Variational Continual Learning

Alejandro Santorum Varela, David Goldfarb, Vishaal Udandarao

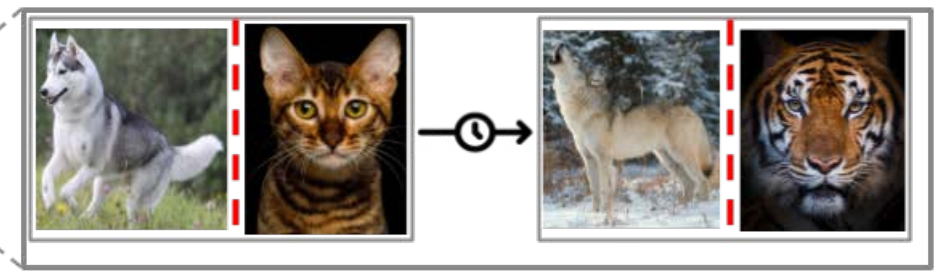
UNIVERSITY OF CAMBRIDGE

Introduction

Variational

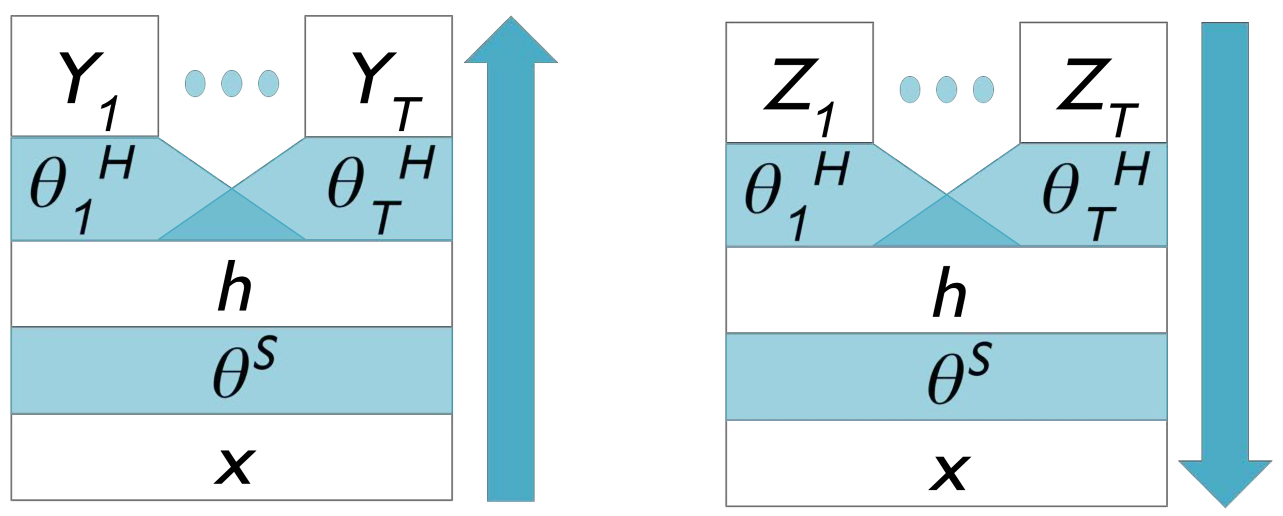
$$q_t(\boldsymbol{\theta}) = \arg\min_{q \in \mathcal{Q}} \mathrm{KL}\Big(q(\boldsymbol{\theta}) \,\|\, \frac{1}{Z_t} q_{t-1}(\boldsymbol{\theta})\, p(\mathcal{D}_t|\boldsymbol{\theta})\Big)$$

Continual Learning

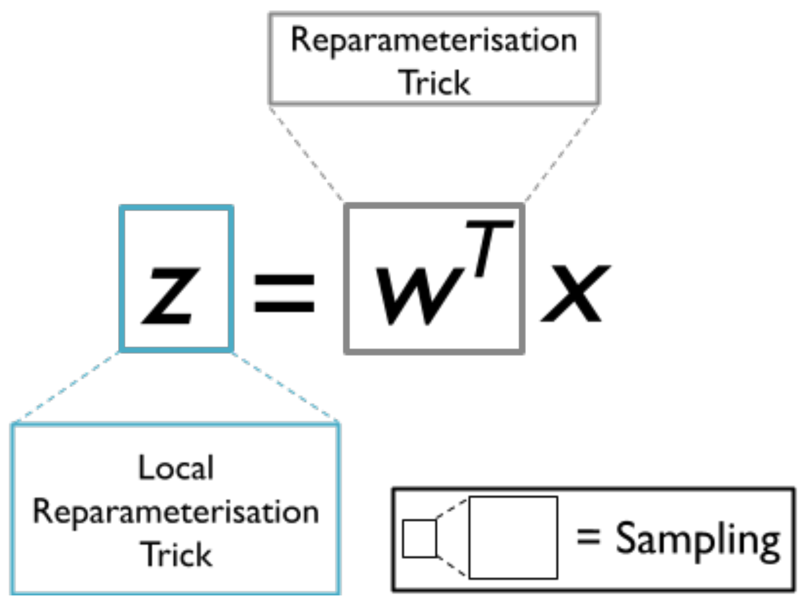


- Continual Learning (CL) requires balance between:
  - Plasticity (Catastrophic Forgetting)
  - Stability (Inability to adapt)

Methods

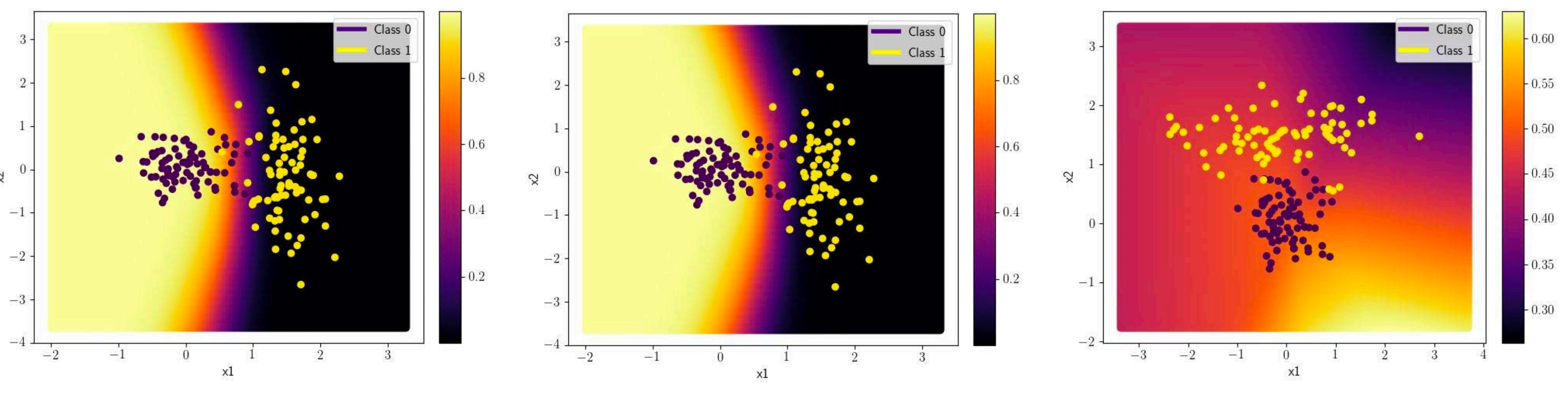


- Sampling:



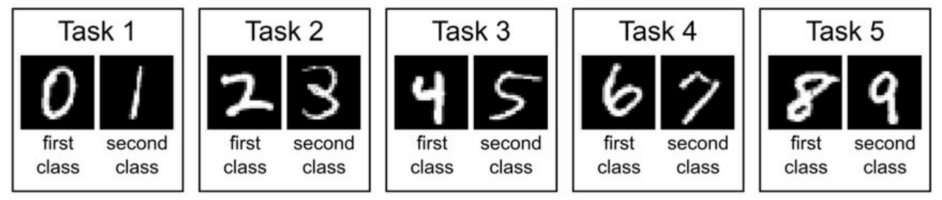
- Architectures:
  - Fully-Connected
  - CNNs
- Coreset Memory Enhancement:
  - Random Coresets
  - K-Center Coresets
  - Variational Generative Replay (VGR)

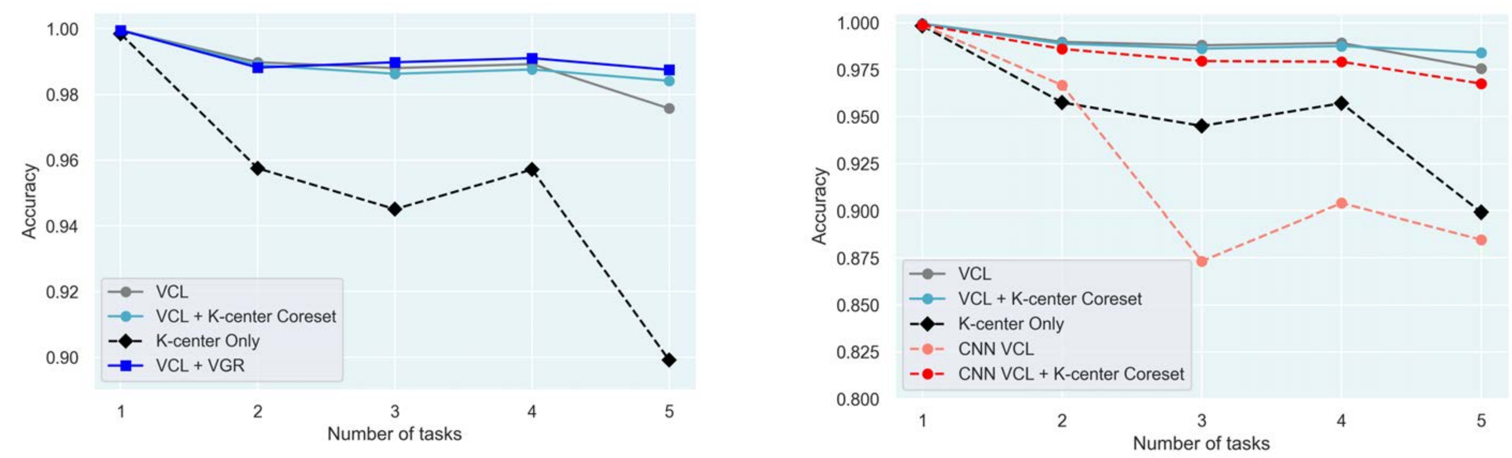
Toy Task



Train T1/Test T1 Train T2/Test T1 Train T2/Test T2

Split-MNIST

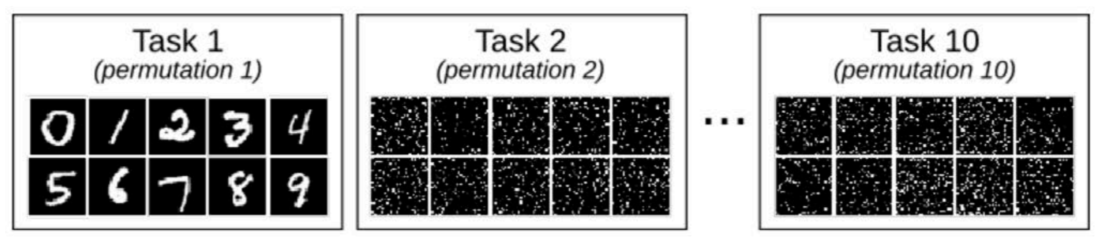


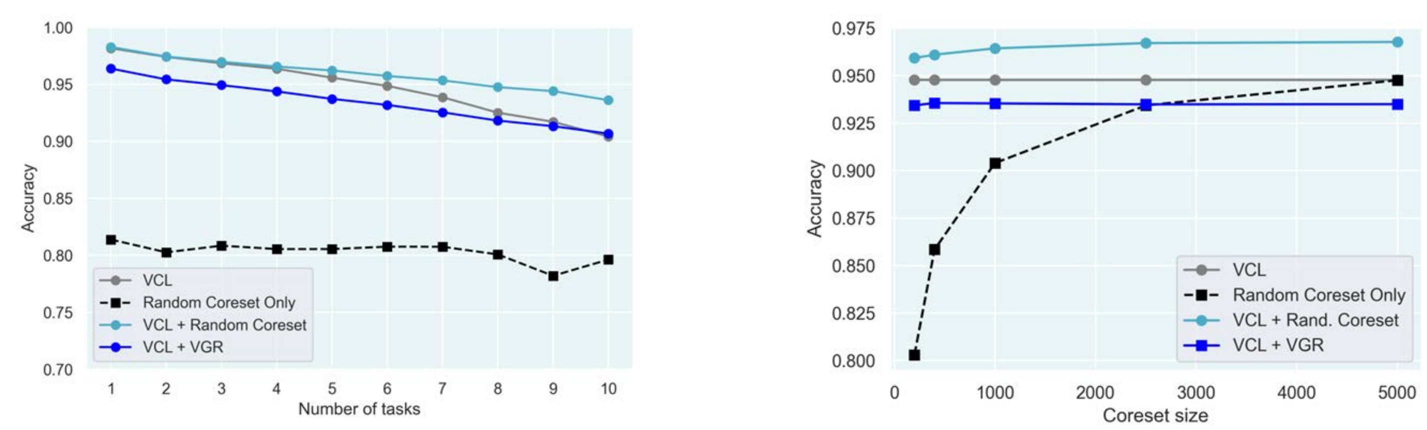


- VCL avoids catastrophic forgetting
- VGR improves on explicit coresets
- CNN: complexity vs. performance

	# of parameters	Accuracy
VCL	268800	97.57 %
VCL + Coreset		98.41 %
CNN VCL	9768	88.48 %
CNN VCL + Coreset		96.76 %

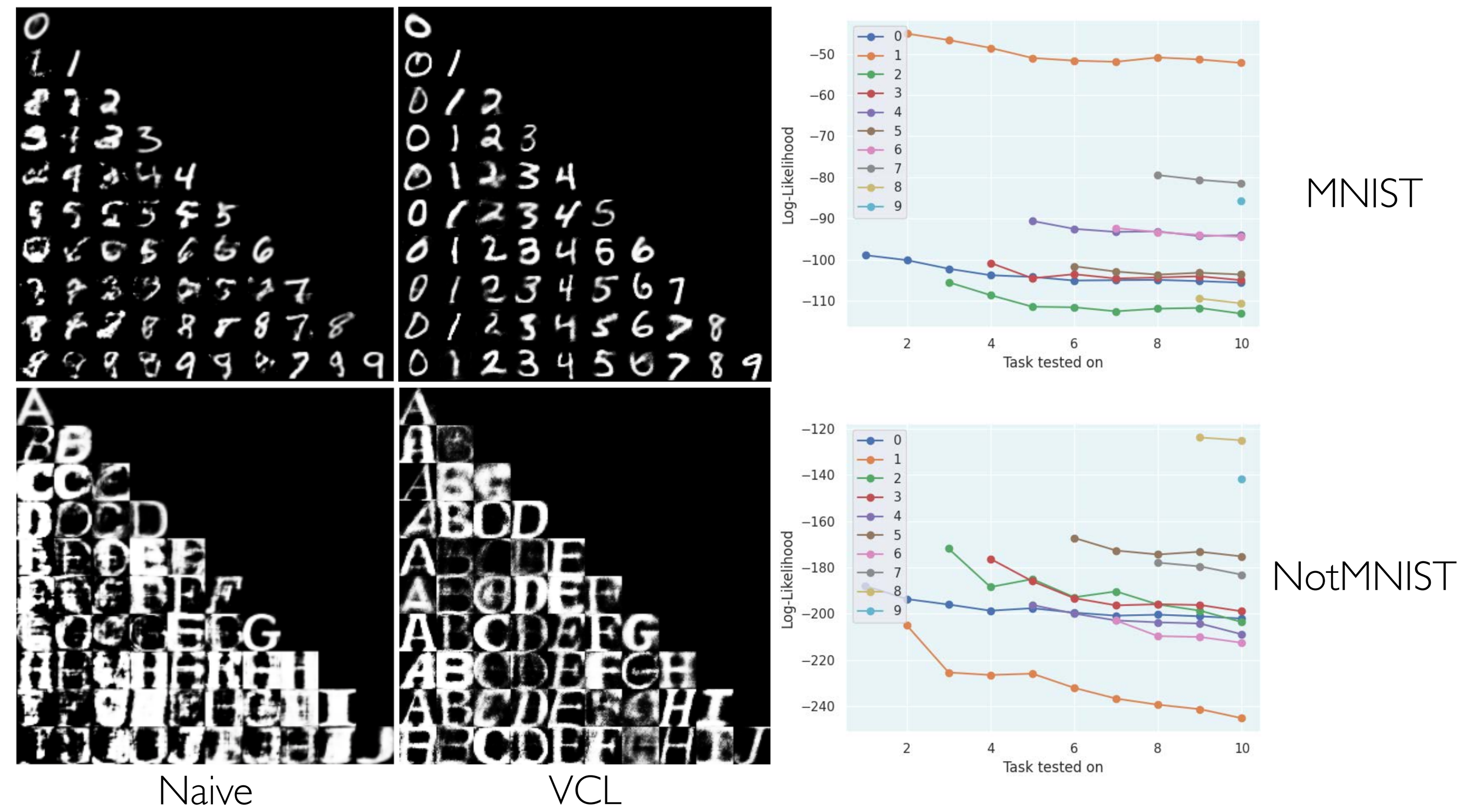
Permuted-MNIST





- VGR: underperforms from unstructured images
- Performance plateaus with coreset size

Generation



- VCL preserves old tasks' structure
- VCL outperforms other generative CL methods

Conclusion

- Variational methods improve over naive CL approaches
- VCL applies to both discriminative and generative models
- Coresets improve both naive and VCL performance
  - Selection algorithm inconsequential
- VGR reduces coresets memory footprint
- VCL can be applied to CNNs
  - CNN matches FC performance with fewer parameters
  - LRT suboptimal for CNNs

References

1. Cuong V. Nguyen et. al, Variational Continual Learning, 2018
2. Sebastian Farquhar et. al, A Unifying Bayesian View of Continual Learning, 2019
3. Diederik P. Kingma et. al, Variational Dropout and the Local Reparameterization Trick, 2015
4. Gido M. van de Ven et. al, Generative replay with feedback connections as a general strategy for continual learning, 2019

Code available at https://github.com/goldfarbDave/vcl