Image Classification using Semi-Supervised Learning

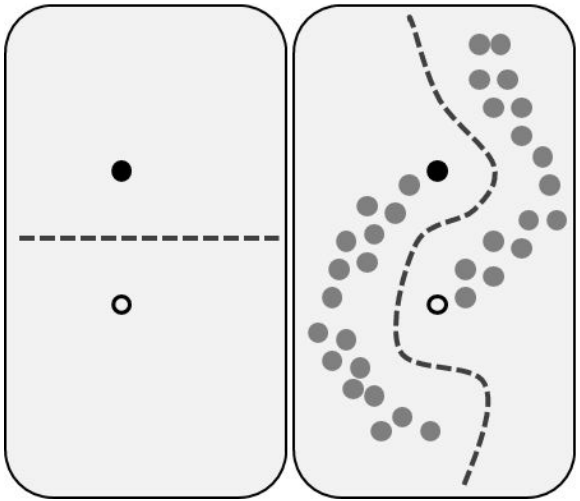
Deep Learning Final Project

Kuixian Zhu (kz1041@nyu.edu)
Marvin Mananghaya (msm796@nyu.edu)
Daniel Rivera Ruiz(drr342@nyu.edu)

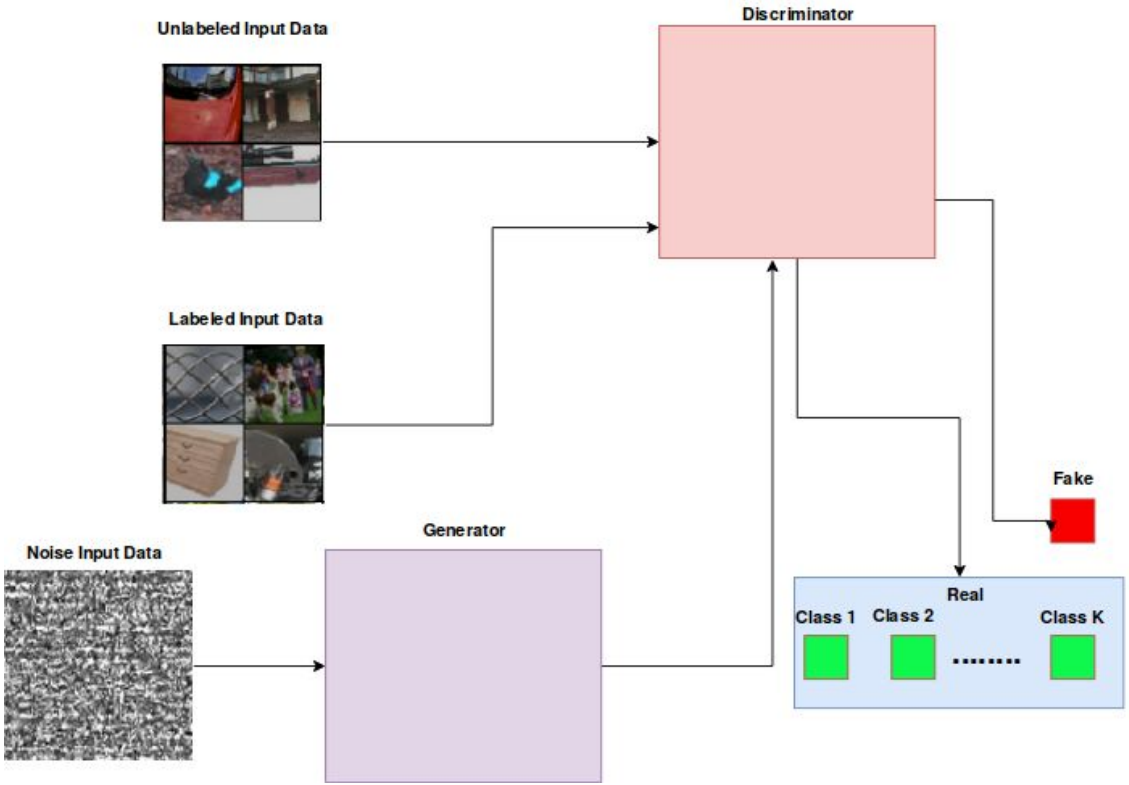
Deep Learning - Spring 2019
New York University

Problem Definition

- Image classification: 512k unlabeled images + 128k labeled images (64k train + 64k val), 1000 classes
- Approach: use SSL techniques to take advantage of the larger set of unlabeled images
- Algorithms explored:
  - SSL-DCGANs
  - Deep Clustering
  - Mean Teacher



SSL-DCGANs



In the style of Guillermo, 2019:

"So DCGANs do well in learning features unsupervised? **What if we do it semi-supervised?**"

- Radford et al., 2015

> "You could totally do that, just change the classification problem **from real vs fake to real k classes and fake**"
>
> -Salimans et al., 2016

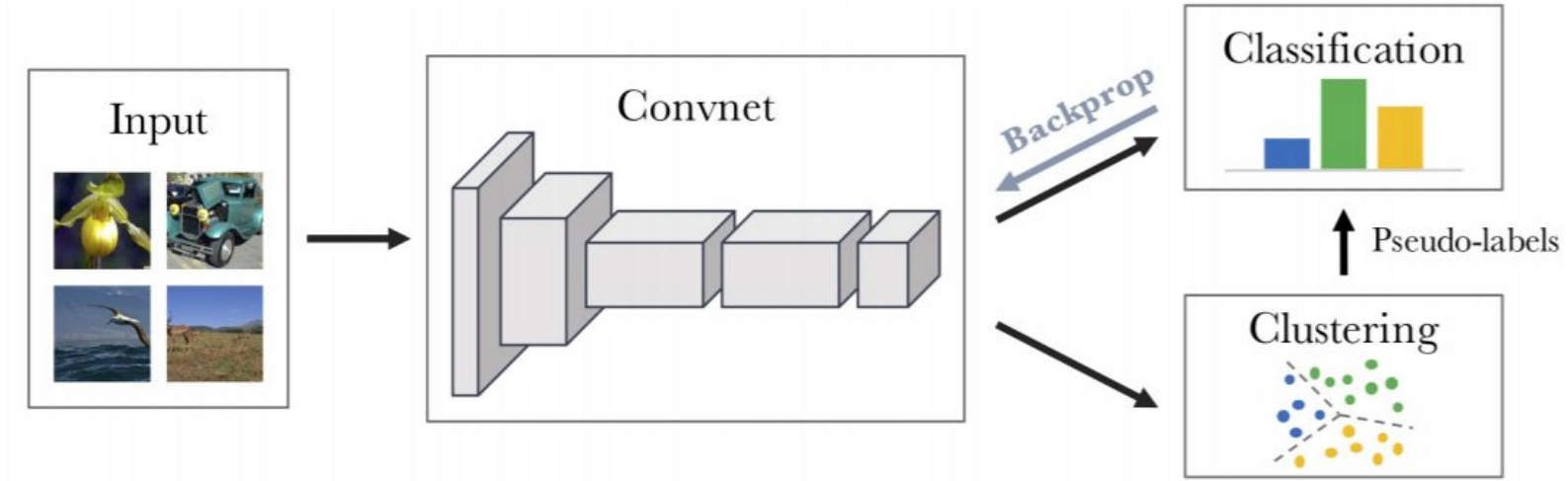
But what happened? Tl;dr It didn't work out so well.

- GANs are **sensitive** to hyperparameter changes
- **Easy to fall** to non-convergence
- One network **can overpower another**

Maybe there are **simpler** yet **better** ways to perform semi-supervised learning?

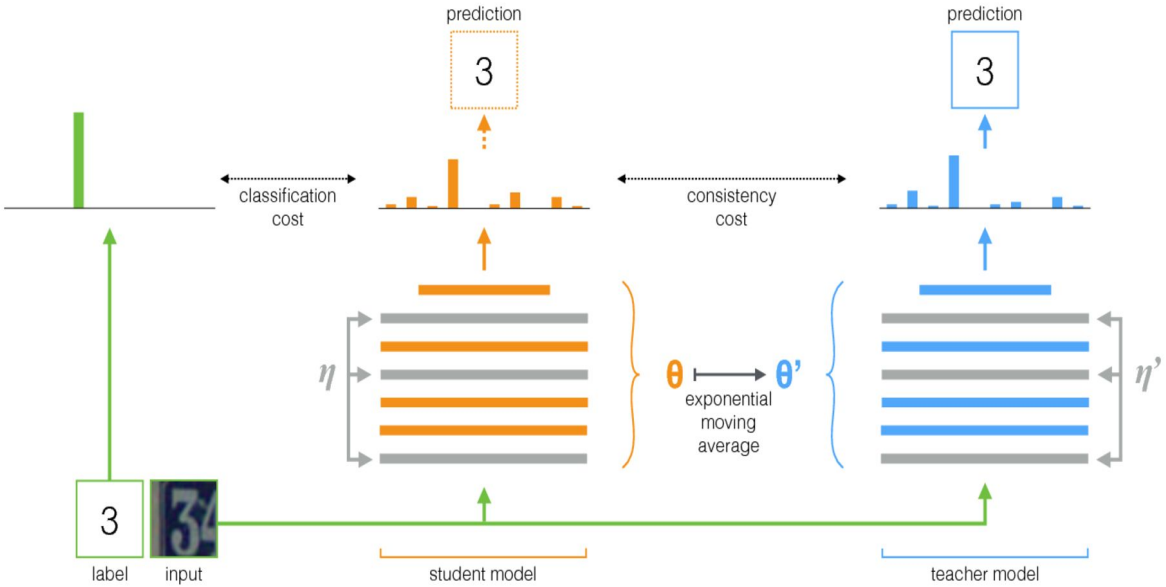
Deep Clustering

- Basic idea: alternate between clustering of the image descriptors and updating the weights of a convnet by predicting the cluster assignments.



Mean Teacher

Mean Teacher is a simple, yet scalable framework for semi-supervised learning.



1. Take any supervised model, call it the **student**.

2. Make a copy of the student model, call it the **teacher**. Let the teacher's weights to be the exponential moving average (EMA) of the student's weights.

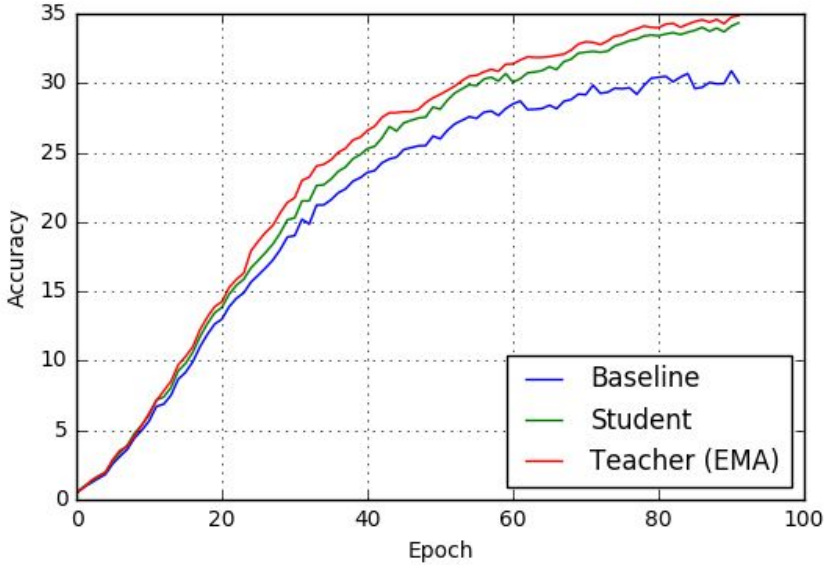
3. While training, use the same images as inputs to both models, but add random noise (augmentation) to them separately.

4. Update the student by the classification loss as usual, yet add an additional *consistency loss* between the student and teacher outputs.

Results and Analysis

- Mean Teacher training curves

- Model comparison (validation accuracy)



Model	Labeled examples per class during training							
	1	2	4	8	16	32	64 acc@1	64 acc@5
Baseline (Resnet152)	–	–	–	–	–	–	30.45	55.20
SSL-DCGAN	–	–	–	–	–	–	10.10	27.04
Mean Teacher (Resnet152)	1.71	1.65	–	–	15.21	20.59	**34.38**	**59.81**
Deep Clustering	1.40	2.38	3.38	5.60	7.71	9.39	13.48	28.61

Conclusions

- SSL can be very powerful because it fully exploits unlabeled data, which is easier and cheaper to collect.
- In general, deep learning models rely greatly in the amount of data available and SSL models are no exception to this rule.
- Due to time and computing resources limitations, we were not able to push any of our models to the limit.