Skanda Koppula

Kingly Building, 18 Woodberry Down
London, UK N42GQ

me@skoppula.com
skoppula.github.io
+1-412-259-3123

Work Experience

Google DeepMind, Research Engineer *September 2018 - Present*

- Pushing forward research on large-scale visual representation learning, self-supervised learning objectives, and new architectures for dense visual tasks such as object detection, segmentation, and optical flow.

NVIDIA Autonomy, PilotNet Team, Intern *May 2018 - September 2018*

- Added support for multi-camera vehicle feeds on the PilotNet training and simulation pipelines.
- Decreased training time by 16% using custom data format to reduce disk reads and CRC overhead.

Google Acoustic Modeling Team, Intern *June 2017 - September 2017*

- Developed approach to characterize and visualize the memory contents and hidden state of recurrent neural networks used in character language models. Work published in ICASSP 2018.
- Tested kernel factorization as a way to improve cross-domain accuracy of neural language models.

Yahoo Login Team, Intern *June 2016 - August 2016*

- Improved malicious login attempt detection by 7% using low-latency, quantized neural networks.
- Implemented online updating of the login classifier by pulling live data from external feed.

Square Security, Intern *June 2015 - Aug. 2015*

- Built an in-production service to collect memory dumps from Square card readers.
- Simplified firmware bug-fixing by parsing dumps into a human readable source error trace.

Projects

ETH Zürich, Fulbright Research Fellowship, Researcher *Sept. 2018 - June 2019*

- Developed CNN fine-tuning procedure to improve bit error resilience. Paired with approximate DRAM modules (voltage and timing scaled), this yielded a 29% power reduction and 5% speed-up while running inference on state-of-art CNNs on CPU, GPU, and state-of-art neural accelerators.

MIT Energy Efficient Systems Lab, Master's Thesis *Sept. 2016 - 2018*

- Achieved a >2x speed-up and >500x decrease in estimated energy consumption for real-time speaker identification by demonstrating a quantized speaker ID neural network and an FPGA design capable of neural network inferencing and mel-frequency coefficient extraction.

MIT Formula Student Driverless, Team Co-Founder *September 2017 - March 2019*

- Helped hire and manage new recruits for a team that built the entire autonomy stack of a Formula Student racecar. Closely directed technical efforts on the perception team, demonstrating an accurate LiDAR-stereo camera system for racetrack landmark localization.
- Won 2nd and 3rd place at the 2019 Formula Student Driverless Italy and Germany competitions.

MIT Formula SAE Racecar Team, Member *September 2015 - 2017*

- Developed PCBs and firmware for an open-source automotive battery management system. Orchestrated battery dis/charging state, balancing, safety checks, and charging algorithms.
- Won 2nd place at the 2017 U.S. Formula Student Electric competition (among 23 teams).

Education

Massachusetts Institute of Technology

Masters of Engineering and BSc, Computer Systems, MEng GPA: 5.0/5.0 *Sept. 2013 - June 2018*

Relevant courses: Digital Architectures for Deep Learning, Applied Cryptography, Compilers, Machine Learning, Operating Systems, Computer Networks and Security, Bayesian Inference

Skills

Software: **Python** (including **Tensorflow**, **PyTorch**, **Jax**, **numpy**), **Java**, **Scala**, **C++**, **bash**
Embedded Systems and Digital Design: **C**, **Altium**, **Vivado HLS**, and **Bluespec Verilog**.

Research Publications

Perceiver IO: A General Architecture for Structured Inputs & Outputs.
International Conference on Learning Representations (ICLR) 2022.
Andrew Jaegle, Sebastian Borgeaud, Jean-Baptiste Alayrac, Carl Doersch, Catalin Ionescu, David Ding, **Skanda Koppula**, Daniel Zoran, Andrew Brock, Evan Shelhamer, Olivier Hénaff, Matthew M. Botvinick, Andrew Zisserman, Oriol Vinyals, João Carreira.

Efficient Visual Pretraining with Contrastive Detection.
International Conference on Computer Vision (ICCV 2021), Oral.
Olivier J. Henaff, **Skanda Koppula**, Jean-Baptiste Alayrac, Aaron van den Oord, Oriol Vinyals, Joao Carreira

A Deep Learning Approach for Characterizing Major Galaxy Mergers.
NeurIPS 2020 Workshop: ML for Physical Science (NeurIPS ML4PS 2020).
Skanda Koppula, Victor Bapst, Marc Huertas-Company, Sam Blackwell, Agnieszka Grabska-Barwinska, Sander Dieleman, Andrea Huber, Natasha Antropova, Mikolaj Binkowski, Hannah Openshaw, Adria Recasens, Fernando Caro, Avishai Dekel, Yohan Dubois, Jesus Vega Ferrero, David C. Koo1, Joel R. Primack, Trevor Back

Accurate, Low-Latency Visual Perception for Autonomous Racing: Challenges, Mechanisms, and Practical Solutions.
The IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2020).
Kieran Strobel, Sibo Zhu, Raphael Chang, and **Skanda Koppula**

EDEN: Enabling Energy-Efficient, High-Performance Deep Neural Network Inference Using Approximate DRAM.
The 52nd Annual IEEE/ACM International Symposium on Microarchitecture (MICRO 2019).
Skanda Koppula, Lois Orosa, A. Giray Yaglicki, Roknoddin Azizi, Taha Shahroodi, Konstantinos Kanellopoulos, and Onur Mutlu

Understanding Recurrent Neural State Using Memory Signatures.
2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2018).
Skanda Koppula, Khe Chai Sim, and Kean Chin

Energy-Efficient Speaker Identification With Low-Precision Networks.
2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2018)..
Skanda Koppula, James Glass, and Anantha P. Chandrakasan