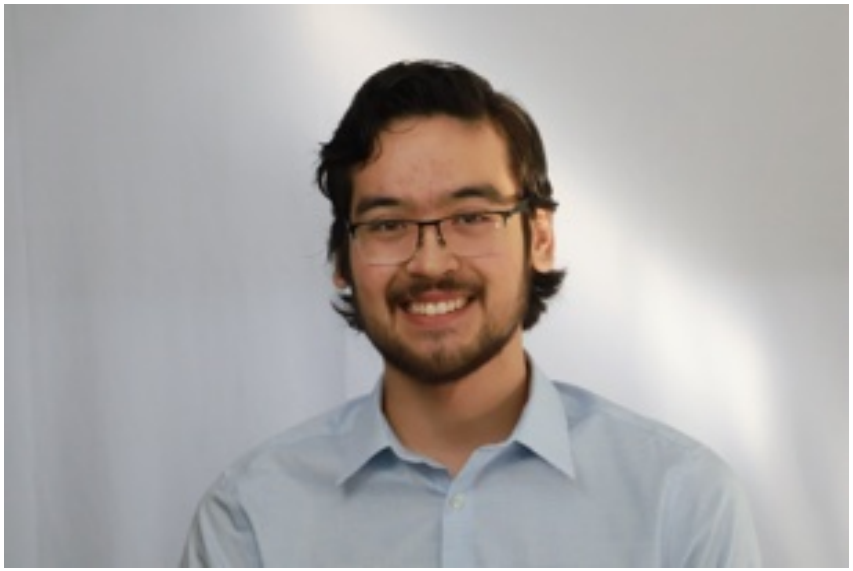
,

Week 3 Presentation

David Gardner
Boston University
CRCV, University of Central Florida

June 4, 2025

Who am I?



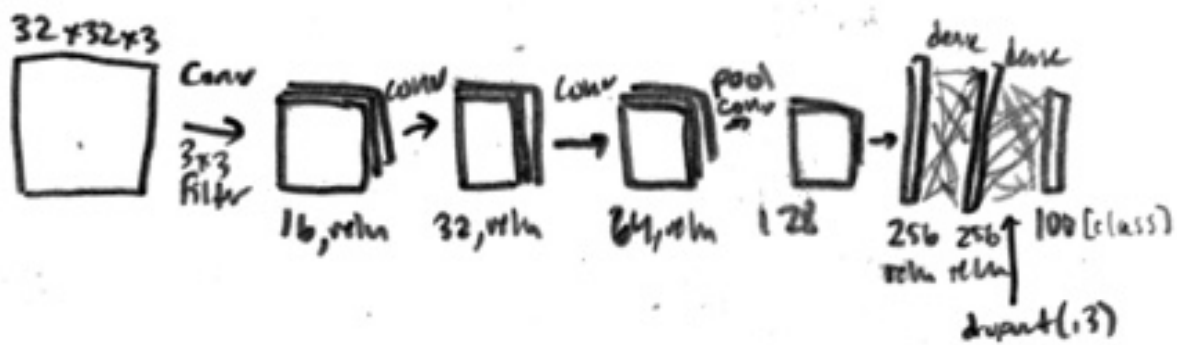
David Gardner

- Electrical Engineering, Math
- Data Science Programming I/II, Control Systems, Real Analysis I/II/III, Probability Theory, H. Linear Algebra, Dynamical Systems, Modern Algebra I

Assignment 1: CIFAR-100

Matching existing implementation

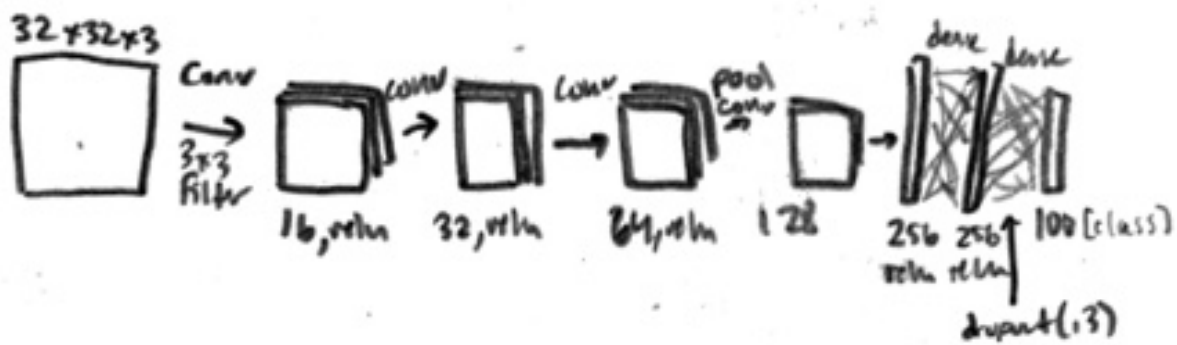
1. "Learning Multiple Layers of Features from Tiny Images" (Krizhevsky 2009): **Whitening**



Assignment 1: CIFAR-100

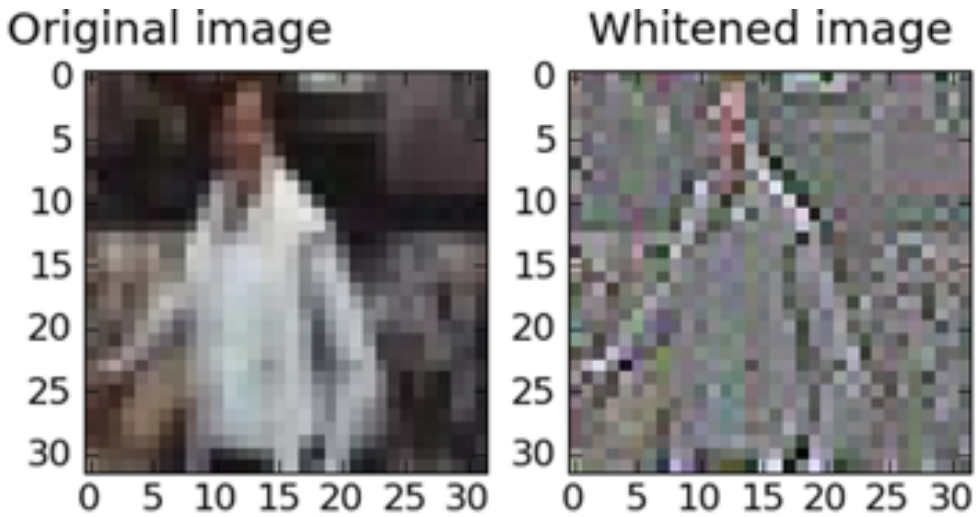
Matching existing implementation

1. "Learning Multiple Layers of Features from Tiny Images" (Krizhevsky 2009): **Whitening**
2. Architecture from Geeks4Geeks, Assignment 0



Methods: Whitening

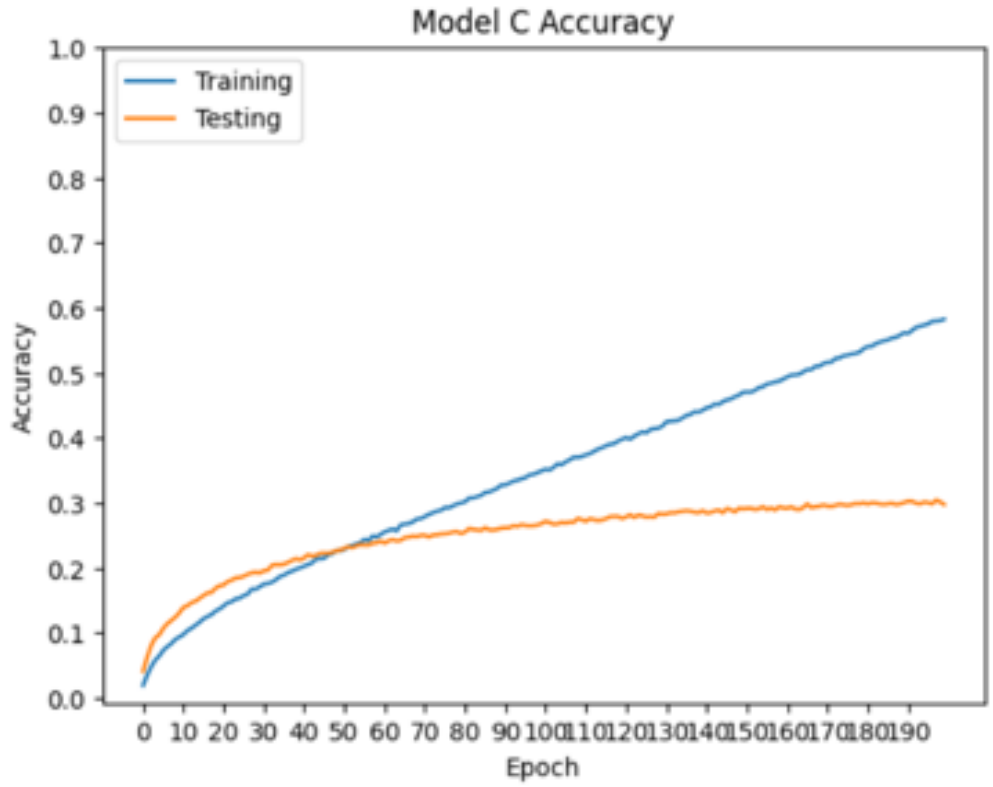
Prevents the model from focusing on irrelevant information (Wiesler et al. 2014).



Credit: Alex Krizhevsky

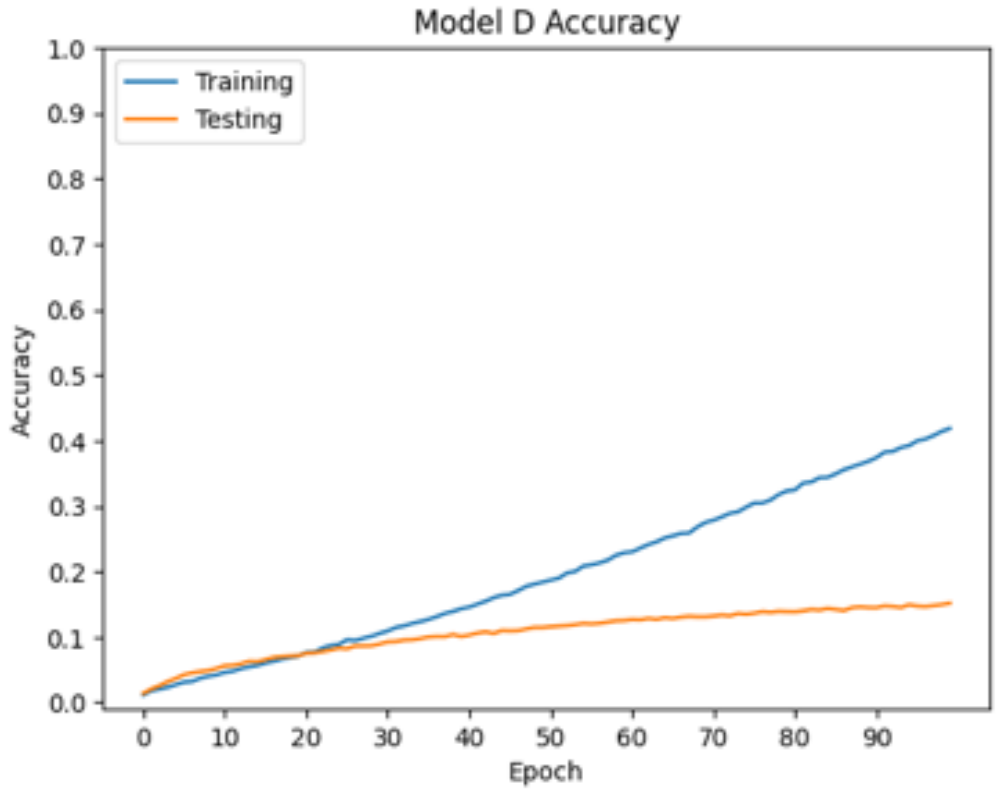
Experiments: Model C (Geeks4Geeks, Original)

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



Experiments: Model D (Geeks4Geeks, Whitened)

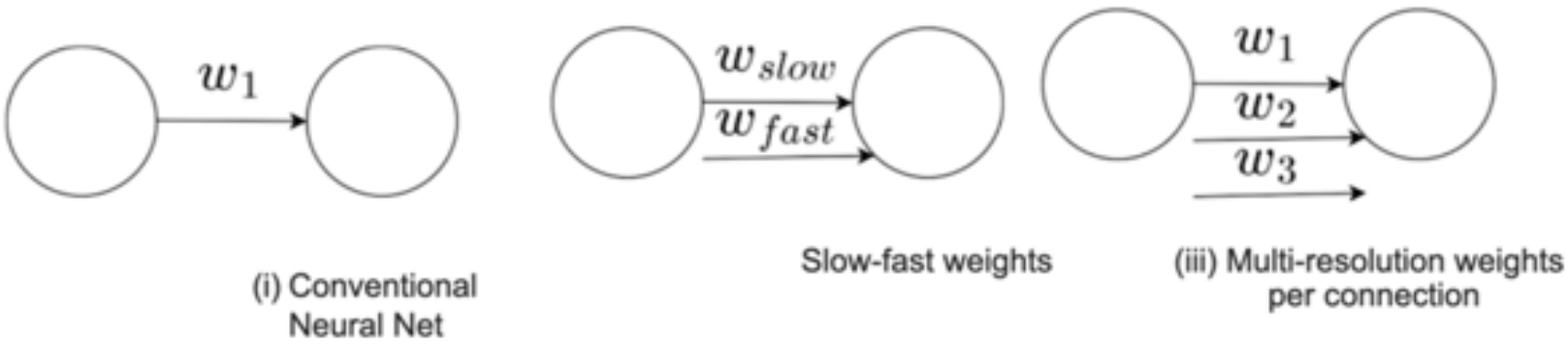
200 epochs, 32 batch size — accuracy: 15.2% (whitening hurt performance)



My Project

Slow-Fast Weights for Long-Term Video Understanding

Rajat Modi, Dr. Rawat



My Project: This Week

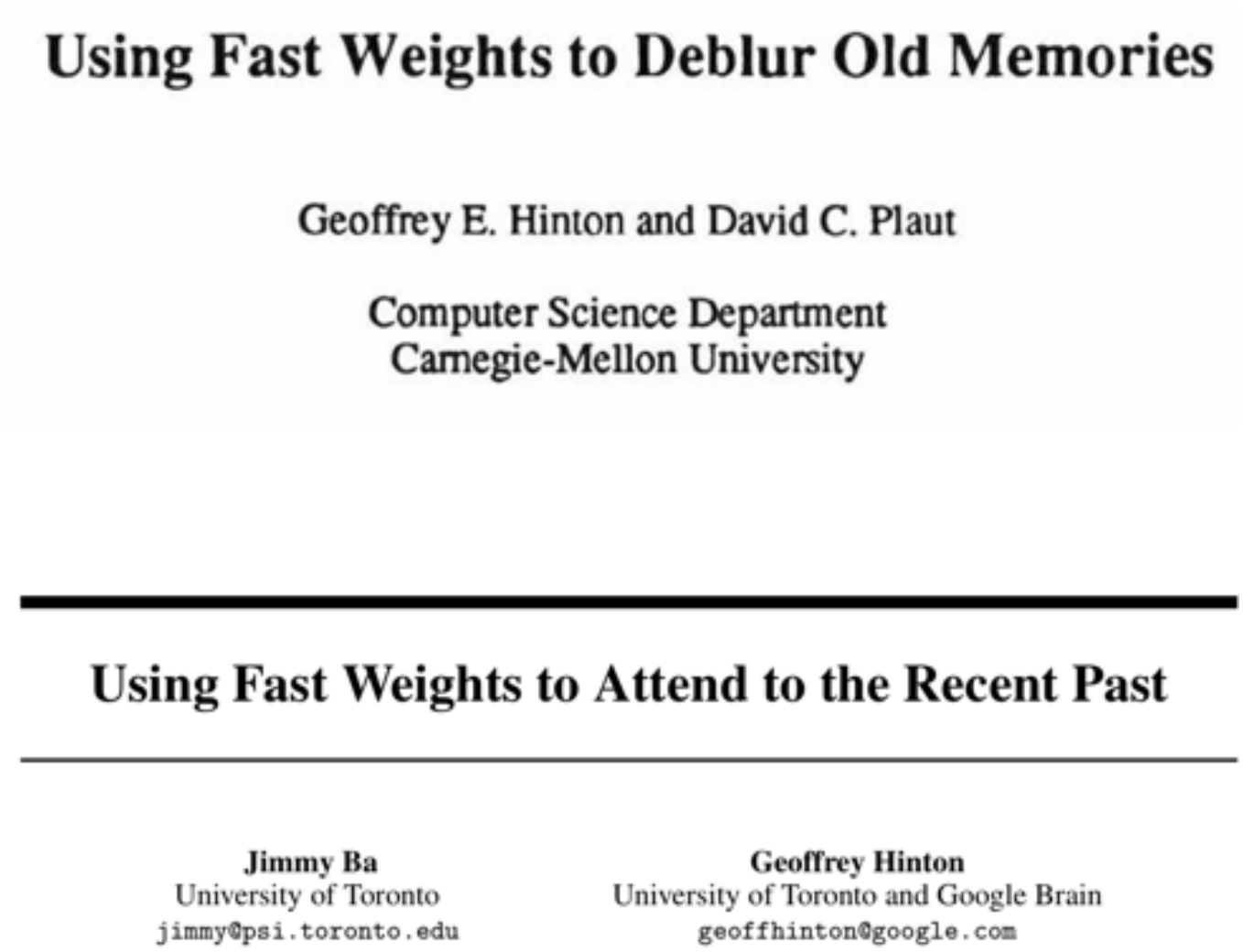
Three major works:

1. "Using fast weights to store temporary memories"



My Project: This Week

Three major works:



Using Fast Weights to Deblur Old Memories

Geoffrey E. Hinton and David C. Plaut

Computer Science Department
Carnegie-Mellon University

Using Fast Weights to Attend to the Recent Past

Jimmy Ba
University of Toronto
jimmy@psi.toronto.edu

Geoffrey Hinton
University of Toronto and Google Brain
geoffhinton@google.com

Bibliography

- Ba, Jimmy et al. (2016). "Using Fast Weights to Attend to the Recent Past". In: *Advances in Neural Information Processing Systems* 29.
- Hinton, Geoffrey (2023). *The Brain as a Bayesian Inference Engine*. https://video-archive.fields.utoronto.ca/view/6180. Fields Institute Lecture, October 25, 2023.
- Hinton, Geoffrey E. and David C. Plaut (1987). "Using Fast Weights to Deblur Old Memories". In: *Proceedings of the Ninth Annual Conference of the Cognitive Science Society*, pp. 177–186.
- 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*