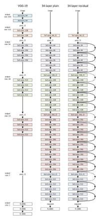
A Close Look at Deep Learning with Small Data

Lorenzo Brigato and Luca Iocchi

SAPIENZA UNIVERSITÀ DI ROMA

Deep Learning reached superior performance in many fields:

1. Lots of **data** (e.g. images, text)
2. High **capacity** neural networks (e.g ResNets)

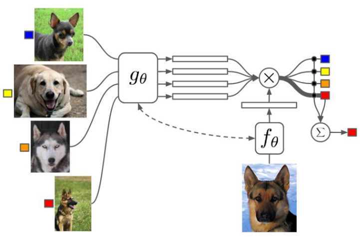


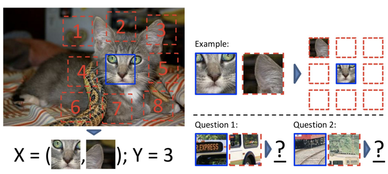
Problem:

1. **Obtaining/Labeling** data at large scales
   a. time-consuming
   b. difficult
   c. expensive

Popular approaches still rely on large source datasets:

1. Transfer/few-shot learning (labeled dataset)
2. Self-supervised learning (unlabeled dataset)





Focus on problems where $\mathcal{D}$ is balanced and relatively small (constraining number of samples per class N):

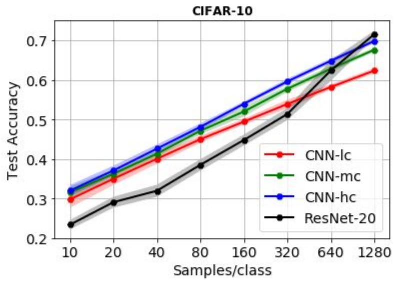
$$N \in \{10, 20, 40, 80, 160, 320, 640, 1280\}$$

Some previous works:

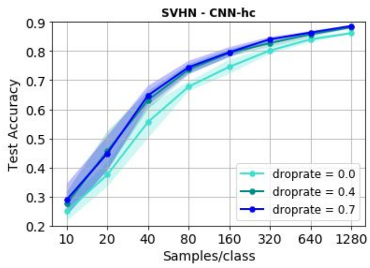
1. *Do we need hundreds of classifiers to solve real world classification problems?* [Fernandez et al. 2014]
2. *Modern neural networks generalize on small data sets* [Olson et al. 2018]
3. *Harnessing the Power of Infinitely Wide Deep Nets on Small-data Tasks* [Arora et al. 2020]
4. *Deep Learning on Small Datasets without Pre-Training using Cosine Loss* [Barz et al. 2020]

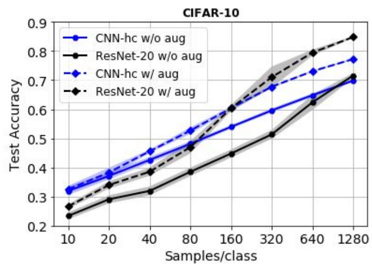
Empirical study:

- **Influence** of **model complexity** on performance
  a. CNN with 4 conv layers, diverse widths
  b. ResNet-20 with 16 base filters



- **Influence** of **regularization** techniques on performance
  c. Dropout -- drop-rates = 0.0/0.4/0.7
  d. Standard data augmentation





- **Comparison** of baseline models with **state-of-the art** approaches:
  e. CNTK [Arora et al. 2020]
  f. Cosine loss [Barz et al. 2020]