Combating Label Noise in Deep Learning using Abstention

Speaker: Sunil Thulasidasan
sunil@lanl.gov

Los Alamos NATIONAL LABORATORY EST. 1943

UNIVERSITY OF WASHINGTON ELECTRICAL ENGINEERING

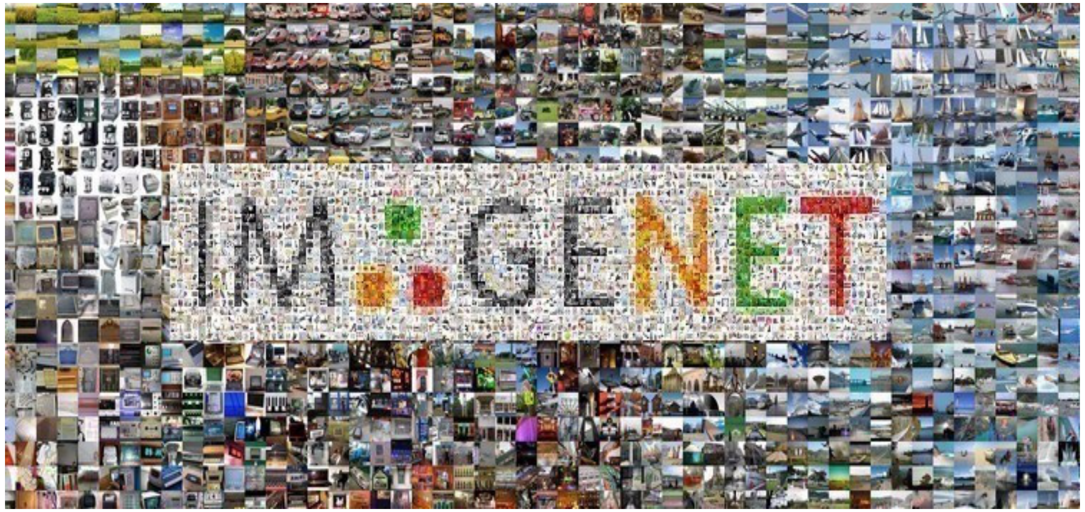
sunil@lanl.gov

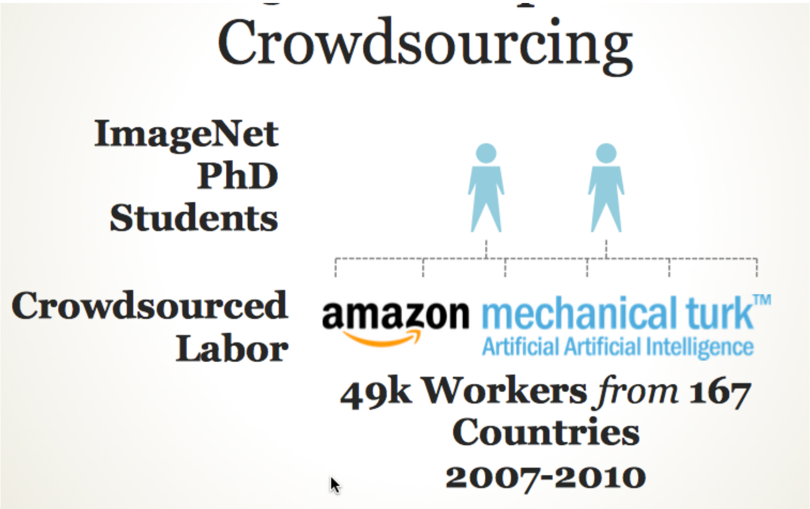
A Practical Challenge for Deep Learning

State-of-the-art models require *large amounts of* ***clean****, annotated data.*

sunil@lanl.gov

Annotation is labor intensive!





Slide from Fei-Fei Li and Jia Deng

ImageNet: 15 million labeled images; over 20,000 classes

The data that transformed AI research—and possibly the world (D. Gershgorn, quartz, magazine, 2017)

- 49k workers
- 167 countries
- 2.5 years to complete!

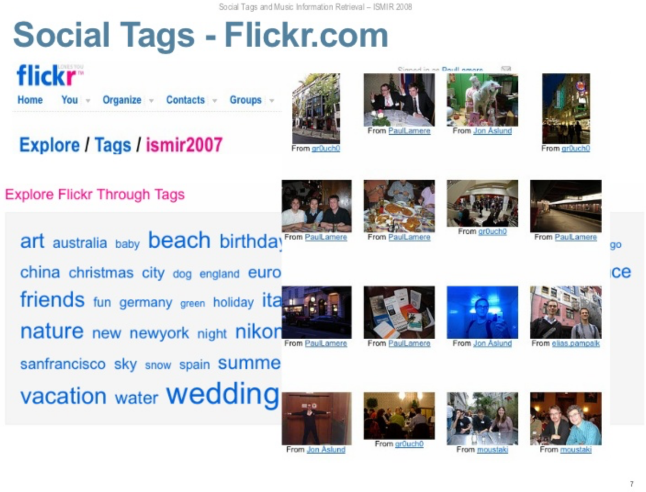
sunil@lanl.gov

Approaches to large-scale labeling

- Crowdsource at scale – labor intensive, but relatively cheap

amazon
beta
mechanical turk

- Use weak labels from queries, user tags and pre-trained classifiers



sunil@lanl.gov

Approaches to large-scale labeling

- Crowdsource at scale – labor intensive, but

Both approaches can lead to significant labeling errors!

- Use weak labels from queries, user tags and pre-trained classifiers

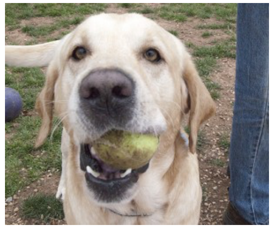
Dog

Taxi

Banana

Slide credit: S Guo et al '2018

- Label noise is an inconsistent mapping from features ***X*** to labels ***Y***



Dog

Dog

Dog

The Deep Abstaining Classifier (DAC)

Approach: Use learning difficulty on incorrectly labeled or confusing samples to defer on learning -- ***"abstain"*** -- till correct mapping is learned.

sunil@lanl.gov

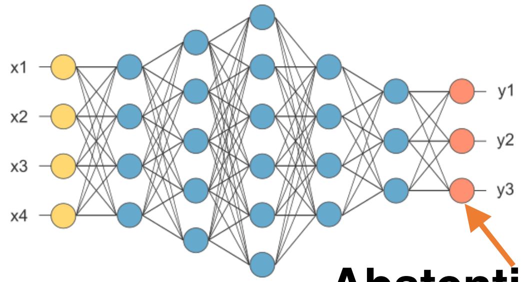
Training a Deep Abstaining Classifier

$$\mathcal{L}(x) = (1 - p(x)_{k+1}) \left(- \sum_{i=1}^{k} t(x)_i \log \frac{p(x)_i}{1 - p(x)_{k+1}} \right) + \alpha \log \frac{1}{1 - p(x)_{k+1}}$$

Cross entropy as usual

sunil@lanl.gov

Training a Deep Abstaining Classifier



Abstention class

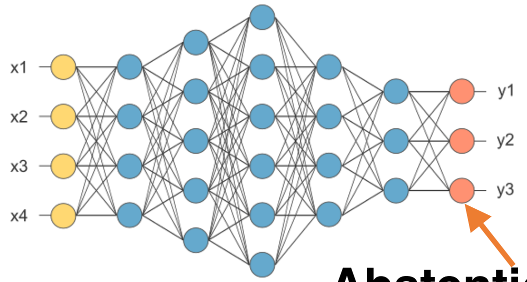
$$\mathcal{L}(x) = (1 - p(x)_{k+1})\left(-\sum_{i=1}^{k} t(x)_i \log \frac{p(x)_i}{1 - p(x)_{k+1}}\right) + \alpha \log \frac{1}{1 - p(x)_{k+1}}$$

Encourages abstention

Cross entropy over actual classes

sunil@lanl.gov

Training a Deep Abstaining Classifier



Abstention class

$$\mathcal{L}(x) = (1 - p(x)_{k+1})\left(-\sum_{i=1}^{k} t(x)_i \log \frac{p(x)_i}{1 - p(x)_{k+1}}\right) + \alpha \log \frac{1}{1 - p(x)_{k+1}}$$

Encourages abstention

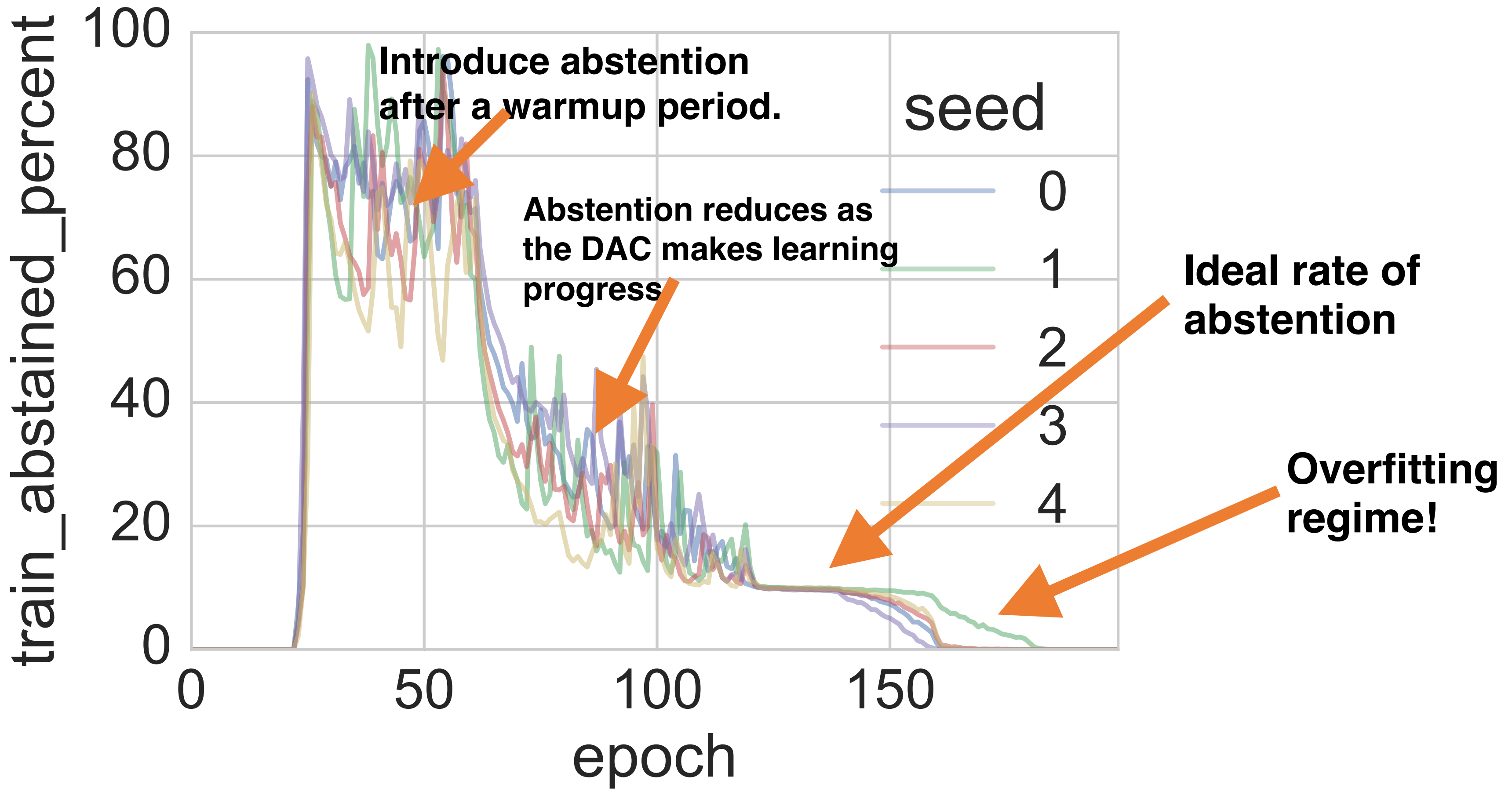
Cross entropy over actual classes

Penalizes abstention

Automatically tuned during learning.

sunil@lanl.gov

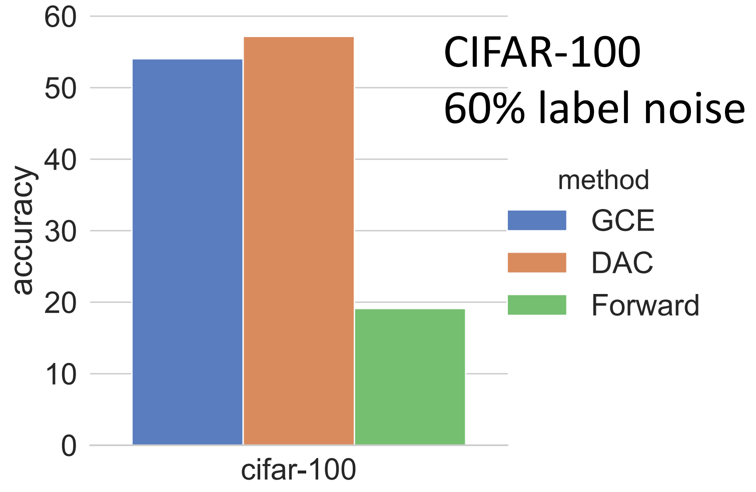
Abstention Dynamics



Abstained percent on training set vs epoch with **10% label noise**.

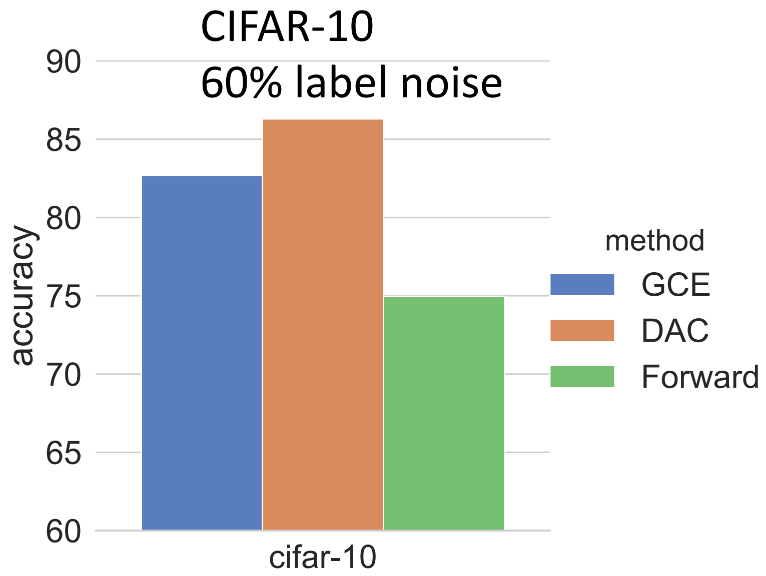
sunil@lanl.gov

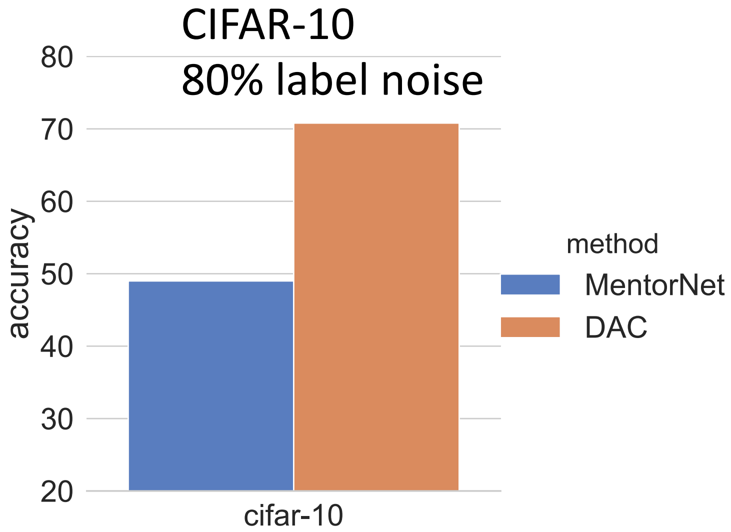
The DAC gives state-of-art results in label-noise experiments.



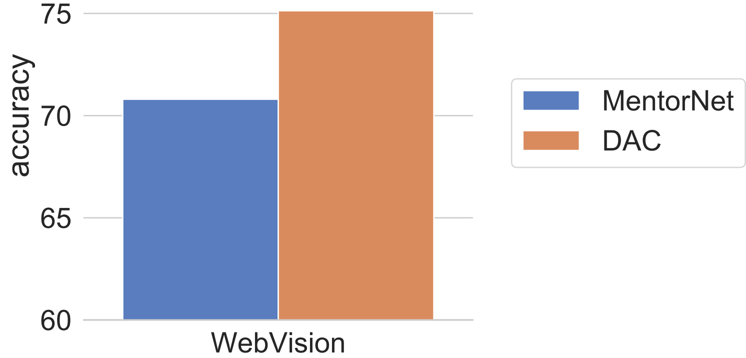
Training protocol:

- **Use DAC to identify and eliminate label noise.**
- **Retrain on cleaner set.**





WebVision: Real-world noisy dataset. ~2.4M images. ~35-40% label noise



GCE: Generalized Cross-Entropy Loss (Zhang et al NIPS '18); Forward (Patrini et al, CVPR '17); MentorNet (Li et al, ICML '18)

sunil@lanl.gov

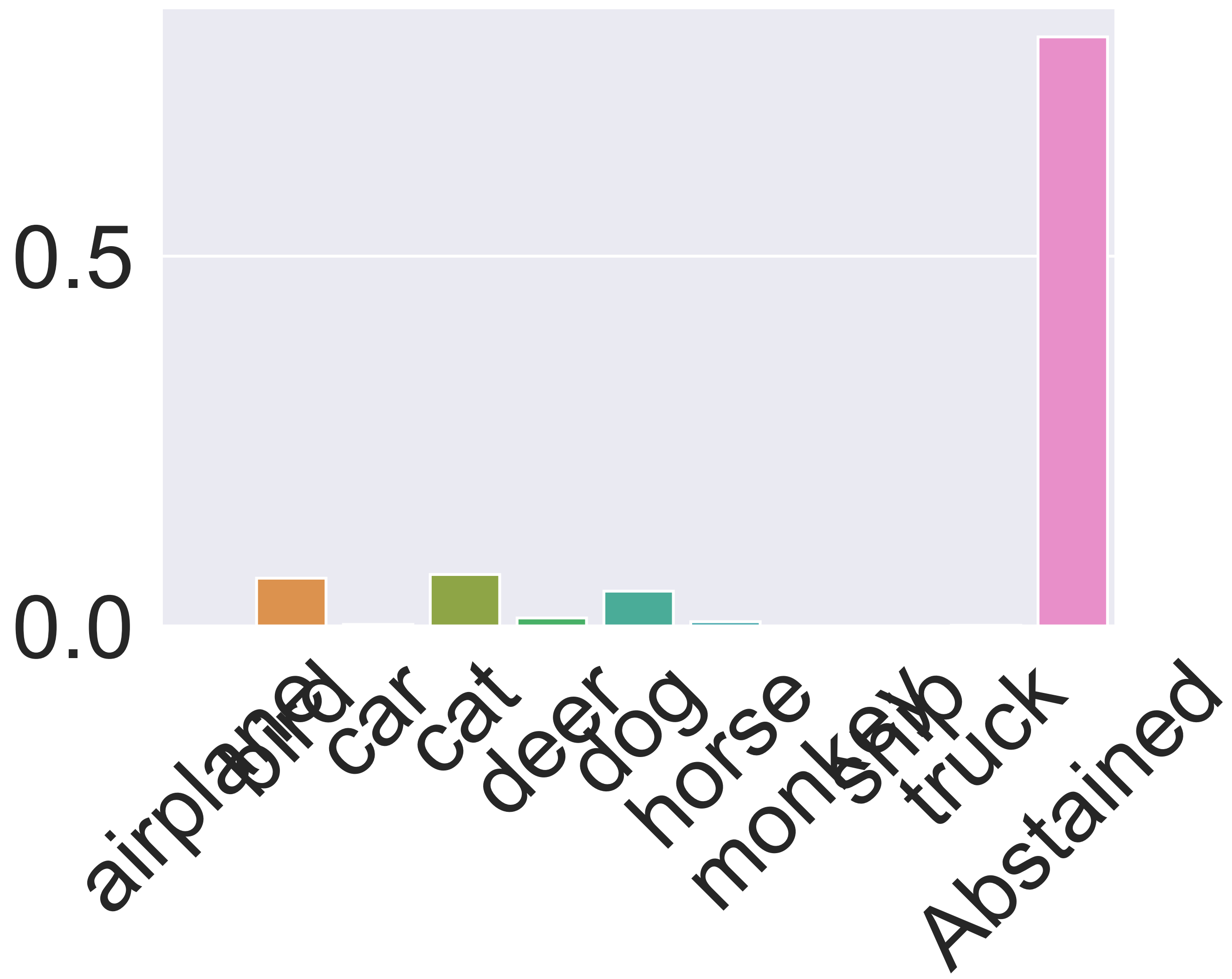
Abstention in the presence of Systematic Label Noise: The Random Monkeys Experiment

All the monkey labels in the training set (STL-10) are randomized.

Can the DAC learn that images containing monkey features have unreliable labels and abstain on monkeys in the ***test set?***

sunil@lanl.gov

Random Monkeys: DAC Predictions on Monkey Images



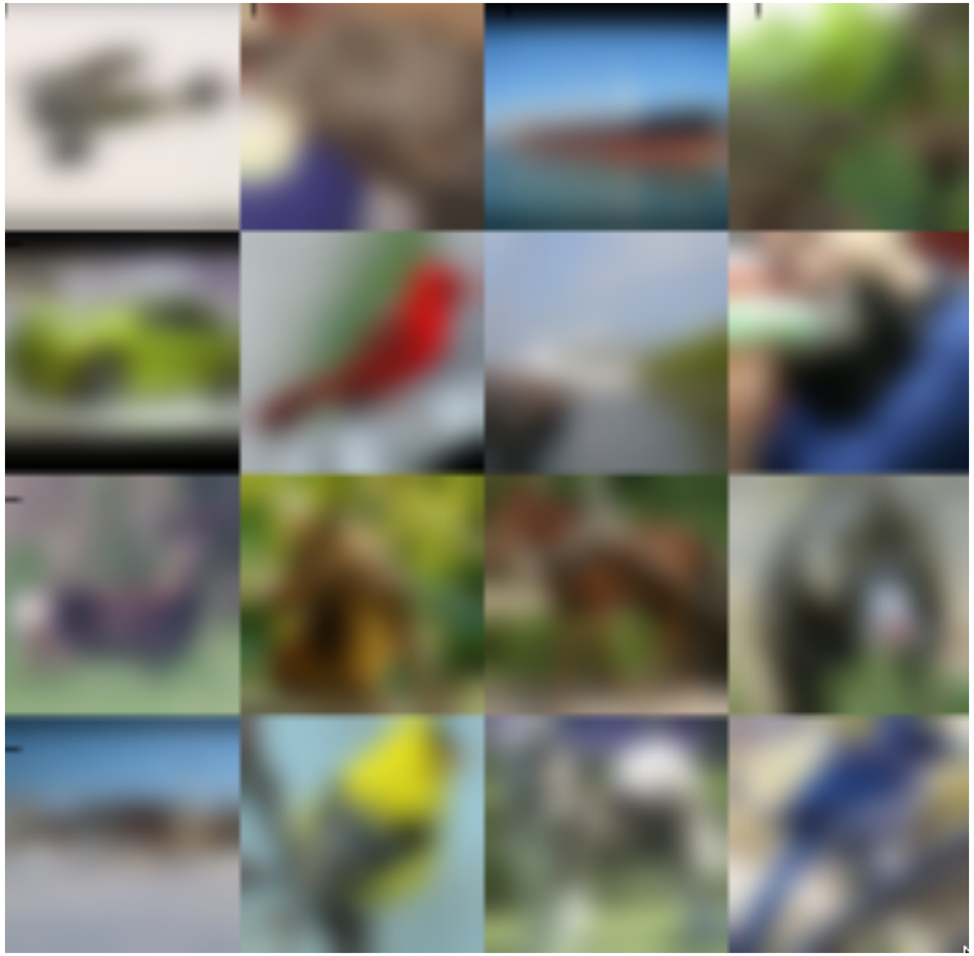
The DAC abstains on most of the monkeys in the test set!

sunil@lanl.gov

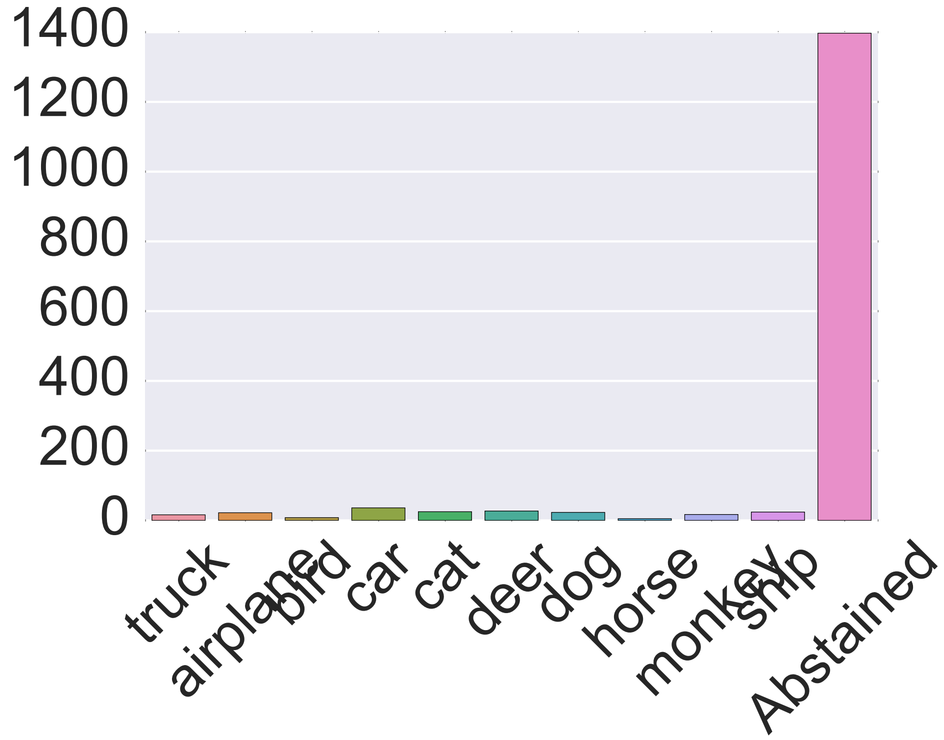
Image Blurring

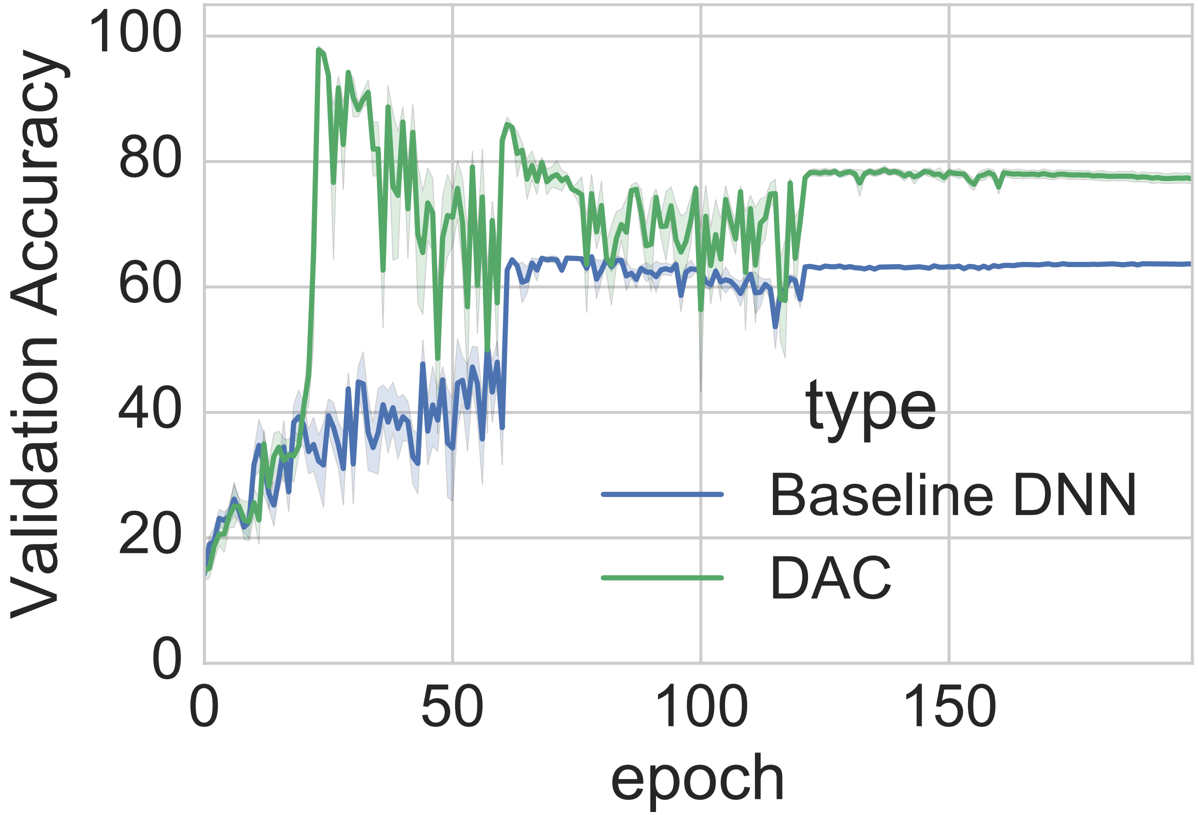
Blur a subset (20%) of the images in the training set and randomize labels

Will the DAC learn to abstain on blurred images in the test set?



DAC Behavior on Blurred Images





DAC abstains on most of the blurred images in the test set

For DAC, validation accuracy is calculated on non-abstained samples.

sunil@lanl.gov

Conclusions

Code available at https://github.com/thulas/dac-label-noise

- **Abstention training is an effective way to clean label noise in a deep learning pipeline**.

- Abstention can also be used as a *representation learner* for label noise.
  - Especially useful for interpretability in “don’t-know” decision situations.

Code available at https://github.com/thulas/dac-label-noise

Joint work with…….

Tanmoy Bhattacharya
Los Alamos National Lab

Jeff Bilmes
University of Washington

Gopinath Chennupati
Los Alamos National Lab

Jamal Mohd-Yusof
Los Alamos National Lab

Poster:
Tue Jun 11th
06:30 -- 09:00 PM @ Pacific Ballroom **#9**

Point of Contact:
Sunil Thulasidasan
(sunil@lanl.gov)