

Abhay Kumar Yadav

CONTACT INFORMATION

4112, Brendan Iribe Center
Dept. of Computer Science
The University of Maryland - College Park

jaiabhay@cs.umd.edu
www.umiacs.umd.edu/~jaiabhay
Google Scholar Profile

RESEARCH INTERESTS

I mainly work on the intersection of **Deep Learning, Computer Vision and Optimization.**

- **Computer Vision:** Image classification, Optimization.
- **Machine Learning:** Efficient and automated training methods for deep networks, Generative Adversarial Networks; Adversarial Training.

EDUCATION

The University of Maryland, College Park, MD Apr. 2017 - present

- Ph.D., Computer Science, *Expected:* Fall 2021
- Advisor: Prof. David W. Jacobs
- GPA: 4.0/4.0

The University of Maryland, College Park, MD Sep. 2014 - Apr. 2017

- M.S., Computer Science
- Advisor: Prof. David W. Jacobs
- GPA: 4.0/4.0

Indian Institute of Science, Bangalore, India Aug. 2007 - Jun. 2009

- M.E., Systems Science And Automation
- Advisors: Ambedkar Dukkipati and M. Narasimha Murty
- GPA: 7.2/8.0 (**First Class with Distinction**)

PUBLICATIONS

1. Amnon Geifman, **Abhay Kumar Yadav**, Yoni Kasten, Meirav Galun, David Jacobs, and Basri Ronen, "On the Similarity between the Laplace and Neural Tangent Kernels", *34th Conference on Neural Information Processing Systems (NeurIPS)*, 2020.
2. **Abhay Kumar Yadav**, Tom Goldstein and David W. Jacobs, "Making L-BFGS Work with Industrial-Strength Nets", *The British Machine Vision Conference (BMVC)*, 2020.
3. **Abhay Kumar Yadav**^{*}, Sohil Shah^{*}, Zheng Xu, David W. Jacobs, and Tom Goldstein, "Stabilizing Adversarial Nets with Prediction Methods", *International Conference on Learning Representations (ICLR)*, 2018. (* equal contribution)
4. Soham De, **Abhay Kumar Yadav**, David Jacobs, and Tom Goldstein, "Automated Inference with Adaptive Batches", *The 20th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2017. (Oral)
5. Sohil Shah^{*}, **Abhay Kumar Yadav**^{*}, Carlos D. Castillo, David W. Jacobs, Christoph Studer, and Tom Goldstein, "Biconvex Relaxation for Semidefinite Programming in Computer Vision", *European Conference on Computer Vision (ECCV)*, 2016. (* equal contribution)
6. **Abhay Kumar Yadav**, R. Ranjan, U. Mahbub and M.C. Rotkowitz, "New Methods for Handling Binary Constraints", *Proceedings of the 54th Annual Allerton Conference on Communication, Control, and Computing*, 2016. (Oral)
7. Ambedkar Dukkipati, **Abhay Kumar Yadav**, M. Narasimha Murty "Maximum Entropy Model Based Classification with Feature Selection", *International Conference on Pattern Recognition (ICPR)*, 2010.

ONGOING
PROJECTS

- **Mitigating Catastrophic Overfitting in Single-step Adversarial Training:**
 - Single-step adversarial training is fast and efficient but susceptible to catastrophic overfitting (unstable training). Many existing works try to address this problem but are computationally expensive and need more memory (**For large-scale datasets like ImageNet, it can reduce the training time from 75 days to 5 days on 4 GPUs**).
 - We learn the class-wise perturbation distributions while training without incurring any additional computation or memory overhead.
 - We are able to match the single-step adversarial accuracy with PGD adversarial training on ImageNet.

SELECTED
PROJECTS

- **Efficient Training for Neural Tangent Kernels:**
 - We show that Neural Tangent Kernel (NTK) has the same set of functions as Laplace kernel (both empirically and theoretically).
 - No need of recursively computing NTK, saving computation and memory for large-scale data.
 - Achieved state-of-the-art results for well-known 102 UCI data sets using a more general gamma-exponential kernel.
- **Making L-BFGS Work for Deep Neural Networks with BatchNorm:**
 - Proposed an L-BFGS method that is more suitable for deep networks with BatchNorm.
 - Beats the existing L-BFGS approaches by a large margin (by around 10% absolute).
 - For some problems, achieved results on par with SGD (manually tuned state-of-the-art results) but with automated line search and preconditioning.
 - Another stepping stone for the practical use of second-order methods to train deep networks.
- **Stable Training for Generative Adversarial Networks (GANs):**
 - Proposed a simple modification of SGD that stabilizes adversarial networks, including GANs.
 - Can enable faster training with large learning rates.
 - Very easy to implement, we discuss both theoretical and empirical results.
 - Achieved better results than the competitive methods with 5x fewer computations.
- **Towards Automated Training of Deep Networks:**
 - Proposed a method that uses automated learning rate and adaptively grows batch size based on the variance in the gradients.
 - Experiments for convex and deep networks achieve performance comparable or better than classical SGD but with less user oversight and more predictability.
 - We also provide some theoretical guarantees.
- **Large Scale Low Rank SDPs in Computer Vision:**
 - Proposed a new and fast method to solve SDPs with PSD constraints.
 - Achieved 4x to 35x faster results with matching or better accuracy than the competitive methods for significant computer vision problems, including image segmentation, metric learning, etc.

RESEARCH
EXPERIENCE

- | | |
|---|-----------------------|
| Intern at Intel Labs | June 2018 -Dec 2018 |
| <ul style="list-style-type: none">• Mentor: Vladlen Koltun• Topic: Second order methods for training deep networks | |
| Research Assistant at University of Maryland, College Park | Sep. 2016 - present |
| <ul style="list-style-type: none">• Mentors: Dr. Tom Goldstein and Prof. David W. Jacobs• Topic: Second order methods for training deep networks | |
| Research Assistant at Indian Institute of Science, Bangalore, India | Aug. 2007 - Jun. 2009 |
| <ul style="list-style-type: none">• Advisors: Ambedkar Dukkipati and M. Narasimha Murty• Topic: Information theoretic approach to data clustering and classification | |

INDUSTRY EXPERIENCE	Senior Software Developer at Server Technologies group Oracle July 2009 to June 2014 <ul style="list-style-type: none"> • Responsibilities: End to end ownership of the projects. Mentoring and training of new hires, interns and contractors. • Product Description: OWSM is a security and monitoring product for Web services. It is based on Java/J2EE technologies which includes EJB, MBeans etc.
REFEREE	Journals: Transactions on Pattern Analysis and Machine Intelligence Conferences: NeurIPS (formerly NIPS), ECCV, ICCV, CVPR, AISTATS, BMVC
TEACHING EXPERIENCE	University of Maryland, College Park <ul style="list-style-type: none"> • Teaching Assistant, Object-Oriented Programming I Fall 2014 • Teaching Assistant, Object-Oriented Programming II Spring 2015
AWARDS	Scholastic Achievements <ul style="list-style-type: none"> • Future Faculty Fellow, University of Maryland, College Park 2018 • Dean's Fellowship, University of Maryland, College Park 2014 • GATE scholarship, IISc Bangalore 2007 • All India Rank 1st in Graduate Aptitude Test in Engineering 2007
RELEVANT COURSEWORKS	Deep Learning, Computer Processing of Pictorial Information, Convex Optimization, Sparsity and Machine Learning, Statistical Pattern Recognition, Computational linguistics
SKILLS	Main Developing Language: Python Other Languages: MATLAB, C, Java, J2EE, R, Lua Tools:Pytorch, Torch
REFERENCES	Available on request