

Adrien Gaidon

PhD, Senior Manager, Machine Learning Research

1133 Karen Way
Mountain View, CA 94040, USA
✉ adrien.gaidon@gmail.com
📧 adriengaidon.com
Lawful US permanent resident

Professional Experience

- 2020 - **Senior Manager, Machine Learning Research**, *Toyota Research Institute (TRI)*, CA, USA.
Head of the Machine Learning Research team at TRI. Applied ML research for Toyota's future products, including World-scale Autonomous Robots that learn from Experience. Our key research areas include Deep Learning (esp. self-supervised learning), Computer Vision (esp. dynamic 3D scene understanding), and using Simulation for ML (esp. differentiable rendering and sim2real).
- 2017 - 2019 **Manager & Senior Research Scientist, Machine Learning**, *Toyota Research Institute (TRI)*.
Manager of the ML team at TRI. Responsible for creating and leading TRI's ML strategy for Automated Driving, from research to ML cloud infrastructure (used in production) and deployment of safety-critical ML models on public roads. Founded and grew the team to a dozen research scientists and ML engineers. Technical advisor for Toyota AI Ventures. External leadership via publications (22 in 2017-2019), invited talks, demonstrations, and scientific collaborations (esp. with Stanford and Preferred Networks).
- 2013 - 2016 **Research Scientist (Computer Vision)**, *Xerox Research Center Europe*, Meylan, France.
Led projects on Deep Learning for video understanding, simulation for perception, combining Computer Vision and Business Process Modeling. Transferred Computer Vision and ML algorithms to production.
- 2008 - 2012 **Doctoral Researcher**, *Microsoft Research - Inria joint center*, Paris, France.
Invented, implemented, and experimentally validated state-of-the-art Computer Vision and Machine Learning algorithms for action recognition in challenging real-world video sources like movies or YouTube videos. Contributions on how to decompose actions into structured, discriminative, and robust models.
- 2008 **R&D Engineer**, *LEAR team, Inria*, Grenoble, France.
Implementation in Python/C/C++ of event and object classification and detection algorithms. Participation to two international Computer Vision competitions: TRECVID and PASCAL VOC. Experimentation on tens of thousands of images and videos, using a cluster of computers, under strong time constraints.
- 06-08/2007 **Research intern**, *Inria Rocquencourt*, Paris, France.
Structure learning of dynamic Bayesian networks using statistical tests and genetic algorithms.
- 06-08/2006 **Research intern**, *LIG research lab, IMAG institute*, Grenoble, France.
Implementation (OCaml) of formal methods in the automatic proof research domain.

Education

- 2008 - 2012 **PhD in Computer Science**, *Microsoft Research - INRIA, Paris & LEAR Team, INRIA Grenoble*, under the supervision of Cordelia Schmid and Zaid Harchaoui, in the fields of Computer Vision and Machine Learning. Title: *Structured Models for Action Recognition in Real-world Videos*.
- 2007 - 2008 **MSc in Artificial Intelligence**, *Institut Polytechnique (INP)*, Grenoble, France.
- 2005 - 2008 **Engineer Diploma in Computer Science and Applied Mathematics**, *ENSIMAG (Ecole Nationale Supérieure d'Informatique et de Mathématiques Appliquées de Grenoble)*, France.
- 2003 - 2005 **"Classes Préparatoires MPSI et MP*"**, *preparation courses (Mathematics and Physics) for the French "Grandes Écoles"*, Clermont-Ferrand, France.
- 2003 **European Scientific Baccalaureate with distinction**, equivalent to "A" levels.

Awards

- 2018 COCO-Mapillary Competition Runner-Up at ECCV 2018 ☞
- 2018 Outstanding reviewer award at CVPR 2018
- 2015 Outstanding reviewer award at CVPR 2015
- 2015 Xerox Innovation Group President's Award for innovative research in Computer Vision
- 2014 Xerox XTIN grant for high risk - high reward project on "Video Analytics in a Virtual World"
- 2008 - 2012 Microsoft Research - Inria PhD scholarship grant
- 2008 Co-winner of the PASCAL VOC 2008 challenge on object classification and detection

Press Coverage & Interviews

TwimLAI podcast: "Advancing Autonomous Vehicle Development Using Distributed Deep Learning", 2019 [↗](#)
DL Summit interview on ML at Toyota, 2019 [↗](#)
W&B interview on ML at TRI, 2019 [↗](#)
AWS Blog: "TRI accelerates safe automated driving with deep learning at a global scale on AWS", 2018 [↗](#)
Forbes: "Artificial Intelligence: The Clever Ways Video Games Are Used To Train AIs", 2018 [↗](#)
Forbes: "How Deep Learning Can Use Virtual Worlds To Solve Real World Problems", 2016 [↗](#)
MIT Tech Review: "To Get Truly Smart, AI Might Need to Play More Video Games", 2016 [↗](#)
Wired: "Making AI Play Lots of Videogames Could Be Huge (No, Seriously)", 2016 [↗](#)
El Espanol: "Máquinas más listas gracias a los videojuegos", 2016 [↗](#)

Communication Skills

— With Humans

Native French

Fluent English & German

Lived in the USA and in Germany

— With Computers

Languages Python (proficient), C/C++, bash

Others PyTorch (main DL framework), Docker, Linux, AWS, HPC

— With Researchers

Computer Vision object detection, tracking, semantic segmentation, action recognition, 3D vision (esp. depth estimation), synthetic data (esp. from simulators and game engines)

Machine Learning Deep Learning (esp. convnets), supervised learning, self-supervised learning, domain adaptation, multi-task learning, optimization, kernel methods, time series analysis

Community

Scientific community Reviewer for the major machine learning, computer vision, and robotics conferences and journals (CVPR, ICLR, ICRA, RSS, NeurIPS, ICML, ICCV, ECCV, BMVC, IJCV, PAMI, TCSVT, ...) External Technical Reviewer for the French National Research Agency (ANR) Annotator for the PASCAL VOC Challenge

Open Source Contributor to and creator of Open Source Python projects, especially on action recognition, camera motion compensation, and kernel methods (<https://github.com/daien>) Organizer of a one-day hackathon on Python and Deep Learning in 2015

Seminars Organization of and presentation at seminars and reading groups on advances in computer vision, deep learning, machine learning, and non-linear optimization

Scientific Activities

40+ publications (1,600+ citations), 12 patents published (7 granted): Google Scholar profile [↗](#) , Patents [↗](#)
Co-organizer of the ICML 2020 workshop on AI for Autonomous Driving (AIAD) [↗](#)
Co-organizer of the ECCV 2020 workshop on Perception for Autonomous Driving (PAD) [↗](#)
Co-organizer of the CVPR 2020 workshop on Frontiers of Monocular 3D Perception [↗](#)
Co-organizer of the ICML 2019 workshop on AI for Autonomous Driving (AIAD) [↗](#)
Guest Editor for International Journal of Computer Vision (IJCV) Special Issue on "Synthetic Visual Data" [↗](#)
Co-organizer of the first international workshop on *Virtual/Augmented Reality for Visual Artificial Intelligence* (VARVAI) at ECCV 2016 & ACM-MM 2016, <http://adas.cvc.uab.es/varvai2016/>

— Publications in peer-reviewed international journals

Spatiotemporal Relationship Reasoning for Pedestrian Intent Prediction,

B. Liu, E. Adeli, Z. Cao, KH Lee, A. Shenoi, A. Gaidon, JC Niebles, RA-L & ICRA, 2020 ↗

Generating Human Action Videos by Coupling 3D Game Engines and Probabilistic Graphical Models,

CR. de Souza, A. Gaidon, Y. Cabon, N. Murray, AM. Lopez, IJCV, 2019, preprint ↗

The Reasonable Effectiveness of Synthetic Visual Data, A. Gaidon, AM. Lopez, F. Perronnin,

Guest Editorial of IJCV Special Issue on "Synthetic Visual Data", 2018, paper ↗

Activity Representation with Motion Hierarchies,

A. Gaidon, Z. Harchaoui, C. Schmid, IJCV, 2014, paper ↗ , code ↗

Temporal Localization of Actions with Actoms,

A. Gaidon, Z. Harchaoui, C. Schmid, PAMI, 2013, paper ↗ , code ↗

— Publications in peer-reviewed international conferences

It Is Not the Journey but the Destination: Endpoint Conditioned Trajectory Prediction, K. Mangalam,

H. Girase, S. Agarwal, K-H. Lee, E. Adeli, J. Malik, A. Gaidon, **ECCV, 2020 (oral, top 2%)** ↗

Monocular Differentiable Rendering for Self-Supervised 3D Object Detection,

D. Beker, H. Kato, MA. Morariu, T. Ando, T. Matsuoka, W. Kehl, A. Gaidon, ECCV, 2020 ↗

Reinforcement Learning based Control of Imitative Policies for Near-Accident Driving,

Z. Cao, E. Biyik, W. Z. Wang, A. Raventos, A. Gaidon, G. Rosman, D. Sadigh, RSS, 2020 ↗

Risk-Sensitive Sequential Action Control with Multi-Modal Human Trajectory Forecasting

for Safe Crowd-Robot Interaction, H. Nishimura, B. Ivanovic, A. Gaidon, M. Pavone, M. Schwager, IROS, 2020

Driving Through Ghosts: Behavioral Cloning with False Positives,

A. Bühler, A. Gaidon, A. Cramariuc, R. Ambrus, G. Rosman, W. Burgard, IROS, 2020

PillarFlow: End-to-end Birds-eye-view Flow Estimation for Autonomous Driving,

KH Lee, M. Kliemann, A. Gaidon, J. Li, C. Fang, S. Pillai, W. Burgard, IROS, 2020 ↗

Game-Theoretic Planning for Risk-Aware Interactive Agents,

M. Wang, N. Mehr, A. Gaidon, M. Schwager, IROS, 2020

Behaviorally Diverse Traffic Simulation via Reinforcement Learning, S. Shiroshita, S. Maruyama, D. Nishiyama,

MY. Castro, K. Hamzaoui, G. Rosman, J. DeCastro, KH. Lee, A. Gaidon, IROS, 2020

Discovering Avoidable Planner Failures of Autonomous Vehicles Using Counterfactual Analysis in Behaviorally Diverse Simulation, D. Nishiyama, MY. Castro, S. Maruyama, S. Shiroshita, K. Hamzaoui, Y. Ouyang,

G. Rosman, J. DeCastro, KH. Lee, A. Gaidon, ITSC, 2020

3D Packing for Self-Supervised Monocular Depth Estimation,

V. Guizilini, R. Ambrus, S. Pillai, A. Raventos, A. Gaidon, **CVPR, 2020 (oral, top 5.7%)** ↗

Autolabeling 3D Objects with Differentiable Rendering of SDF Shape Priors,

S. Zakharov, W. Kehl, A. Bhargava, A. Gaidon, **CVPR, 2020 (oral, top 5.7%)** ↗

Real-Time Panoptic Segmentation from Dense Detections, R. Hou, J. Li, A. Bhargava, A. Raventos, V. Guizilini,

C. Fang, J. Lynch, A. Gaidon, **CVPR, 2020 (oral, top 5.7%)** ↗

Spatio-Temporal Graph for Video Captioning with Knowledge Distillation,

B. Pan, H. Cai, DA Huang, KH Lee, A. Gaidon, E. Adeli, JC Niebles CVPR, 2020 ↗

Semantically-Guided Representation Learning for Self-Supervised Monocular Depth,

V. Guizilini, R. Hou, J. Li, R. Ambrus, A. Gaidon, ICLR, 2020 ↗

Disentangling Human Dynamics for Pedestrian Locomotion Forecasting with Noisy Supervision,

K. Mangalam, E. Adeli, KH. Lee, A. Gaidon, JC. Niebles, WACV (**oral**), 2020 ↗ , video ↗

Learning Imbalanced Datasets with Label-Distribution-Aware Margin Loss,

K. Cao, C. Wei, A. Gaidon, N. Arechiga, T. Ma, NeurIPS, 2019 ↗ , also **oral** at BayLearn 2019 ↗

Exploring the Limitations of Behavior Cloning for Autonomous Driving,

F. Codevilla, E. Santana, AM. Lopez, A. Gaidon, **ICCV, 2019 (oral, top 4.3%)** ↗ , video ↗

Robust Semi-Supervised Monocular Depth Estimation with Reprojected Distances,

V. Guizilini, R. Ambrus, S. Pillai, A. Gaidon, **CoRL, 2019 (spotlight)** ↗

Two Stream Networks for Self-Supervised Ego-Motion Estimation,

R. Ambrus, V. Guizilini, J. Li, S. Pillai, A. Gaidon, **CoRL, 2019 (spotlight)** ↗

An Attention-based Recurrent Convolutional Network for Vehicle Taillight Recognition,
KH. Lee, T. Tagawa, J.M. Pan, A. Gaidon, B. Douillard, IV, 2019

ROI-10D: Monocular Lifting of 2D Detection to 6D Pose and Metric Shape,
F. Manhardt, W. Kehl, A. Gaidon, CVPR, 2019 ↗

SPIGAN: Privileged Adversarial Learning from Simulation,
KH. Lee, G. Ros, J. Li, A. Gaidon, ICLR, 2019 ↗

SuperDepth: Self-Supervised, Super-Resolved Monocular Depth Estimation,
S. Pillai, R. Ambrus, A. Gaidon, ICRA, 2019 ↗

Procedural Generation of Videos to Train Deep Action Recognition Networks,
C. de Souza, A. Gaidon, Y. Cabon, A.M. López, CVPR, 2017, paper ↗ , arXiv ↗ PHAV dataset ↗

Sympathy for the Details: Dense Trajectories and Hybrid Classification Architectures for Action Recognition,
C. de Souza, A. Gaidon, E. Vig, A.M. López, ECCV, 2016, paper ↗ , arXiv ↗

Virtual Worlds as Proxy for Multi-Object Tracking Analysis,
A. Gaidon, Q. Wang, Y. Cabon, E. Vig, CVPR, 2016, paper ↗ , arXiv ↗ , VKITTI dataset ↗

Online Domain Adaptation for Multi-Object Tracking,
A. Gaidon, E. Vig, **BMVC, 2015 (oral, top 7%)**, paper ↗ , arXiv ↗

Deep Fishing: Gradient Features from Deep Nets,
A. Gordo, A. Gaidon, F. Perronnin, BMVC, 2015, paper ↗ , arXiv ↗

Extending Generic BPM with Computer Vision Capabilities,
A. Mos, A. Gaidon, E. Vig, RMSOC, 2015, paper ↗

Recognizing Activities with Cluster-trees of Tracklets,
A. Gaidon, Z. Harchaoui, C. Schmid, BMVC, 2012, paper ↗

A Time Series Kernel for Action Recognition,
A. Gaidon, Z. Harchaoui, C. Schmid, BMVC 2011, paper ↗ , code ↗

Atom Sequence Models for Efficient Action Detection,
A. Gaidon, Z. Harchaoui, C. Schmid, CVPR, 2011, paper ↗ , project ↗

Mining Visual Actions from Movies,
A. Gaidon, M. Marszalek, C. Schmid, **BMVC, 2009 (oral, top 9%)**, paper ↗ , project ↗ , talk ↗

— Other publications

Differentiable Rendering: A Survey,
H. Kato, D. Beker, M. Morariu, T. Ando, T. Matsuoka, W. Kehl, A. Gaidon, arXiv, 2020 ↗

Self-Supervised 3D Keypoint Learning for Ego-motion Estimation,
J. Tang, R. Ambrus, V. Guizilini, S. Pillai, H. Kim, A. Gaidon, arXiv, 2019 ↗

Learning to Fuse Things and Stuff, J. Li, A. Raventos, A. Bhargava, T. Tagawa, A. Gaidon, arXiv, 2018 ↗

Online Learning to Sample, G. Bouchard, G. T. Trouillon, J. Perez, A. Gaidon, arXiv, 2015 ↗

Self-Learning Camera: Autonomous Adaption of Object Detectors to Unlabeled Video Streams,
A. Gaidon, G. Zen, J-A. Rodriguez-Serrano, arXiv, 2014 ↗

Automatic Recognition of Human Activities in Realistic Videos,
A. Gaidon, Z. Harchaoui, C. Schmid, ERCIM News Magazine, 2013 ↗

Classification aided two stage localization,
H. Harzallah, C. Schmid, F. Jurie, A. Gaidon, PASCAL VOC Challenge Workshop, ECCV 2008 ↗

Inria-LEAR's video copy detection system,
M. Douze, A. Gaidon, H. Jegou, M. Marszalek, C. Schmid, TRECVID Workshop 2008 ↗

— Patents

Systems and methods for conditional image translation
German Ros Sanchez, Adrien D. Gaidon, Kuan-Hui Lee, Jie Li, US Patent US-20190354804-A1 ↗

System and method for leveraging end-to-end driving models for improving driving task modules
Shyamal D. Buch, Adrien D. Gaidon, US Patent US-20190113917-A1 ↗

Adversarial learning of photorealistic post-processing of simulation with privileged information
Kuan-Hui Lee, German Ros, Adrien D. Gaidon, Jie Li, US Patent US-2019147582-A1 ↗

Online domain adaptation for multi-object tracking,

A Gaidon, E Vig, US Patent 9,984,315, 2018 [↗](#) , 2018

System and method for video classification using a hybrid unsupervised and supervised multi-layer architecture,

CR De Souza, A Gaidon, E Vig, AM Lopez, US Patent 9,946,933, 2018 [↗](#)

Extracting gradient features from neural networks,

AG Soldevila, A Gaidon, FC Perronnin, US Patent 9792492 [↗](#) , 2017

Efficient object detection with patch-level window processing,

A Gaidon, D Larlus-Larrondo, FC Perronnin, US Patent US9697439 [↗](#) , 2017

Multi-object tracking with generic object proposals,

A. Gaidon, E. Vig, US Patent 9443320 [↗](#) , 2016

Deep data association for online multi-class multi-object tracking,

A Gaidon, US Patent App. 15/089,867 and EP20170163444 [↗](#) , 2016

Generating a virtual world to assess real-world video analysis performance,

Q Wang, A Gaidon, E Vig, US Patent App. 15/051,005 and EP3211596A1 [↗](#) , 2016

Self-learning object detectors from unlabeled videos using multi-task learning,

A. Gaidon, G. Zen, J-A. Rodriguez-Serrano, US Patent 9158971 and EP2916265A1 [↗](#) , 2015

Online domain adaptation for multi-object tracking,

A Gaidon, E Vig US Patent App. 14/704,350 [↗](#) , 2015

Extending generic business process management with computer vision capabilities,

AC Mos, A Gaidon, E Vig US Patent App. 14/697,167 [↗](#) , 2015

— Invited talks and demos

- 2020 "The 3 R's and P's of Autonomous Driving: Robustness, Randomness, and Risk in Perception, Prediction, and Planning", talk at RSS 2020 workshop on Interaction and Decision-Making in Autonomous-Driving [↗](#)

"Scaling up ML for Autonomy", talk at Wayve.ai

"Spatiotemporal Relationship Reasoning for Pedestrian Intent Prediction", Long-term Human Motion Prediction workshop, ICRA'20, [↗](#) , video [↗](#)

- 2019 "Self-Supervised Pseudo-Lidar Networks", Stanford Robotics seminar [↗](#) and ICCV 2019 Workshop on Autonomous Driving - Beyond Single-Frame Perception [↗](#)

"Learning from Structure for World-Scale Automated Driving", ICCV workshop on Autonomous Navigation in Unconstrained Environment (AutoNUE) [↗](#)

"Self-Supervised Monocular Depth for Automated Driving", talk at CVPR 2019 Workshop on Workshop on Autonomous Driving - Beyond Single-Frame Perception [↗](#)

"Beyond Supervised Driving", talk at Element AI & MILA, Montreal

"Predicting and Using Monocular Depth for Deep Driving", talk at IV 2019 workshop on 3D Deep Learning for Automated Driving (3D-DLAD) [↗](#)

"Beyond Supervised Driving", talk at IV 2019 workshop on Unsupervised Learning for Automated Driving (ULAD) [↗](#)

"Scaling Beyond Supervised Driving with PyTorch", talk at F8, San Jose, USA, video [↗](#)

"Beyond Supervised Driving", invited talk at Stanford

"Predicting and Using Monocular Depth for Deep Driving", invited talk at TTIC, Chicago

"Beyond Supervised Driving", oral presentation at GTC'19, San Jose, [↗](#)

"Beyond Supervised Driving", talk at the ML Tokyo Meetup, Tokyo [↗](#)

"Beyond Supervised Driving", talk at the Deep Learning Summit, San Francisco [↗](#) video [↗](#)

"Beyond Supervised Driving", talk at AMLD, EPFL [↗](#)

"Predicting and Using Monocular Depth for Deep Driving", invited talk at Mapillary, Graz, Austria

"Beyond Supervised Driving", invited talk at Naver Labs Europe, Grenoble, France

- 2018 "Advancing Autonomous Vehicle Development Using Distributed Deep Learning", talk at AWS re:Invent, Las Vegas [↗](#)

Hacker Dojo Course on Autonomous Vehicles, Santa Clara [↗](#)

- "Beyond Supervised Driving" invited talk at University of Amsterdam, NL, video [↗](#)
- "Bridging the gap between fundamental research and product development at TRI by applying techniques from AI and ML" invited talk at World Summit AI, Amsterdam, NL [↗](#)
- "Beyond Supervised Driving" invited talk at the Beyond Supervised Learning workshop, CVPR 2018, Salt Lake City, [↗](#)
- "Living on the edge case: how autonomous vehicles can deal with an extremely messy world", panel discussion at The Autonomous Vehicle Summit, San Francisco [↗](#) , Recording [↗](#)
- "Advancing State-of-the-Art of Autonomous Vehicles and Robotics Research using AWS GPU Instances", invited talk at GTC 2018 [↗](#)
- "Deep Learning and Simulation for Automated Driving at Scale" invited talk at the Auto AI Conference, San Francisco
- "Simulation and Deep Learning for Autonomous Driving" invited talk at University of Washington
- 2017 Panel discussion and invited talk on "Deep Learning and Simulation for real-world Machine Intelligence in Autonomous Driving" at NIPS 2017 ML for Intelligent Transportation Systems [↗](#)
- "Need for Sim: Procedural Generation of Realistic Driving Environments and Human Actions for Deep Learning", invited talk at Google Brain, Mountain View
- "Need for Sim: Procedural Generation of Realistic Driving Environments and Human Actions for Deep Learning", IV 2017 Deep Driving workshop
- 2016 "Realistic Virtual Worlds and Human Actions for Video Understanding" demo at NIPS 2016
- Invited talks at Zoox, San Francisco; THOTH, Inria, France; Philips Research, Paris, France
- Virtual Worlds for Computer Vision demo at CVPR 2016, Las Vegas
- 2014 Self-Learning Camera demo at ECCV 2014
- Invited presentation at the TASK-CV workshop at ECCV 2014
- Invited talk at LEAR - Inria Grenoble, France
- 2013 Invited talk at ERMITES Summer School, France
- Invited talk at Xerox Research Center Europe, France
- 2012 Invited talk at the MSR - Inria CVML workshop, Microsoft Research, Cambridge, UK
- Invited talk at ETH Zürich, Switzerland
- Invited talk at the Xerox Research Center Europe (XRCE), Meylan, France
- 2011 Invited talk at the GdR ISIS, CNRS, Paris
- Invited talk at the Computer Vision Center (CVC), Autonomous University of Barcelona, Spain
- Technical demo at the Microsoft Research - Inria forum, Paris, France
- 2010 Invited talk at the Visual Geometry Group, University of Oxford, UK
- Invited talk at the Microsoft Research - Inria vision workshop, Paris, France

— Collaboration

- 2017- Collaboration with Stanford on ML for Automated Driving on multiple projects spanning perception, prediction, planning, control
- 2018 Internship supervisor of Felipe Codevilla with Eder Santana and Antonio Lopez (CVC)
- 2017 Internship supervisor of Shyamal Buch with Juan Carlos Nieves (Stanford)
- 2015-2017 PhD co-supervisor of Cesar de Souza on Learning Representations for Behavior Understanding in Videos, collaboration with Professor Antonio Lopez at the Computer Vision Center of the Autonomous University of Barcelona (UAB), Spain
- 2015-2016 Co-supervisor of Vladyslav Sydorov on human pose estimation in videos, collaboration with Kartteek Alahari and Cordelia Schmid at Inria, France
- 2015-2016 Internship supervisor of Yohann Cabon on virtual worlds for video analytics
- 2015 Internship co-supervisor of Qiao Wang on virtual worlds for multi-object tracking analysis
- 2014 Internship co-supervisor of Gloria Zen on self-learning cameras