Ameya Prabhu

https://ameya.prabhu.be
github.com/drimpossible
ameya@prabhu.be

Research Interests

- Continual Learning
- Model Editing
- Never-Ending Learning (NEL)
- Continual Dataset Curation
- Scalable Active Learning

Past Interests

- Heirarchical Classification
- Multi-Object Tracking
- Network Pruning at Init
- Binary Deep Networks
- Large Language Modelling
- Codemixed NLP
- Knowledge Graphs
- Efficient Multimodal Representation Learning

Selected Courses

- Project: Modeling Deep Networks as Graphs
- Project: Differential Privacy meets Machine Learning
- Machine Learning (Grad)
- Computer Vision (Grad)
- Advances in Natural Language Processing
- Digital Image Processing
- Computational Linguistics
- Advanced Cryptography
- Functional Programming
- Advanced Data Structures
- Advanced Algorithms

Education

D. Phil. in Engineering Science Oct 2020 - Present
Advisors: Philip Torr and Adel Bibi, Torr Vision Group (TVG), University of Oxford

MS by Research in Computer Science Aug 2018 - Aug 2019
Advisor: Anoop Namboodiri, Center for Visual Information Technology (CVIT, IIIT-H).

B. Tech. (Honors) in Computer Science Aug 2014 - Aug 2018
International Institute of Information Technology, Hyderabad (IIIT-H), India

D.Phil. Thesis: Compute Efficient Continual Learning on Datastreams from Internet

The aim of my thesis is enable neural networks to continually align classification models with feedback and new data encountered after deployment in the real world. Our focus is on learning efficiency, as repeated training becomes prohibitively expensive at scale, by developing efficient kNN augmented classifiers. We design components of a never-ending image learning system by downloading and automatically annotating images from internet searches using foundational models. Parts of this project got adopted by Meta AI.

MS Thesis: Compact and Compute Efficient Deep Neural Networks

I explored simple but effective methods to reduce space and compute costs in deep models at deployment using pruning and binarization (weights and activations quantized to ± 1). I asked the question “where to binarize" which gives large performance boosts to binary models with little tradeoff in compute costs. I additionally demonstrated the power of random pruning beating state-of-the-art pruning methods, with theoretical insights with expander graphs. Parts of this project got adopted by Texas Instruments.

Selected Publications

*Ian McKenzie et. al. (inc **Ameya Prabhu**). Inverse Scaling: When Bigger Isn't Better. TMLR.*
We study the evidence supporting the existence of inverse scaling in Language Models (LMs), indicating decreased task performance with increased scale. We identify four potential reasons for this phenomenon: (i) favoring memorized sequences over in-context instructions, (ii) imitation of undesirable patterns from the training data, (iii) distraction by an easier task present in the actual task, and (iv) accurate but misleading few-shot demonstrations.

Hasan Hammoud, **Ameya Prabhu***, Ser-Nam Lim, Philip Torr, Adel Bibi, Bernard Ghanem. Rapid Adaptation in Online Continual Learning: Are We Evaluating It Right? ICCV 2023.*
We discover that the online accuracy metric in CL literature can be misleading due to spurious label correlations. We show methods can simply be overfitting to the latest samples, achieve high online accuracy despite not learning the intended task. To address this, we propose to evaluate accuracy on near-future samples, which removes spurious correlations.

***Ameya Prabhu**, Zhipeng Cai, Puneet Dokania, Philip Torr, Vladlen Koltun, Ozan Sener. Online Continual Learning Without the Storage Constraint.*
We reformulate OCL with relaxed storage constraints and a fixed, limited economical budget. We propose a simple algorithm that efficiently utilizes incoming data using a kNN classifier and universal pretrained feature extractors, ensuring consistency by never forgetting past data, vastly outperforming existing OCL methods on two large-scale OCL datasets.

***Ameya Prabhu**, Hasan Hammoud*, Puneet Dokania, Philip Torr, Ser-Nam Lim, Bernard Ghanem, Adel Bibi. Compute Budgeted Continual Learning: What Does Matter? CVPR 2023.*
We explore Continual Learning when computate/time are constrained rather than storage. We conduct extensive experiments designing a large-scale benchmark and find that traditional CL approaches fail to outperform a simple finetuning baseline, indicating that most existing compute-intensive CL methods may not be practical for real-world scenarios.

Shashwat Goel, **Ameya Prabhu***, Amartya Sanyal, Ser-Nam Lim, Philip Torr, Ponnurangam Kumaraguru. Towards Adversarial Evaluations for Inexact Machine Unlearning.*
We demonstrate the necessity of adversarial evaluation strategies for analyzing unlearning algorithms. We propose a simple, black-box evaluation called the Interclass Confusion (IC) test and two baseline methods (EU-k and CF-k), outperforming popular unlearning methods.

Awards

Outstanding Reviewer
CVPR 2021
CVPR 2020
CVPR2019

*Aurobindo Munagala, Sidhant Subramanian, Shyamgopal Karthik, **Ameya Prabhu**, Anoop Namboodiri. CLActive: Episodic Memories for Rapid Active Learning. CoLLAs 2022.*

We highlight an underexplored problem in deep active learning: Reducing time for repeated training of a model for each round of active selection. We propose to simply learn a small proxy model continually, bringing significant speedups with retaining past information.

---

*Shyamgopal Karthik, **Ameya Prabhu**, Puneet Dokania, Vineet Gandhi. No Cost Likelihood Manipulation at Test Time for Making Better Mistakes in Deep Networks. ICLR 2021.*

This paper introduces a parameter-free capability to amend mistakes by leveraging class hierarchies at inference time by CRM, yielding drastically lower top-k hierarchical error. We point out critical issues in average mistakes metric, which has mislead this subfield, causing the state-of-the-art hierarchy-aware methods being worse than a CrossEntropy baseline.

***Ameya Prabhu**, Philip Torr, Puneet Dokania. GDumb: A Simple Approach that Questions Our Progress in Continual Learning. ECCV 2020 (Oral).*

We introduce a simple baseline GDumb, which shows that by simply restarting from the initialization can still obtains state-of-the-art performance on diverse benchmarks. This shows that the multitude of complex CL algorithms do not really leverage past seen samples, effectively forgetting all data not stored in replay buffer. This raises concerns about progress in Continual Learning, and the need for new settings.

---

*Shyamgopal Karthik, **Ameya P.**, Vineet Gandhi. Simple unsupervised multi-object tracking.*

We remove the need for annotated tracking datasets by developing an unsupervised re-identification deep model. Given unlabeled videos, our simple baseline generates tracking labels using SORT and trains a ReID network to predict the generated labels using crossentropy loss. We demonstrate this performs comparable to its supervised counterpart consistently across diverse tracking frameworks, with limited scope for further improvement.

---

*Aurobindo Munagala, **Ameya Prabhu** and Anoop Namboodiri. STQ-Nets: Unifying Network Binarization and Structured Pruning. BMVC 2020.*

We show that weights/activations closer to zero have higher binarization error making them good pruning candidates. Combining these two complementary paradigms helps us obtain highly compact deep models with little accuracy loss across several benchmarks.

Ameya Prabhu, Charles Dognin and Maneesh Singh. Sampling Bias in Deep Active Classification: An Empirical Study. EMNLP 2019.*

We show that the assumptions behind recent trend towards diversity-based active learning do not hold in large datasets: Datasets we tested had negligible class, feature correlations. Simple uncertianty sampling was robust to critical algorithmic factors and Samples collected by it correlated with supports of a SVM, and can effectively generate compact surrogate datasets (5x-40x compression). Work was adopted by Verisk Analytics.

Ameya Prabhu, Girish Varma*, Anoop Namboodiri. Deep Expander Networks: Efficient Deep Networks from Graph Theory. ECCV 2018 (Oral)*

We introduce a capability of pruning networks at intialization without using data by using intelligent inductive biases. We leverage Expander Graphs from TCS to develop sparse but highly connected networks. This property allows randomly pruned networks to beat several state-of-the-art pruning methods, discover better connectivity structures and allow training of larger sparse networks.

***Ameya Prabhu**, Vishal Batchu, Rohit Gajawada, Aurobindo Munagala, Anoop Namboodiri. Hybrid Binary Networks: Optimizing for Accuracy, Efficiency and Memory. WACV 2018 (Oral)*

We investigate the question of *where* to binarize inputs and show that binarizing the right areas in the network could contribute significantly to speed-ups, without damaging the overall accuracy as compared to binary networks (weights and activations ± 1).

Ameya Prabhu, Aditya Joshi*, Manish Shrivastava, Vasudeva Varma. Towards Sub-Word Level Compositions for Sentiment Analysis of Hindi-English Code Mixed Data. COLING 2016.*

We introduced Subword-LSTMs to incorporate linguistic priors in neural network architectures and show that it learns information about sentiment value of important morphemes. We present the important subwords learnt by our model in morpheme-level feature maps.

***Ameya Prabhu**, Riddhiman Dasgupta, Anush Sankaran, Srikanth Tamilselvam and Senthil KK Mani. Recommending Deep Networks by Accuracy Prediction for Unknown Datasets*

For unknown classification datasets, choosing a base deep learning architecture is often time-taking and laborious process. We propose a novel technique to recommend suitable architecture from a repository of models. Further, we also predict the performance accuracy of the recommended architecture on the given query dataset, without training the model. Work got incorporated into IBM Watson. Awarded best intern poster for outstanding research done in the internship.

---

*Koustav Ghosal, **Ameya Prabhu**, Riddhiman Dasgupta, Anoop Namboodiri. Learning Clustered Subspaces for Sketch Based Image Retrieval. ACPR 2015 (Oral).*

We conjectured that sketches and images belong to different subspaces and obtain a cross-modal correspondence between the two. We use Cluster-CCA to project them onto a correlated lower dimensional subspace, for performing semantic-multimodal retrieval.

Work Experience

ML Intern, Intel AI, Germany Under: Ozan Sener and Vladlen Koltun **Sep 2021 - Jun 2022**

AI Resident, Verisk Analytics, USA Under: Maneesh Singh **Aug 2018 - Oct 2019**

ML Intern, IBM Research, India Under: Riddhiman Dasgupta **May 2018 - Sep 2018**

Teaching (Assistantships)

I contributed by teaching in tutorial sessions, setting and grading assignments, supervising course projects in the following courses:

- **Computer Vision** *(Spring '18)*: The graduate-level introductory machine learning course (Instructor: Prof. Anoop Namboodiri, CVIT)
- **Foundations of AI and ML** *(Sping '18)*: Introduction to machine learning in an educational initiative by IIIT-H for professionals. (Instructor: Prof. C.V. Jawahar, CVIT)
- **Statistical Methods in AI** *(Monsoon '17)*: The graduate-level introductory machine learning course (Instructor: Prof. Vineet Gandhi, CVIT)
- **Digital Signal Analysis and App.** *(Spring '17)*: The introductory undergraduate signal processing course (Instructor: Prof. Vineet Gandhi, CVIT)
- **Computer Programming** *(Monsoon '16)*: The introductory undergraduate programming course (Instructor: Prof. Anoop Namboodiri, CVIT)
- **Electronics Workshop-1** *(Spring '16)*: The undergraduate freshman year course primarily designing analog circuits with a hands-on component involving handling lab equipment. (Instructor: Prof. Madhava Krishna, RRC)
- **Digital Logic and Processor Design** *(Monsoon '15)*: The undergraduate digital processor design course consisting of boolean logic to sequential logic, ALU designs, and programming basic 8085-like processor (Instructor: Prof. Vijay Prakash, RRC)

Programming Projects

- Implemented a distributed banking sytem, and cryptographic protocols using Diffie-Hellman Key Exchange Protocol using Java RMI.
- Implemented various parallel algorithms like sorting, computing MSTs and the games like Game of Life in MPI.
- Implemented a compiler for a subset of the C language using Flex and Bison for parsing, followed by generating ASTs and conversion to LLVM intermediate representations.
- Implemented graph and string processing algorithms along Data Structures like Seg-Trees, AVL-trees, Hash-Maps for Data Structure and Algorithms courses.
- Developed a mobile app security framework with a pipeline consisting of automatic decompilation from .apk to .java code, with static and dynamic code flow analysis along with various signature detection algorithms.
- Redeveloped a basic bash shell in C++.
