Utku Evci

Montreal, QC

February 20, 2023

✉ *utkuevci@gmail.com*   ✈ *utkuevci.com*   ⚲   🎓   in

SELECTED LEAD-AUTHOR PUBLICATIONS

ICML 2022	**Head2Toe: Utilizing Intermediate Features for Better Transfer [1]** , paper / code / video
ICLR 2022	**GradMax: Growing NNs using Gradient Information [2]** , paper / code
AAAI 2022	**Gradient Flow in Sparse NNs and How LTs Win [3]** , Oral (3%) / paper / code / video
ICML 2020	**Rigging the Lottery: Making All Tickets Winners [4]** , paper / code / blog

TALKS AND SERVICE

2021/2022	**Sparsity in Neural Networks Workshop,** lead organizer of the inaugural (and 2nd) workshop which had 200+ live views and 60+ submissions.
2022	**Google Workshop on Sparsity and Adaptive Computation,** How Sparsity?
2022	**Cohere For AI,** On Sparsity and Beyond Static Network Architectures
2022	**MILA Tea Talks,** Beyond Static Network Architectures [recording]
2021	**MLCollective, DLCT Talk Series,** Difficulty of Sparse Training and RigL
2019	**MicroNet Challenge @ Neurips,** co-wrote the evaluation code.
2019-2022	**Reviewer,** ICML 20,21,22 / Neurips 20,21 / ICLR 21,22,23 / JMLR 21,22 / TMLR 22,23

ACHIEVEMENTS

2018	**Google AI Residency,** Selected from over 5k applications ($< 1\%$).
2016	**Fulbright Scholarship & NYU GSAS Tuition Scholarship,** for M.Sc. at NYU.
2011	**Semahat Arsel Scholarship,** most prestigious full scholarship for the B.Sc. at Koc University.
2011	**Ranked 1st in Turkey,** in College Entrance Exam (LYS) out of more than a million people.

EDUCATION

May 2018	**New York University,** Courant Institute	New York, NY
	M.Sc. in Computer Science, GPA:3.95/4	
June 2016	**Koc University,** College of Engineering	Istanbul, Turkey
	B.Sc. in Electrical and Electronics Engineering, GPA: 3.99/4.30, 2nd in class	
	B.Sc. in Computer Engineering, GPA: 4.02/4.30, 2nd in class	

WORK & RESEARCH EXPERIENCE

Present	**Google,** Brain Team	Montreal, Canada

Researcher

• Researching efficient training methods for neural networks. Led research projects on (1) growing neural networks [2] (2) understanding and improving sparse training [3] (3) efficient transfer learning [1, 5] and contributed to a number of other projects on few-shot learning [6, 7].

• Co-created an internal interview series for highlighting research careers and lessons learned.

2018-2020	**Google,** Brain Team	Montreal, Canada

2018 AI Residency Program

• Led two projects on training sparse neural networks. Results on the loss energy landscape of sparse training are presented at ICML 2019 Deep Phenomena Workshop [8]. Later we developed a novel sparse training method for training sparse neural which is published in ICML 2020 [4]. Code open-sourced here.

• Led a project on developing better pruning algorithms that reduces the Δ loss due pruning [9].

• Learned Tensorflow and checked-in 20k+ lines of peer-reviewed code in the first 12 months.

Summer 2017	**Amazon,** AWS EC2	Seattle, United States

Software Development Engineer (SDE) Intern: Auditing Big-Data

• Wrote 3000+ lines of spark/python-code for auditing TBs of data on AWS reaching 50mb/s per node.

Spring 2017	**NYU**, Courant Institute	New York, United States

Research Assistant: 2 different projects
- Published a paper on the geometry of the loss landscape of deep neural networks [10].
- Worked with Alex Rives on predicting protein structure from sequence information.

Summer 2015	**Swiss Federal Institute of Technology (EPFL)**, IIG	Lausanne, Switzerland

Research Intern: Modeling Human Stepping
- Modelled human stepping with neural networks using motion capture data [11].

OTHER PROJECTS

Spring 2018	**Detecting Dead Weights and Units [12]**, Python/Bash

M.Sc. Thesis advised by Prof. Léon Bottou
- Implemented pytorchpruner: pruning library for pyTorch with 1k+ lines of code.
- Wrote exp-bootstrp for managing large scale experiments.

Fall 2015	**Facial Expression Detection**, Matlab/Bash

B.Sc. Graduation Project
- Built a Rasperry-Pi based facial expression detecting art-installation, which is exhibited on campus.
- Created a dataset of facial expressions from 80 students and trained a NN based model.

CODING

> **5000 lines**	C ∘ **Python** ∘ Java ∘ Bash ∘ JAX ∘ Tensorflow ∘ pyTorch
> **1000 lines**	CUDA ∘ Pandas

PUBLICATIONS

[1] Utku Evci, Vincent Dumoulin, H. Larochelle, and Michael Curtis Mozer. Head2Toe: Utilizing Intermediate Representations for Better Transfer Learning. In *Proceedings of the 39th International Conference on Machine Learning*, 2022.

[2] Utku Evci, Max Vladymyrov, Thomas Unterthiner, Bart van Merrienboer, and Fabian Pedregosa. GradMax: Growing Neural Networks using Gradient Information. *ICLR*, 2022.

[3] Utku Evci, Yani Andrew Ioannou, Cem Keskin, and Yann N. Dauphin. Gradient Flow in Sparse Neural Networks and How Lottery Tickets Win. *AAAI*, 2022.

[4] Utku Evci, Trevor Gale, Pablo Samuel Castro Rivadeneira, and Erich Elsen. Rigging The Lottery: Making All Tickets Winners. In *ICML*, 2020.

[5] Laura Graesser, Utku Evci, Erich Elsen, and Pablo Samuel Castro. The State of Sparse Training in Deep Reinforcement Learning. In *Proceedings of the 39th International Conference on Machine Learning*, 2022.

[6] Eleni Triantafillou, Tyler Zhu, Vincent Dumoulin, Pascal Lamblin, Utku Evci, Kelvin Xu, Ross Goroshin, Carles Gelada, Kevin Swersky, Pierre-Antoine Manzagol, and Hugo Larochelle. Meta-Dataset: A Dataset of Datasets for Learning to Learn from Few Examples. In *ICLR*, 2020.

[7] Vincent Dumoulin, Neil Houlsby, Utku Evci, Xiaohua Zhai, Ross Goroshin, Sylvain Gelly, and Hugo Larochelle. Comparing Transfer and Meta Learning Approaches on a Unified Few-Shot Classification Benchmark. In *Neurips Datasets and Benchmarks Track*, 2021.

[8] Utku Evci, Fabian Pedregosa, Aidan N. Gomez, and Erich Elsen. The Difficulty of Training Sparse Neural Networks. In *ICML Workshop Deep Phenomena*, 2019.

[9] Utku Evci, Nicolas Le Roux, Pablo Castro, and Léon Bottou. Mean Replacement Pruning. *Openreview*, 2018.

[10] Levent Sagun, Utku Evci, V. Ugur Güney, Yann Dauphin, and Léon Bottou. Empirical Analysis of the Hessian of Over-Parameterized Neural Networks. In *ICLR Workshop Track*, 2018.

[11] Ronan Boulic, Utku Evci, Eray Molla, and Phanindra Pisupati. One Step from the Locomotion to the Stepping Pattern. In *Proceedings of the 29th International Conference on Computer Animation and Social Agents*, 2016.

[12] Utku Evci. Detecting Dead Weights and Units in Neural Networks. *arXiv*, 2018.