,

Assignment 1

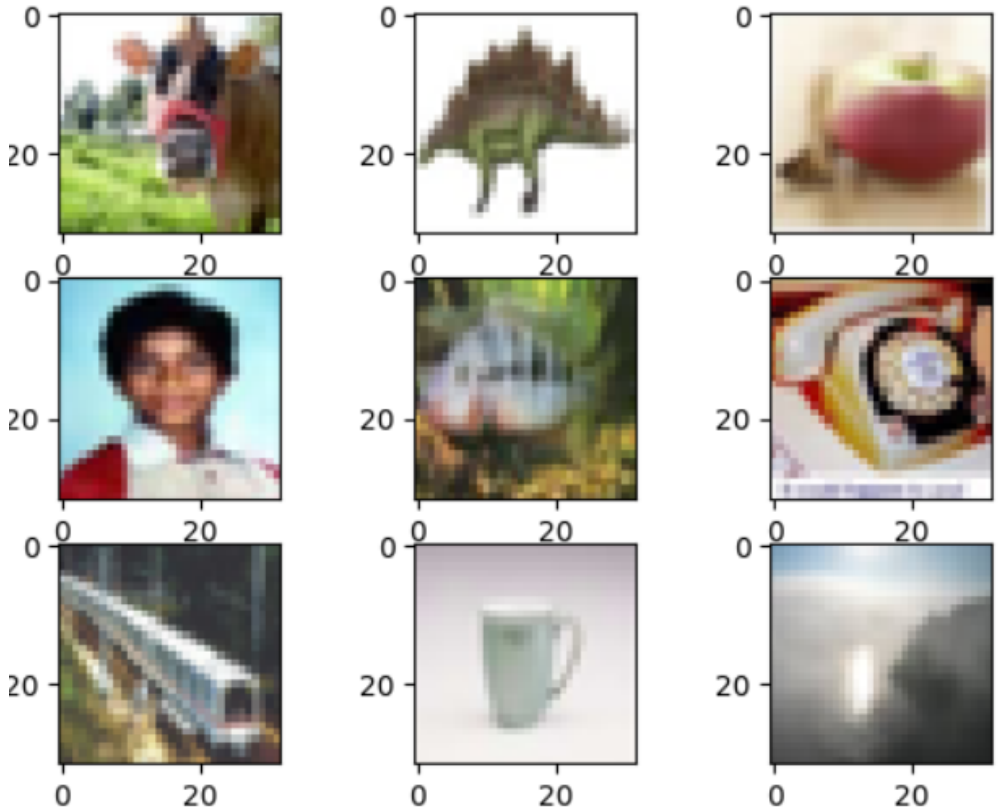
David Gardner
Boston University
CRCV, University of Central Florida

June 2, 2025

Dataset: CIFAR-100

100 classes of 32x32 color photorealistic images from the internet.

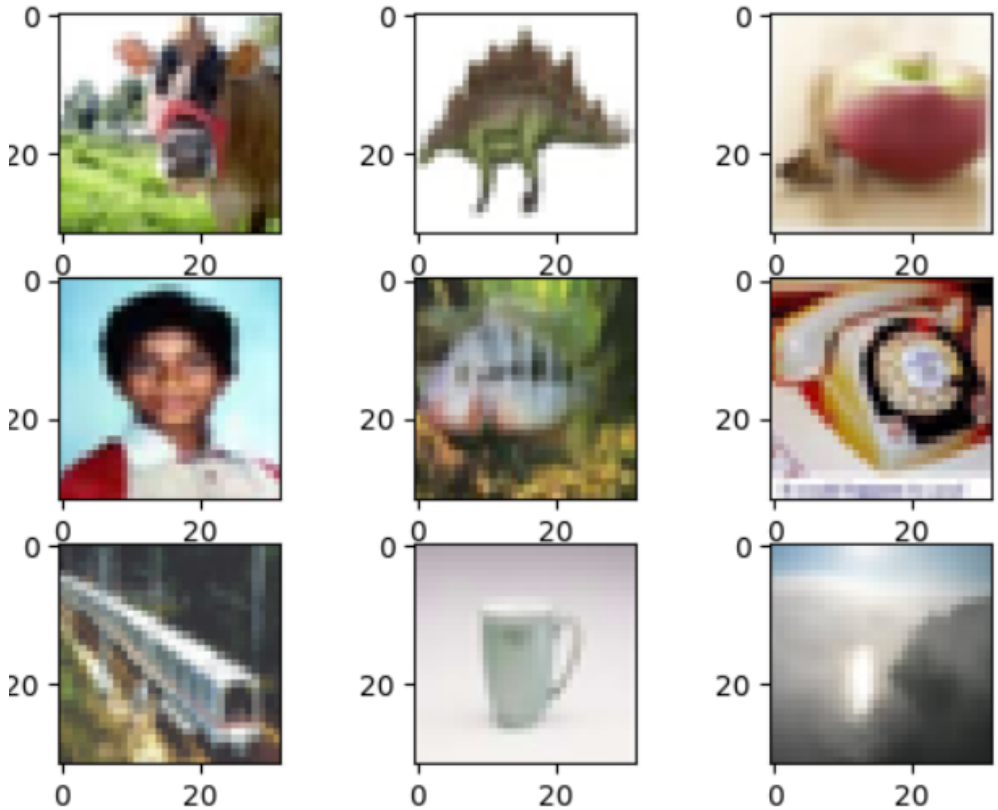
1. 80 Million Tiny Images dataset in 2008



Dataset: CIFAR-100

100 classes of 32x32 color photorealistic images from the internet.

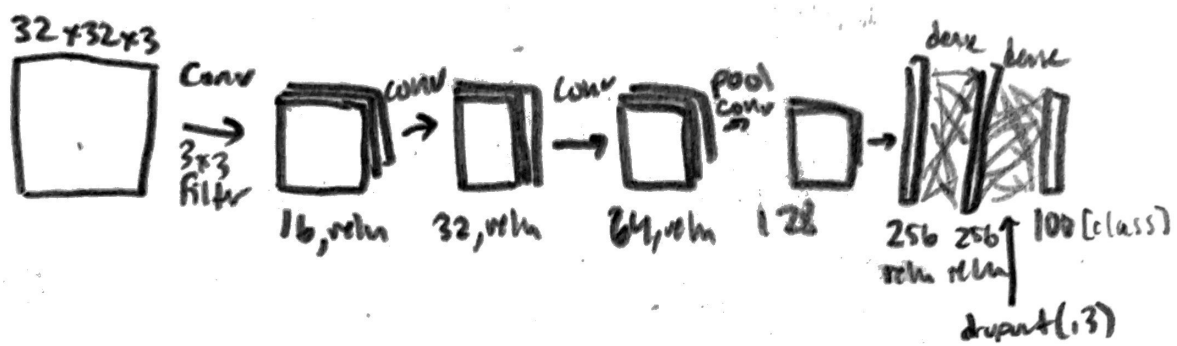
1. 80 Million Tiny Images dataset in 2008
2. Alex Krizhevsky labeled 6,000 in 2009 Krizhevsky 2009



Methods

Matching existing implementation

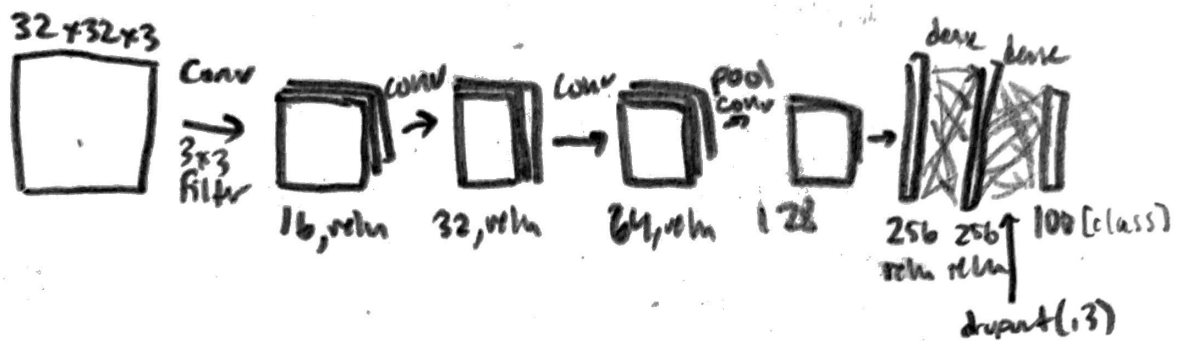
1. "Learning Multiple Layers of Features from Tiny Images" (2007)



Methods

Matching existing implementation

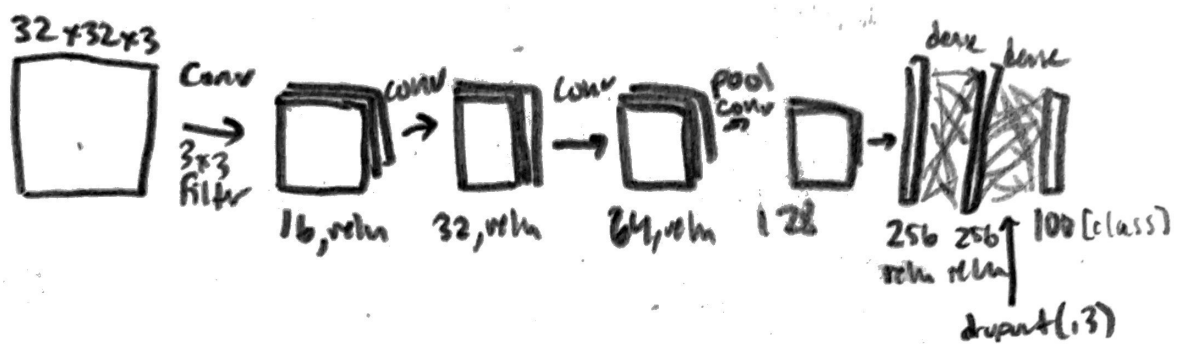
1. "Learning Multiple Layers of Features from Tiny Images" (2007)
    (i) **Whitening**
    (ii) Restricted Boltzmann Machine (RBM)



Methods

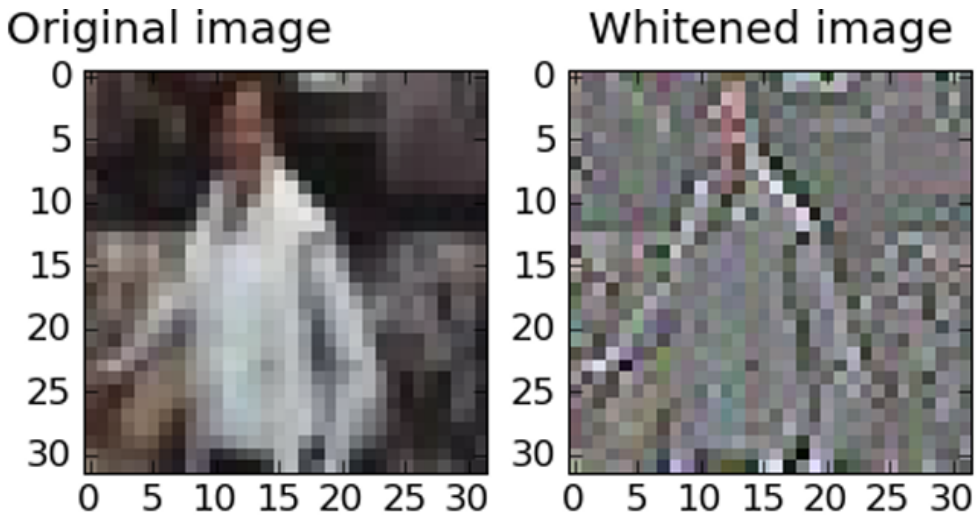
Matching existing implementation

1. "Learning Multiple Layers of Features from Tiny Images" (2007)
    (i) **Whitening**
    (ii) Restricted Boltzmann Machine (RBM)
2. Architecture from Geeks4Geeks, Assignment 0



Methods: Whitening

Prevents the model from focusing on irrelevant information Wiesler et al. 2014.



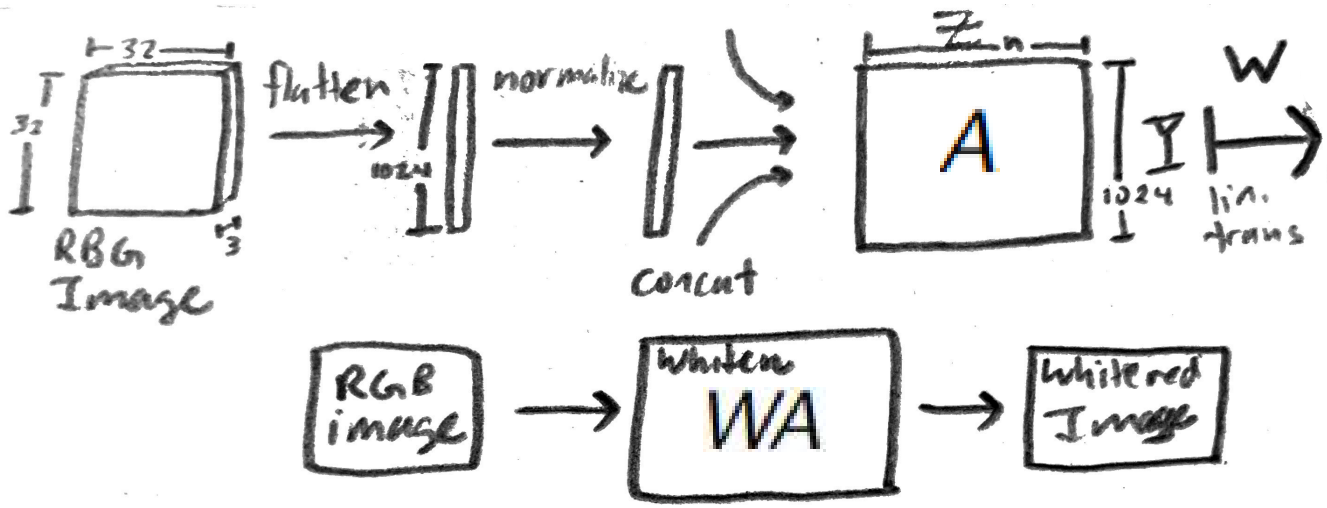
Credit: Alex Krizhevsky

Methods: Whitening

The sample covariance matrix of Y given n samples is

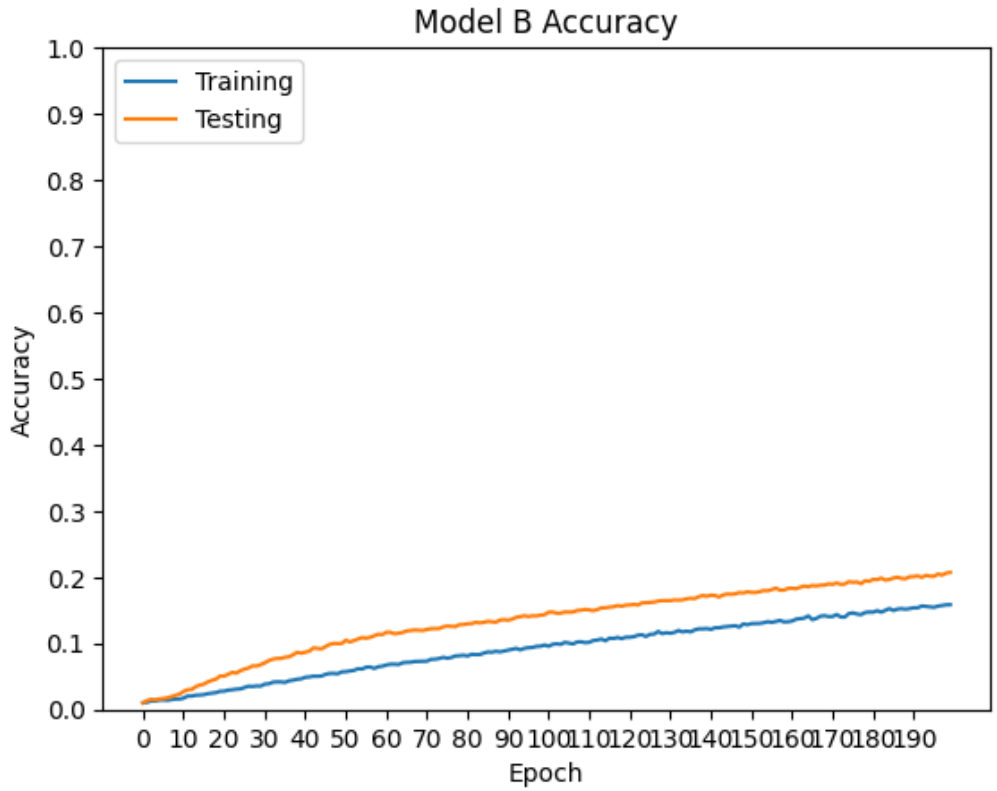
$$C = \frac{1}{n-1} AA^T.$$

We want to find W such that the sample covariance matrix of WA is the identity matrix.



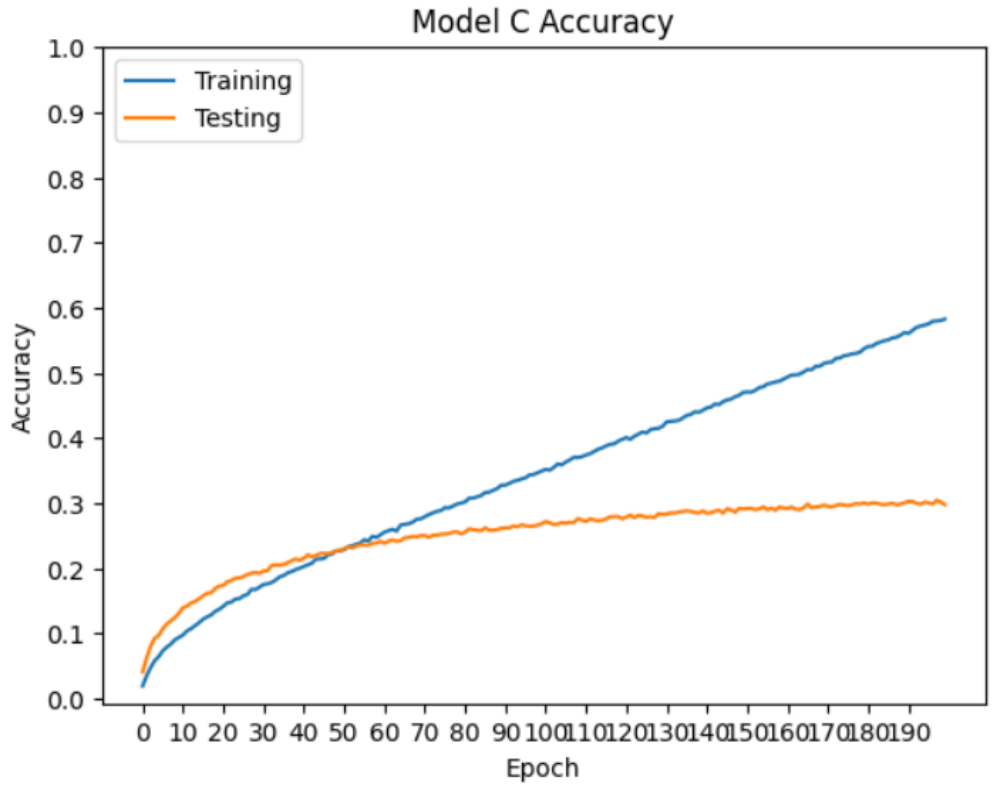
Experiments: Model B (Assignment 0)

200 epochs, 32 batch size — accuracy: 20.7% (ok...)



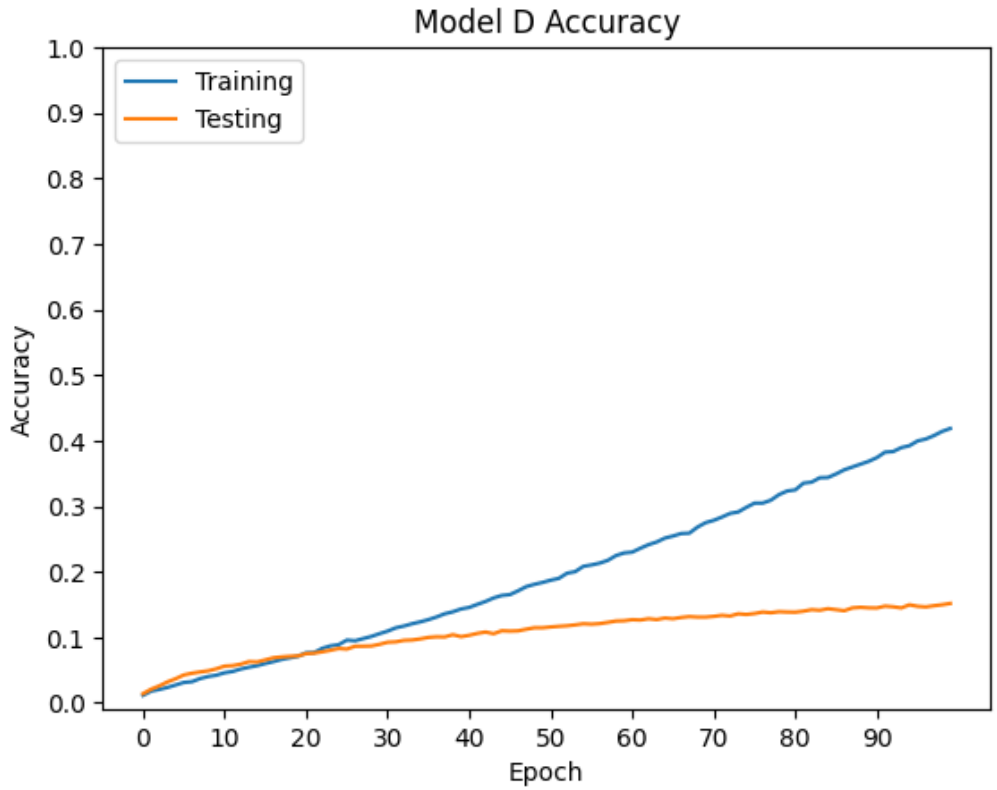
Experiments: Model C (Geeks4Geeks, Original)

200 epochs, 32 batch size — accuracy: 30.4% (overfitting)



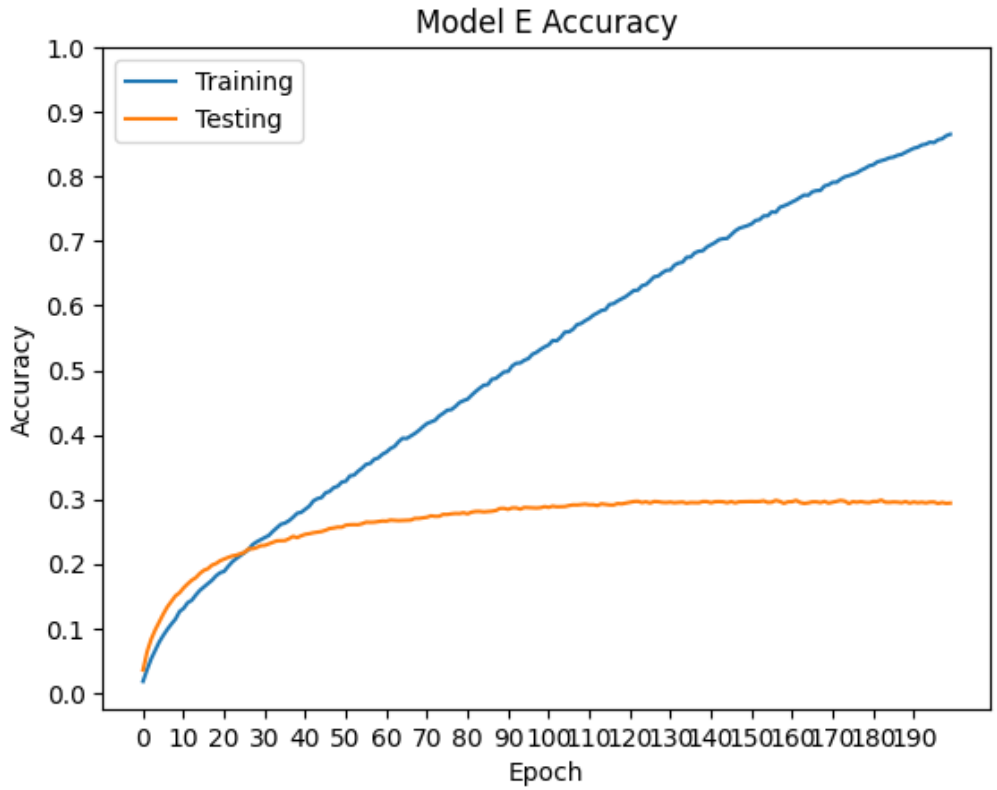
Experiments: Model D (Geeks4Geeks, Whitened)

200 epochs, 32 batch size — accuracy: 15.2% (whitening hurt performance)



Experiments: Model E (Geeks4Geeks, Batch Normalizations)

200 epochs, 32 batch size — accuracy: 29.6% (overfitting!!)



Discussion

Matching existing implementation

1. Whitening, batch normalization doesn't always help.

Discussion

Matching existing implementation

1. Whitening, batch normalization doesn't always help.

Discussion

Matching existing implementation

1. Whitening, batch normalization doesn't always help.
2. Different architectures to include whitening

Bibliography

- Krizhevsky, Alex (2009). *Learning Multiple Layers of Features from Tiny Images*. Technical Report. https://www.cs.toronto.edu/ kriz/learning-features-2009-TR.pdf. Toronto, ON, Canada: University of Toronto.
- Wiesler, Simon et al. (2014). "Mean-normalized stochastic gradient for large-scale deep learning". In: *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 180–184. DOI: 10.1109/ICASSP.2014.6853582.