dbolya@gatech.edu
scholar

Daniel Bolya

ML Research Scientist @ Meta

dbolya.github.io
dbolya

Research Interests

I'm particularly interested in making the state-of-the-art in computer vision more efficient. My goal is to avoid common strategies such as pruning and quantization and instead search for orthogonal methods to increase efficiency—with the hope that these methods can all be used together for a multiplicative effect.

Technical Experience

Languages	Python, C++, Java, JavaScript, Lua
Frameworks	PyTorch, TensorFlow, NumPy
Areas	Vision, Diffusion, Detection, Efficiency, Self Supervision
Skills	Pushing SotA, Training Large Models (e.g., 1024 GPUs)

Publications

2025 ***Perception encoder: The best visual embeddings are not at the output of the network*** — Preprint
Daniel Bolya*, Po-Yao Huang*, Peize Sun*, Jang Hyun Cho*, Andrea Madotto*, Chen Wei, Tengyu Ma, Jiale Zhi, Jathushan Rajasegaran, Hanoona Rasheed, Junke Wang, Marco Monteiro, Hu Xu, Shiyu Dong, Nikhila Ravi, Daniel Li, Piotr Dollár, Christoph Feichtenhofer

PerceptionLM: Open-Access Data and Models for Detailed Visual Understanding — Preprint
Jang Hyun Cho*, Andrea Madotto*, Effrosyni Mavroudi*, Triantafyllos Afouras*, Tushar Nagarajan*, Muhammad Maaz*, Yale Song*, Tengyu Ma*, Shuming Hu*, Suyog Jain, Miguel Martin, Huiyu Wang, Hanoona Rasheed, Peize Sun, Po-Yao Huang, **Daniel Bolya**, Nikhila Ravi, Shashank Jain, Tammy Stark, Shane Moon, Babak Damavandi, Vivian Lee, Andrew Westbury, Salman Khan, Philipp Krähenbühl, Piotr Dollár, Lorenzo Torresani, Kristen Grauman, Christoph Feichtenhofer

Gaze-LLE: Gaze Target Estimation via Large-Scale Learned Encoders — Highlight CVPR 2025
Fiona Ryan, Ajay Bati, Sangmin Lee, **Daniel Bolya**, Judy Hoffman, James M Rehg

2024 ***Window Attention is Bugged: How not to Interpolate Position Embeddings*** — ICLR 2024
Daniel Bolya, Chaitanya Ryali, Judy Hoffman, Christoph Feichtenhofer

ZipIt! Merging Models from Different Tasks without Training — ICLR 2024
George Stoica*, **Daniel Bolya***, Jakob Bjorner, Taylor Hearn, Judy Hoffman

2023 ***Hiera: A Hierarchical Vision Transformer without the Bells-and-Whistles*** — Oral ICML 2023
Chaitanya Ryali*, Yuan-Ting Hu*, **Daniel Bolya***, Chen Wei, Haoqi Fan, Po-Yao Huang, Vaibhav Aggarwal, Arkabandhu Chowdhury, Omid Poursaeed, Judy Hoffman, Jitendra Malik, Yanghao Li*, Christoph Feichtenhofer*

Token Merging for Fast Stable Diffusion — Oral CVPR 2023 (ECV Workshop)
Daniel Bolya, Judy Hoffman

Token Merging: Your ViT But Faster — Oral Top 5% ICLR 2023
Daniel Bolya, Cheng-Yang Fu, Xiaoliang Dai, Peizhao Zhang, Christoph Feichtenhofer, Judy Hoffman

2022 ***Hydra attention: Efficient attention with many heads*** — Best Paper ECCV 2022 (CADL Workshop)
Daniel Bolya, Cheng-Yang Fu, Xiaoliang Dai, Peizhao Zhang, Judy Hoffman

2021 ***Scalable diverse model selection for accessible transfer learning*** — NeurIPS 2021
Daniel Bolya*, Rohit Mittapalli*, Judy Hoffman

2020 ***TIDE: A General Toolbox for Identifying Object Detection Errors*** — Spotlight ECCV 2020
Daniel Bolya, Sean Foley, James Hays, Judy Hoffman

Likelihood Landscapes: A Unifying Principle Behind Many Adversarial Defenses — Oral ECCV 2020 (AROW Workshop)
Fu Lin, Rohit Mittapalli, Prithvijit Chattopadhyay, **Daniel Bolya**, Judy Hoffman

YOLACT++: Better Real-time Instance Segmentation — TPAMI 2020
Daniel Bolya*, Chong Zhou*, Fanyi Xiao, Yong Jae Lee

2019 ***YOLACT: Real-time Instance Segmentation*** — Oral ICCV 2019
Daniel Bolya, Chong Zhou, Fanyi Xiao, Yong Jae Lee

2016 ***Using Artificial Intelligence Systems for Autonomous Visual Comprehension and Handwriting Generation*** — ISEF 2016
Daniel Bolya*, Dylan McLeod*

Education

May 2024 – August 2019 — **Machine Learning**
PH.D. · Georgia Institute of Technology
Advised by Judy Hoffman.

March 2019 – September 2016 — **Computer Science**
B.S. · University of California Davis
Math Minor. Research advised by Yong Jae Lee.

Awards

2024	**GaTech College of Computing Outstanding Graduate Research Assistant Award**	
2022	**Best Paper Award** (ECCV 2022 CADL Workshop)	*Hydra Attention: Efficient Attention with Many Heads*
2020	**National Science Foundation Graduate Research Fellowship**	
2020	**Best Paper Runner-Up** (ECCV 2020 AROW Workshop)	*Likelihood Landscapes: A Unifiying Principle...*
2019	**COCO Challenge Most Innovative Award**	*YOLACT: Real-Time Instance Segmentation*
2019	**Chancellor's Award for Excellence in Undergraduate Research Honorable Mention**	*YOLACT: Real-Time Instance Segmentation*
2017	**HackDavis Honorable Mention**	*Proton: Positivity Generator*
2016	**Intel International Science and Engineering Fair (ISEF) Finalist**	*Handwritten Math Equation Solver*
2016	**Sacramento STEM Fair 1st Place Category Award in Math and CS**	*Handwritten Math Equation Solver*
2016	**Sacramento STEM Fair 3rd Place Grand Prize Award**	*Handwritten Math Equation Solver*
2016	**California State Science Fair Honorable Mention**	*Handwritten Math Equation Solver*
2016	**Intel Excellence in Computer Science**	*Handwritten Math Equation Solver*

Employment

Ongoing – June 2024
Research Scientist (FAIR)
Meta · New York, New York
Pushing the boundaries of Computer Vision.

May 2024 – August 2023
Graduate Research Assistant
Georgia Tech · Atlanta, Georgia
Advised by Judy Hoffman.

August 2023 – May 2020
NSF Graduate Research Fellow
Georgia Tech · Atlanta, Georgia
Worked on PARC, ToMe for SD, and ZipIt! among others. Advised by Judy Hoffman.

August 2023 – May 2023
Research Scientist Intern (FAIR)
Meta · San Francisco, California
Released Hiera and worked to further push the state-of-the-art with Hiera under Christoph Feichtenhofer.

August 2022 – May 2022
Research Scientist Intern (Meta AI)
Meta · San Francisco, California
Worked on and released Hydra Attention and Token Merging under Cheng-Yang Fu.

August 2021 – May 2021
Research Scientist Intern (FAIAR)
Meta · Remote
Worked on grounded unsupervised part segmentation under Vignesh Ramanathan.

May 2020 – August 2019
Graduate Research Assistant
Georgia Tech · Atlanta, Georgia
Developed, released, and supported TIDE. Advised by Judy Hoffman.

August 2019 – April 2019
Research Assistant
UC Davis · Davis, California
Released and supported YOLACT, as well as prepared for YOLACT++. Advised by Yong Jae Lee.

March 2019 – June 2018
Undergraduate Student Researcher
UC Davis · Davis, California
Developed YOLACT, the first real-time instance segmentation method. Advised by Yong Jae Lee.

Academic Talks

June 2025	Invited	**Large Scale Holistic Video Understanding** (CVPR)	*Perception Encoder: SotA Image-Video CLIP*
June 2025	Invited	**Efficient Large Vision Models** (CVPR)	*Perception Encoder: SotA Image-Video CLIP*
June 2025	Invited	**Pixel-level Video Understanding in the Wild** (CVPR)	*Perception Encoder: SotA Image-Video CLIP*
June 2025	Invited	**Multimodal Models Forum** (BAAI)	*Perception Encoder: SotA Image-Video CLIP*
April 2024	Lecture	**Georgia Tech CS 6476 Advanced Computer Vision**	*Accelerating Vision by Eliminating Redundancy*
Nov 2023	Invited	**Meta FAIR**	*Accelerating Vision by Eliminating Redundancy*
Sep 2023	Invited	**Georgia Tech AI Synapse**	*Accelerating Vision by Eliminating Redundancy*

Aug 2023	INVITED	**Runway ML**	*Accelerating Vision by Eliminating Redundancy*
Aug 2023	INVITED	**NVIDIA Research**	*Accelerating Vision by Eliminating Redundancy*
Jun 2023	ORAL	**Efficient Deep Learning for Computer Vision** (CVPR)	*Token Merging for Fast Stable Diffusion*
May 2023	ORAL	**International Conference on Learning Representations**	*Token Merging: Your ViT but Faster*
Oct 2022	ORAL	**Computational Aspects of Deep Learning** (ECCV)	*Hydra Attention: Efficient Attn w/ Many Heads*
Aug 2020	SPOTLIGHT	**European Conference on Computer Vision**	*TIDE: An Object Detection Evaluation Toolkit*
Oct 2019	ORAL	**COCO + Mapillary Joint Recognition Challenge** (ICCV)	*YOLACT: Real-time Instance Segmentation*
Oct 2019	ORAL	**International Conference on Computer Vision**	*YOLACT: Real-time Instance Segmentation*