsymbol{\theta} \sim q_t(\boldsymbol{\theta})} \Big[\sum_{i=0}^{n-1} \frac{\lambda^i (1 - \lambda^{n-i})}{1 - \lambda^n} \log p(\mathcal{D}_{t-i} \mid \boldsymbol{\theta}) \Big] - \sum_{i=0}^{n-1} \frac{\lambda^i (1 - \lambda)}{1 - \lambda^n} \mathscr{D}_{KL}(q_t(\boldsymbol{\theta}) \mid\mid q_{t-i-1}(\boldsymbol{\theta})). \quad (5)$$

Temporal-Difference VCL



Temporal-Difference VCL

- We can define a “n-Step TD VCL target” as:

$$\mathrm{TD}_t(n) = \mathbb{E}_{\boldsymbol{\theta} \sim q_t(\boldsymbol{\theta})}\left[\sum_{i=0}^{n-1} \log p(\mathcal{D}_{t-i} \mid \boldsymbol{\theta})\right] - \mathscr{D}_{KL}(q_t(\boldsymbol{\theta}) \,||\, q_{t-n}(\boldsymbol{\theta}))$$

- 1) The optimization problem can be equivalently defined as, $\forall n \in \mathbb{N}_0,\ n \leq t$

$$q_t(\boldsymbol{\theta}) = \arg\max_{q \in \mathcal{Q}} \mathrm{TD}_t(n)$$

- 2) TD-VCL objective is a discounted sum of TD targets:

$$q_t(\boldsymbol{\theta}) = \arg\max_{q \in \mathcal{Q}} \frac{1-\lambda}{1-\lambda^n} \underbrace{\left[\sum_{k=0}^{n-1} \lambda^k \mathrm{TD}_t(k+1)\right]}_{\text{Discounted sum of TD targets}}.$$

- 3) **TD-VCL objectives generalize a spectrum of Continual Learning algorithms**

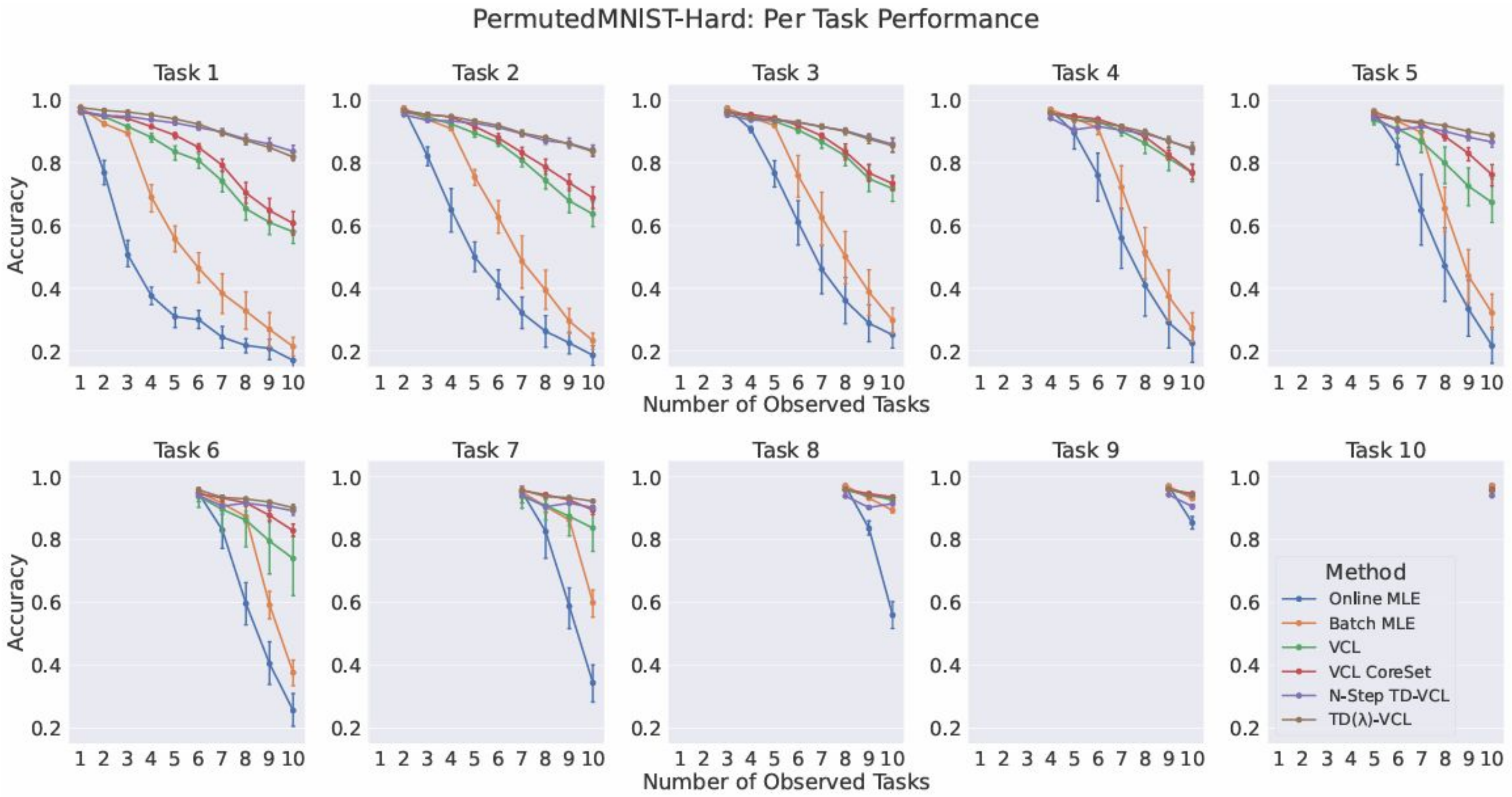
Results

	PermutedMNIST-Hard								
	t = 2	t = 3	t = 4	t = 5	t = 6	t = 7	t = 8	t = 9	t = 10
Online MLE	0.87±0.07	0.77±0.06	0.73±0.08	0.69±0.08	0.65±0.13	0.57±0.16	0.51±0.14	0.46±0.11	0.40±0.08
Batch MLE	0.95±0.01	0.93±0.01	0.88±0.04	0.83±0.04	0.77±0.10	0.71±0.13	0.64±0.12	0.57±0.11	0.51±0.06
VCL	0.95±0.00	0.94±0.01	0.93±0.02	0.91±0.02	0.89±0.03	0.86±0.03	0.83±0.04	0.80±0.06	0.78±0.04
VCL CoreSet	**0.96±0.00**	**0.95±0.00**	**0.94±0.00**	**0.93±0.02**	0.91±0.01	0.89±0.02	0.86±0.03	0.84±0.04	0.81±0.03
n-Step TD-VCL	0.95±0.01	0.94±0.00	**0.94±0.00**	**0.93±0.01**	**0.92±0.01**	**0.91±0.01**	**0.90±0.02**	**0.89±0.01**	**0.88±0.02**
TD(λ)-VCL	**0.97±0.00**	**0.96±0.00**	**0.95±0.00**	**0.94±0.01**	**0.93±0.01**	**0.92±0.01**	**0.91±0.01**	**0.90±0.01**	**0.89±0.02**

	SplitMNIST-Hard				SplitNotMNIST-Hard			
	t = 2	t = 3	t = 4	t = 5	t = 2	t = 3	t = 4	t = 5
Online MLE	0.86±0.02	0.61±0.03	0.75±0.04	0.57±0.06	0.72±0.02	0.61±0.05	0.61±0.00	0.51±0.04
Batch MLE	0.95±0.04	0.65±0.04	0.82±0.04	0.59±0.03	0.71±0.02	0.65±0.03	0.61±0.00	0.50±0.06
VCL	0.87±0.02	0.66±0.04	0.82±0.03	0.64±0.11	0.69±0.04	0.63±0.03	0.60±0.00	0.51±0.06
VCL CoreSet	0.93±0.04	0.68±0.07	0.84±0.04	0.62±0.03	0.69±0.04	0.65±0.02	0.60±0.01	0.51±0.07
n-Step TD-VCL	**0.98±0.01**	**0.79±0.08**	**0.88±0.04**	**0.67±0.04**	**0.72±0.04**	**0.73±0.05**	**0.70±0.04**	**0.58±0.08**
TD(λ)-VCL	**0.98±0.01**	**0.81±0.07**	**0.89±0.03**	**0.66±0.02**	**0.74±0.02**	**0.73±0.03**	**0.69±0.03**	**0.58±0.09**

	CIFAR100-10				TinyImageNet-10			
	t = 4	t = 6	t = 8	t = 10	t = 4	t = 6	t = 8	t = 10
Online MLE	0.57±0.06	0.56±0.03	0.53±0.06	0.52±0.04	0.45±0.02	0.44±0.01	0.45±0.02	0.44±0.03
Batch MLE	0.58±0.04	0.58±0.05	0.56±0.06	0.54±0.07	0.48±0.02	0.48±0.02	0.50±0.02	0.51±0.03
VCL	0.63±0.02	0.60±0.02	0.61±0.05	0.66±0.01	0.51±0.03	0.51±0.03	0.51±0.02	0.51±0.02
VCL CoreSet	0.63±0.03	0.63±0.02	0.61±0.02	0.65±0.02	0.51±0.02	0.51±0.02	0.54±0.02	0.54±0.02
n-Step TD-VCL	**0.67±0.02**	**0.65±0.01**	**0.68±0.04**	**0.69±0.02**	**0.55±0.02**	**0.54±0.02**	**0.56±0.02**	**0.56±0.02**
TD(λ)-VCL	**0.66±0.04**	**0.66±0.02**	**0.67±0.01**	**0.71±0.01**	**0.56±0.02**	**0.55±0.03**	**0.56±0.02**	**0.56±0.02**

Results



PermutedMNIST-Hard: Per Task Performance
Task 1
Task 2
Task 3
Task 4
Task 5
Task 6
Task 7
Task 8
Task 9
Task 10
Accuracy
Number of Observed Tasks
Method
Online MLE
Batch MLE
VCL
VCL CoreSet
N-Step TD-VCL
TD(λ)-VCL

Limitations / Future Work

- Limitations
    - Introduction of new hyperparameters, n and lambda
    - Relies on past posterior estimations, which increase memory requirements

- Future Work
    - Further connections with MDPs
    - Apply TD-VCL in problems that involves sequential VI, such as probabilistic meta-RL

Temporal-Difference Variational Continual Learning

Luckeciano Melo, Alessandro Abate, Yarin Gal

University of Oxford