Lorenzo Torresani

https://ltorresa.github.io   torresani@gmail.com   +1(650)690-0070

Education

9/2005	**Ph.D. in Computer Science**, Stanford University. Thesis advisor: Christoph Bregler.
6/2001	**M.S. in Computer Science**, Stanford University.
4/1996	**Laurea in Computer Science**, *Summa cum Laude*, University of Milan, Italy.

Professional Experience

3/2023 – present	**Research Director**, Fundamental AI Research (FAIR), Meta, New York, NY.
8/2021 – 2/2023	**Research Manager**, Fundamental AI Research (FAIR), Meta, New York, NY.
1/2022 – present	**Adjunct Professor**, Computer Science, Dartmouth College.
7/2020 – 12/2021	**Professor**, Computer Science, Dartmouth College.
7/2018 – 7/2021	**Research Scientist**, Facebook AI Research (FAIR), Cambridge, MA.
7/2014 – 6/2020	**Associate Professor with Tenure**, Computer Science, Dartmouth College.
1/2018 – 5/2018	**Visiting Faculty & Fulbright U.S. Scholar**, Ashesi University, Berekuso, Ghana.
7/2017 – 12/2017	**Visiting Research Scientist**, Facebook, Menlo Park, CA.
1/2009 – 6/2014	**Assistant Professor**, Computer Science, Dartmouth College.
6/2007 – 12/2008	**Associate Researcher**, Microsoft Research, Cambridge, United Kingdom.
12/2005 – 4/2007	**Senior Researcher**, Riya/Like.com, Inc., San Mateo, CA, USA.
9/2005 – 11/2005	**Member of Technical Staff**, Sarnoff Corporation, Princeton, NJ, USA.
3/2003 – 9/2005	**Research Scientist**, Computer Science Department, New York University.
7/2002 – 9/2002	**Visiting Scholar**, Department of EECS, U.C. Berkeley.
1/2001 – 2/2003	**Research Assistant**, Computer Science Department, Stanford University.
8/1996 – 12/2002	**Researcher**, Digital Persona, Inc., Redwood City, CA, USA.
1/1995 – 4/1996	**Research Assistant**, Istituto per la Ricerca Scientifica e Tecnologica, Trento, Italy.

Selected Awards

2018, 2016, 2015	**Facebook Faculty Research Award**, Facebook.
2017	**Fulbright U.S. Scholar Award**, J. William Fulbright Foreign Scholarship Board.
2017	**Neukom Institute CompX Faculty Grants Award**, Dartmouth College.
2016	**Scholarly Innovation and Advancement Award**, Dartmouth College.
2014	**Google Faculty Research Award**, Google.
2014	**Susan and Gib Myers 1964 Dartmouth Faculty Fellowship**, Dartmouth College.
2014	**Neukom Institute CompX Faculty Grants Award**, Dartmouth College.
2011	**Microsoft Research Faculty Fellowship Finalist**, Selected as one of the 16 most promising academic researchers worldwide for research and innovation in computer science.
2010	**Junior Faculty Fellowship**, Dartmouth College.
2010	**Burke Research Initiation Award**, Dartmouth College.
2010	**CAREER Award**, National Science Foundation.
2001	**Best Student Paper Award**, IEEE CVPR.

Publications

Refereed journal articles

[J9] Y. Xian, B. Korbar, M. Douze, L. Torresani, B. Schiele, and Z. Akata. "Generalized Few-Shot Video Classification with Video Retrieval and Feature Generation," to appear in *IEEE Transactions on Pattern Analysis and Machine Intelligence, TPAMI*, 2022.
https://arxiv.org/abs/2007.04755

[J8] M.H. Baig, and L. Torresani. "Multiple Hypothesis Colorization and Its Application to Image Compression," *Computer Vision and Image Understanding, CVIU*, 164:111-123, Nov. 2017.
JCR ISI 2015 impact factor = 2.13.
http://dx.doi.org/10.1016/j.cviu.2017.01.010

[J7] D. Tran, and L. Torresani. "EXMOVES: Mid-level Features for Efficient Action Recognition and Video Analysis," *International Journal of Computer Vision, IJCV*, 119(3):239-253, Sept. 2016.
JCR ISI 2015 impact factor = 4.27.
http:/dx.doi.org/10.1007/s11263-016-0905-6

[J6] S. Rodríguez-Vaamonde, L. Torresani and A.W. Fitzgibbon. "What can pictures tell us about web pages? Improving document search using images," *IEEE Transactions on Pattern Analysis and Machine Intelligence, TPAMI*, 37(6):1274-1285, June 2015.
JCR ISI 2015 impact factor = 6.07.
http://dx.doi.org/10.1109/TPAMI.2014.2366761

[J5] A. Bergamo, and L. Torresani. "Classemes and Other Classifier-based Features for Efficient Object Categorization," *IEEE Transactions on Pattern Analysis and Machine Intelligence, TPAMI*, 36(10):1988-2001, Oct. 2014.
JCR ISI 2014 impact factor = 5.78.
http://dx.doi.org/10.1109/TPAMI.2014.2313111

[J4] M.H. Nguyen, L. Torresani, F. de la Torre, and C. Rother. "Learning Discriminative Localization from Weakly Labeled Data." *Pattern Recognition*, 47(3):1523-1534, March 2014.
JCR ISI 2011 impact factor = 3.09.
http://dx.doi.org/10.1016/j.patcog.2013.09.028

[J3] A. Chandrashekar, L. Torresani, and R. Granger. "Learning what is where from unlabeled images: joint localization and clustering of foreground objects." *Machine Learning Journal*, Springer, Feb. 2014.
JCR ISI 2014 impact factor = 1.88.
http://dx.doi.org/10.1007/s10994-013-5330-2

[J2] L. Torresani, V. Kolmogorov, and C. Rother. "A Dual Decomposition Approach to Feature Correspondence." *IEEE Transactions on Pattern Analysis and Machine Intelligence, TPAMI*, 35(2):259-271, Feb. 2013.
JCR ISI 2013 impact factor = 5.69.
http://dx.doi.org/10.1109/TPAMI.2012.105

[J1] L. Torresani, A. Hertzmann, and C. Bregler. "Non-Rigid Structure-From-Motion: Estimating Shape and Motion with Hierarchical Priors." *IEEE Transactions on Pattern Analysis and Machine Intelligence, TPAMI*, 30(5):878-892, May 2008.
JCR ISI 2008 impact factor = 5.96.
http://dx.doi.org/10.1109/TPAMI.2007.70752

Refereed conference papers

[C82] K. Grauman, A. Westbury, L. Torresani, K. Kitani, J. Malik, T. Afouras, K. Ashutosh, V. Baiyya, S. Bansal, B. Boote, E. Byrne, Z. Chavis, J. Chen, F. Cheng, F.J. Chu, S. Crane, A. Dasgupta, J. Dong, M. Escobar, C. Forigua, A. Gebreselasie, S. Haresh, J. Huang, M.M. Islam, S. Jain, R. Khirodkar,

D. Kukreja, K.J. Liang, J.W. Liu, S. Majumder, Y. Mao, M. Martin, E. Mavroudi, T. Nagarajan, F. Ragusa, S.K. Ramakrishnan, L. Seminara, A. Somayazulu, Y. Song, S. Su, Z. Xue, E. Zhang, J. Zhang, A. Castillo, C. Chen, X. Fu, R. Furuta, C. Gonzalez, P. Gupta, J. Hu, Y. Huang, Y. Huang, W. Khoo, A. Kumar, R. Kuo, S. Lakhavani, M. Liu, M. Luo, Z. Luo, B. Meredith, A. Miller, O. Oguntola, X. Pan, P. Peng, S. Pramanick, M. Ramazanova, F. Ryan, W. Shan, K. Somasundaram, C. Song, A. Southerland, M. Tateno, H. Wang, Y. Wang, T. Yagi, M. Yan, X. Yang, Z. Yu, S.C. Zha, C. Zhao, Z. Zhao, Z. Zhu, J. Zhuo, P. Arbelaez, G. Bertasius, D. Crandall, D. Damen, J. Engel, G.M. Farinella, A. Furnari, B. Ghanem, J. Hoffman, C.V. Jawahar, R. Newcombe, H.S. Park, J.M. Rehg, Y. Sato, M. Savva, J. Shi, M.Z. Shou, M. Wray, "Ego-Exo4D: Understanding Skilled Human Activity from First- and Third-Person Perspectives," to appear in *IEEE Conference On Computer Vision and Pattern Recognition, CVPR*, June 2024.
Selected for oral presentation ($< 1\%$ accept rate).
https://arxiv.org/abs/2311.18259

[C81] Y. Shen, H. Wang, X. Yang, M. Feiszli, E. Elhamifar, L. Torresani, and E. Mavroudi, "Learning to Segment Referred Objects from Narrated Egocentric Videos," to appear in *IEEE Conference On Computer Vision and Pattern Recognition, CVPR*, June 2024.
Selected for oral presentation ($< 1\%$ accept rate).

[C80] T. Nagarajan, and L. Torresani, "Step Differences in Instructional Video," to appear in *IEEE Conference On Computer Vision and Pattern Recognition, CVPR*, June 2024.

[C79] M.M. Islam, N. Ho, X. Yang, T. Nagarajan, L. Torresani, and G. Bertasius, "Video ReCap: Recursive Captioning of Hour-Long Videos," to appear in *IEEE Conference On Computer Vision and Pattern Recognition, CVPR*, June 2024.
https://arxiv.org/abs/2402.13250

[C78] Y. Song, G. Byrne, T. Nagarajan, H. Wang, M. Martin, and L. Torresani, "Ego4D Goal-Step: Toward Hierarchical Understanding of Procedural Activities," *Advances in Neural Information Processing Systems, NeurIPS*, December 2023.
Selected as spotlight paper ($\sim 3\%$ of submitted papers).
https://openreview.net/forum?id=3BxYAaovKr

[C77] T. Afouras, E. Mavroudi, T. Nagarajan, H. Wang, and L. Torresani, "HT-Step: Aligning Instructional Articles with How-To Videos," *Advances in Neural Information Processing Systems, NeurIPS*, December 2023.
https://openreview.net/forum?id=vv3cocNsEK

[C76] H. Wang, M.K. Singh, and L. Torresani, "Ego-Only: Egocentric Action Detection without Exocentric Transferring," *IEEE International Conference on Computer Vision, ICCV*, October 2023.
https://openaccess.thecvf.com/content/ICCV2023/papers/Wang_Ego-Only_Egocentric_Action_Detection_without_Exocentric_Transferring_ICCV_2023_paper.pdf

[C75] E. Mavroudi, T. Afouras, and L. Torresani,"Learning to Ground Instructional Articles in Videos through Narrations," *IEEE International Conference on Computer Vision, ICCV*, October 2023.
https://openaccess.thecvf.com/content/ICCV2023/papers/Mavroudi_Learning_to_Ground_Instructional_Articles_in_Videos_through_Narrations_ICCV_2023_paper.pdf

[C74] Z. Xue, Y. Song, K. Grauman, and L. Torresani, "Egocentric Video Task Translation," *IEEE Conference On Computer Vision and Pattern Recognition, CVPR*, June 2023.
Selected as highlight paper ($\sim 2.6\%$ of submitted papers).
https://openaccess.thecvf.com/content/CVPR2023/papers/Xue_Egocentric_Video_Task_Translation_CVPR_2023_paper.pdf

[C73] K. Ashutosh, R. Girdhar, L. Torresani, and K. Grauman, "HierVL: Learning Hierarchical Video-Language Embeddings," *IEEE Conference On Computer Vision and Pattern Recognition, CVPR*, June 2023.
Selected as highlight paper ($\sim 2.6\%$ of submitted papers).

https://openaccess.thecvf.com/content/CVPR2023/papers/Ashutosh_HierVL_Learning_Hierarchical_Video-Language_Embeddings_CVPR_2023_paper.pdf

[C72] X. Yang, F.J. Chu, M. Feiszli, R. Goyal, L. Torresani, and D. Tran, "Relational Space-Time Query in Long-Form Videos," *IEEE Conference On Computer Vision and Pattern Recognition, CVPR*, June 2023.
Selected as highlight paper ($\sim$ 2.6% of submitted papers).
https://openaccess.thecvf.com/content/CVPR2023/papers/Yang_Relational_Space-Time_Query_in_Long-Form_Videos_CVPR_2023_paper.pdf

[C71] Y. Jian, and L. Torresani, "Label hallucination for few-shot classification," *AAAI Conference on Artificial Intelligence*, September 2022.
https://ojs.aaai.org/index.php/AAAI/article/view/20659/20418

[C70] K. Grauman, A. Westbury, E. Byrne, Z. Chavis, A. Furnari, R. Girdhar, J. Hamburger, H. Jiang, M. Liu, X. Liu, M. Martin, T. Nagarajan, I. Radosavovic, S.K. Ramakrishnan, F. Ryan, J. Sharma, M. Wray, M. Xu, E.Z. Xu, C. Zhao, S. Bansal, D. Batra, V. Cartillier, S. Crane, T. Do, M. Doulaty, A. Erapalli, C. Feichtenhofer, A. Fragomeni, Q. Fu, A. Gebreselasie, C. González, J. Hillis, X. Huang, Y. Huang, W. Jia, W. Khoo, J. Kolar, S. Kottur, A. Kumar, F. Landini, C. Li, Y. Li, Z. Li, K. Mangalam, R. Modhugu, J. Munro, T. Murrell, T. Nishiyasu, W. Price, P. Ruiz, M. Ramazanova, L. Sari, K. Somasundaram, A. Southerland, Y. Sugano, R. Tao, M. Vo, Y. Wang, X. Wu, T. Yagi, Z. Zhao, Y. Zhu, P. Arbelaez[†], D. Crandall[†], D. Damen[†], G.M. Farinella[†], C. Fuegen[†], B. Ghanem[†], V.K. Ithapu[†], CV Jawahar[†], H. Joo[†], K. Kitani[†], H. Li[†], R. Newcombe[†], A. Oliva[†], H.S. Park[†], J.M. Rehg[†], Y. Sato[†], J. Shi[†], M.Z. Shou[†], A. Torralba[†], L. Torresani[†], M. Yan[†], and J. Malik[†], "Ego4D: Around the World in 3,000 Hours of Egocentric Video," *IEEE Conference On Computer Vision and Pattern Recognition, CVPR*, June 2022.
Authors with daggers ([†]) are faculty PIs and working group leads in the project.
Paper award nominee ($\sim$ 1.6% of accepted papers).
Selected for oral presentation. Acceptance rate for oral presentation = $\sim$ 4%.
https://openaccess.thecvf.com/content/CVPR2022/papers/Grauman_Ego4D_Around_the_World_in_3000_Hours_of_Egocentric_Video_CVPR_2022_paper.pdf

[C69] X. Lin, F. Petroni, G. Bertasius, M. Rohrbach, S.F. Chang, and L. Torresani, "Learning To Recognize Procedural Activities with Distant Supervision," *IEEE Conference On Computer Vision and Pattern Recognition, CVPR*, June 2022.
https://openaccess.thecvf.com/content/CVPR2022/papers/Lin_Learning_To_Recognize_Procedural_Activities_With_Distant_Supervision_CVPR_2022_paper.pdf

[C68] J. Wang, and L. Torresani, "Deformable Video Transformer," *IEEE Conference On Computer Vision and Pattern Recognition, CVPR*, June 2022.
https://openaccess.thecvf.com/content/CVPR2022/papers/Wang_Deformable_Video_Transformer_CVPR_2022_paper.pdf

[C67] J. Wang, G. Bertasius, D. Tran, and L. Torresani, "Long-short temporal contrastive learning of video transformers," *IEEE Conference On Computer Vision and Pattern Recognition, CVPR*, June 2022.
https://openaccess.thecvf.com/content/CVPR2022/papers/Wang_Long-Short_Temporal_Contrastive_Learning_of_Video_Transformers_CVPR_2022_paper.pdf

[C66] G. Bertasius, H. Wang, and L. Torresani, "Is Space-Time Attention All You Need for Video Understanding?," *International Conference on Machine Learning, ICML*, July 2021.
https://arxiv.org/abs/2102.05095

[C65] M.M. Aladago, and L. Torresani, "Slot Machines: Discovering Winning Combinations of Random Weights in Neural Networks," *International Conference on Machine Learning, ICML*, July 2021.
https://arxiv.org/abs/2101.06475

[C64] X. Lin, G. Bertasius, J. Wang, S.F. Chang, D. Parikh, and L. Torresani, "VX2TEXT: End-to-End Learning of Video-Based Text Generation From Multimodal Inputs," *IEEE Conference On Computer*

Vision and Pattern Recognition, CVPR, June 2021.
https://arxiv.org/abs/2101.12059

[C63] X. Yang, H. Fan, L. Torresani, L. Davis, and H. Wang, "Beyond Short Clips: End-to-End Video-Level Learning with Collaborative Memories," *IEEE Conference On Computer Vision and Pattern Recognition, CVPR*, June 2021.
https://arxiv.org/abs/2104.01198

[C62] J. Wei, A. Suriawinata, B. Ren, X. Liu, M. Lisovsky, L. Vaickus, C. Brown, M. Baker, N. Tomita, L. Torresani, J. Wei, and S. Hassanpour. "A Petri Dish for Histopathology Image Analysis," *19th International Conference on Artificial Intelligence in Medicine, AIME*, June 2021.
https://arxiv.org/abs/2101.12355

[C61] L. Sari, K. Singh, J. Zhou, L. Torresani, N. Singhal, Y. Saraf. "A Multi-View Approach to Audio-Visual Speaker Verification," *IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP*, June 2021.
https://arxiv.org/abs/2102.06291

[C60] L. Sevilla-Lara, S. Zha, Z. Yan, V. Goswami, M. Feiszli, and L. Torresani. "Only Time Can Tell: Discovering Temporal Data for Temporal Modeling," *IEEE Winter Conference on Applications of Computer Vision, WACV*, January 2021.
Accepted in round 1 of Algorithms track.
https://arxiv.org/abs/1907.08340

[C59] Y. Wang, G. Bertasius, T.H. Oh, A. Gupta, M. Hoai, and L. Torresani. "Supervoxel Attention Graphs for Long-Range Video Modeling," *IEEE Winter Conference on Applications of Computer Vision, WACV*, January 2021.
Accepted in round 2 of Algorithms track.
https://openaccess.thecvf.com/content/WACV2021/papers/Wang_Supervoxel_Attention_Graphs_for_Long-Range_Video_Modeling_WACV_2021_paper.pdf

[C58] J. Wei, A. Suriawinata, B. Ren, X. Liu, M. Lisovsky, L. Vaickus, C. Brown, M. Baker, M. Nasir-Moin, N. Tomita, L. Torresani, J. Wei, and S. Hassanpour. "Learn like a Pathologist: Curriculum Learning by Annotator Agreement for Histopathology Image Classification," *IEEE Winter Conference on Applications of Computer Vision, WACV*, January 2021.
Accepted in round 2 of Applications track.
https://arxiv.org/abs/2009.13698

[C57] G. Bertasius, and L. Torresani. "COBE: Contextualized Object Embeddings from Narrated Instructional Video," *Advances in Neural Information Processing Systems, NeurIPS*, December 2020.
Acceptance rate = 20%.
https://arxiv.org/abs/2007.07306

[C56] H. Alwassel, D. Mahajan, B. Korbar, L. Torresani, B. Ghanem, and D. Tran. "Self-Supervised Learning by Cross-Modal Audio-Video Clustering," *Advances in Neural Information Processing Systems, NeurIPS*, December 2020.
Selected for spotlight presentation. Acceptance rate for spotlight presentation = 3%
https://arxiv.org/abs/1911.12667

[C55] Y. Wang, V. Tran, G. Bertasius, L. Torresani, and M. Hoai. "Attentive Action and Context Factorization," *British Machine Vision Virtual Conference, BMVC*, September 2020.
Acceptance rate = 29.1%.
https://arxiv.org/abs/1904.05410

[C54] J. Han, F. Ding, X. Liu, L. Torresani, J. Peng, and Q. Liu. "Stein Variational Inference for Discrete Distributions," *International Conference on Artificial Intelligence and Statistics, AISTATS*, August 2020.
https://arxiv.org/abs/2003.00605

[C53] G. Bertasius, and L. Torresani. "Classifying, Segmenting, and Tracking Object Instances in Video with Mask Propagation," *IEEE Conference On Computer Vision and Pattern Recognition, CVPR*, June 2020.
Paper award nominee ($\sim$ 2% of accepted papers).
Selected for oral presentation. Acceptance rate for oral presentation = $\sim$ 5%.
https://arxiv.org/abs/1912.04573

[C52] R. Gao, T.H. Oh, K. Grauman, and L. Torresani. "Listen to Look: Action Recognition by Previewing Audio," *IEEE Conference On Computer Vision and Pattern Recognition, CVPR*, June 2020.
Acceptance rate = 22%.
https://arxiv.org/abs/1912.04487

[C51] H. Wang, D. Tran, L. Torresani, and M. Feiszli. "Correlation Networks for Video Classification," *IEEE Conference On Computer Vision and Pattern Recognition, CVPR*, June 2020.
Acceptance rate = 22%.
https://arxiv.org/abs/1906.03349

[C50] G. Bertasius, C. Feichtenhofer, D. Tran, J. Shi, and L. Torresani. "Learning Temporal Pose Estimation from Sparsely-Labeled Videos," *Advances in Neural Information Processing Systems, NeurIPS*, December 2019.
Acceptance rate = 21.2%.
https://papers.nips.cc/paper/8567-learning-temporal-pose-estimation-from-sparsely-labeled-videos

[C49] K. Ahmed, and L. Torresani. "STAR-Caps: Capsule Networks with Straight-Through Attentive Routing," *Advances in Neural Information Processing Systems, NeurIPS*, December 2019.
Acceptance rate = 21.2%.
https://papers.nips.cc/paper/9110-star-caps-capsule-networks-with-straight-through-attentive-routing

[C48] B. Korbar, D. Tran, and L. Torresani. "SCSampler: Sampling Salient Clips from Video for Efficient Action Recognition," *IEEE International Conference on Computer Vision, ICCV*, October 2019.
Selected for oral presentation. Acceptance rate for oral presentation = 4.3%.
http://openaccess.thecvf.com/content_ICCV_2019/papers/Korbar_SCSampler_Sampling_Salient_Clips_From_Video_for_Efficient_Action_Recognition_ICCV_2019_paper.pdf

[C47] D. Tran, H. Wang, L. Torresani, and M. Feiszli. "Video Classification with Channel-Separated Convolutional Networks," *IEEE International Conference on Computer Vision, ICCV*, October 2019.
Acceptance rate = 25.0%.
http://openaccess.thecvf.com/content_ICCV_2019/papers/Tran_Video_Classification_With_Channel-Separated_Convolutional_Networks_ICCV_2019_paper.pdf

[C46] R. Girdhar, D. Tran, L. Torresani, and D. Ramanan. "DistInit: Learning Video Representations without a Single Labeled Video," *IEEE International Conference on Computer Vision, ICCV*, October 2019.
Acceptance rate = 25.0%.
http://openaccess.thecvf.com/content_ICCV_2019/papers/Girdhar_DistInit_Learning_Video_Representations_Without_a_Single_Labeled_Video_ICCV_2019_paper.pdf

[C45] H. Zhao, A. Torralba, L. Torresani, Z. Yan. "HACS: Human Action Clips and Segments Dataset for Recognition and Temporal Localization," *IEEE International Conference on Computer Vision, ICCV*, October 2019.
Acceptance rate = 25.0%.
http://openaccess.thecvf.com/content_ICCV_2019/papers/Zhao_HACS_Human_Action_Clips_and_Segments_Dataset_for_Recognition_and_ICCV_2019_paper.pdf

[C45] B. Korbar, D. Tran, and L. Torresani. "Cooperative Learning of Audio and Video Models from Self-Supervised Synchronization," *Advances in Neural Information Processing Systems, NeurIPS*, 2018.

Acceptance rate = 20.8%.
`http://papers.nips.cc/paper/8002-cooperative-learning-of-audio-and-video-models-fro`
`m-self-supervised-synchronization.pdf`

[C44] G. Bertasius, L. Torresani, and J. Shi. "Object Detection in Video with Spatiotemporal Sampling Networks," *European Conference on Computer Vision, ECCV*, 2018.
Acceptance rate = 27.7%.
`http://openaccess.thecvf.com/content_ECCV_2018/papers/Gedas_Bertasius_Object_Detection`
`_in_ECCV_2018_paper.pdf`

[C43] K. Ahmed, and L. Torresani. "MaskConnect: Connectivity Learning by Gradient Descent," *European Conference on Computer Vision, ECCV*, 2018.
Acceptance rate = 27.7%.
`http://openaccess.thecvf.com/content_ECCV_2018/papers/Karim_Ahmed_MaskConnect_Connect`
`ivity_Learning_ECCV_2018_paper.pdf`

[C42] J. Ray, H. Wang, D. Tran, Y. Wang, M. Feiszli, L. Torresani, and M. Paluri. "Scenes-Objects-Actions: A Multi-Task, Multi-Label Video Dataset," *European Conference on Computer Vision, ECCV*, 2018.
Acceptance rate = 27.7%.
`http://openaccess.thecvf.com/content_ECCV_2018/papers/Heng_Wang_Scenes-Objects-Action`
`s_A_Multi-Task_ECCV_2018_paper.pdf`

[C41] S. Singh, A. Batra, G. Pang, L. Torresani, S. Basu, M. Paluri, and C.V. Jawahar. "Self-Supervised Feature Learning for Semantic Segmentation of Overhead Imagery," *British Machine Vision Conference, BMVC*, 2018.
Acceptance rate = 29.5%.
`http://bmvc2018.org/contents/papers/0345.pdf`

[C40] D. Tran, H. Wang, L. Torresani, J. Ray, Y. LeCun, and M. Paluri. "A Closer Look at Spatiotemporal Convolutions for Action Recognition," *IEEE Conference On Computer Vision and Pattern Recognition, CVPR*, 2018.
Acceptance rate = 29.1%.
`http://openaccess.thecvf.com/content_cvpr_2018/papers/Tran_A_Closer_Look_CVPR_2018_pap`
`er.pdf`

[C39] D.A. Huang, V. Ramanathan, D. Mahajan, L. Torresani, M. Paluri, J.C. Niebles, and L. Fei-Fei. "What Makes a Video a Video: Analyzing Temporal Information in Video Understanding Models and Datasets," *IEEE Conference On Computer Vision and Pattern Recognition, CVPR*, 2018.
Selected for spotlight oral presentation. Acceptance rate for spotlight oral presentation = 6.6%.
`http://openaccess.thecvf.com/content_cvpr_2018/papers/Huang_What_Makes_a_CVPR_2018_pap`
`er.pdf`

[C38] R. Girdhar, G. Gkioxari, L. Torresani, M. Paluri, and D. Tran. "Detect-and-Track: Efficient Pose Estimation in Videos," *IEEE Conference On Computer Vision and Pattern Recognition, CVPR*, 2018.
Acceptance rate = 29.1%.
`http://openaccess.thecvf.com/content_cvpr_2018/papers/Girdhar_Detect-and-Track_Effic`
`ient_Pose_CVPR_2018_paper.pdf`

[C37] K. Ahmed, and L. Torresani. "BranchConnect: Large-Scale Visual Recognition with Learned Branch Connections," *IEEE Winter Conference on Applications of Computer Vision, WACV*, 2018.
Accepted in round 1 of Algorithms track. Acceptance rate = 37%.
`https://arxiv.org/abs/1704.06010`

[C36] M.H. Baig, V. Koltun, and L. Torresani. "Learning to Inpaint for Image Compression," *Advances in Neural Information Processing Systems, NIPS*, 2017.
Acceptance rate = 20.9%.
`http://papers.nips.cc/paper/6724-learning-to-inpaint-for-image-compression.pdf`

7

[C35] G. Bertasius, L. Torresani, S. Yu, and J. Shi. "Convolutional Random Walk Networks for Semantic Image Segmentation," *IEEE Conference On Computer Vision and Pattern Recognition, CVPR*, 2017. Acceptance rate = 29.9%.
https://doi.org/10.1109/CVPR.2017.650

[C34] L. Bazzani, H. Larochelle, and L. Torresani. "Recurrent Mixture Density Network for Spatiotemporal Visual Attention," *International Conference on Learning Representations, ICLR*, 2017. Acceptance rate = 38.6%.
https://openreview.net/pdf?id=SJRpRfKxx

[C33] G. Bertasius, Q. Liu, L. Torresani, and J. Shi. "Local Perturb-and-MAP for Structured Prediction," *International Conference on Artificial Intelligence and Statistics, AISTATS*, 2017. Acceptance rate = 31.7%.
http://proceedings.mlr.press/v54/bertasius17a/bertasius17a.pdf

[C32] K. Ahmed, M.H. Baig, and L. Torresani. "Network of Experts for Large-Scale Image Categorization," *European Conference On Computer Vision, ECCV*, October 2016, 516-532. Acceptance rate = 26.6%.
http://dx.doi.org/10.1007/978-3-319-46478-7_32

[C31] G. Bertasius, J. Shi, and L. Torresani. "Semantic Segmentation with Boundary Neural Fields," *IEEE Conference On Computer Vision and Pattern Recognition, CVPR*, June 2016, 3602-3610. Acceptance rate = 29.9%.
http://dx.doi.org/10.1109/CVPR.2016.392

[C30] L. Bazzani, A. Bergamo, D. Anguelov, and L. Torresani. "Self-Taught Object Localization with Deep Networks," *IEEE Winter Conference on Applications of Computer Vision, WACV*, March 2016. Accepted in Algorithms track. Acceptance rate = 42.3%.
http://arxiv.org/abs/1409.3964

[C29] M.H. Baig, and L. Torresani. "Coupled Depth Learning," *IEEE Winter Conference on Applications of Computer Vision, WACV*, March 2016. Accepted in round 1 of Algorithms track. Acceptance rate = 30%.
http://arxiv.org/abs/1501.04537

[C28] G. Bertasius, J. Shi, and L. Torresani. "High-for-Low and Low-for-High: Efficient Boundary Detection from Deep Object Features and its Applications to High-Level Vision," *IEEE International Conference on Computer Vision, ICCV*, December 2015, 504-512. Acceptance rate = 30.9%.
http://arxiv.org/abs/1504.06201

[C27] D. Tran, L. Bourdev, R. Fergus, L. Torresani, and M. Paluri. "Learning Spatiotemporal Features with 3D Convolutional Networks," *IEEE International Conference on Computer Vision, ICCV*, December 2015, 4489-4497. Acceptance rate = 30.9%.
http://arxiv.org/abs/1412.0767

[C26] G. Bertasius, J. Shi, and L. Torresani. "DeepEdge: A Multi-Scale Bifurcated Deep Network for Top-Down Contour Detection," *IEEE Conference On Computer Vision and Pattern Recognition, CVPR*, June 2015, 4380-4389. Acceptance rate = 28%.
http://dx.doi.org/10.1109/CVPR.2015.7299067

[C25] D. Tran, and L. Torresani. "EXMOVES: Classifier-based Features for Scalable Action Recognition," *International Conference on Learning Representations, ICLR*, 2014. Acceptance rate = 50%.
http://arxiv.org/pdf/1312.5785.pdf

[C24] K. Ramnath, S. Baker, L. Vanderwende, M. El-Saban, S. Sinha, A. Kannan, N. Hasan, M. Galley, Y. Yang, D. Ramanan, A. Bergamo, and L. Torresani. "AutoCaption: Automatic Caption Generation for Personal Photos," *IEEE Winter Conference on Applications of Computer Vision, WACV*, pages 1050-1057, 2014.
Acceptance rate = 40%.
http://dx.doi.org/10.1109/WACV.2014.6835988

[C23] S. Rodríguez-Vaamonde, L. Torresani, and A.W. Fitzgibbon. "What can pictures tell us about web pages? Improving document search using images," *The 36th International ACM SIGIR conference on research and development in Information Retrieval, SIGIR '13*, pages 849-852, 2013.
Acceptance rate = 35%.
http://dx.doi.org/10.1145/2484028.2484144

[C22] C.W. You, N.D. Lane, F. Chen, R. Wang, Z. Chen, T.J. Bao, Y. Cheng, M. Lin, L. Torresani, and A.T. Campbell. "CarSafe App: Alerting Drowsy and Distracted Drivers using Dual Cameras on Smartphones," *The 11th Annual International Conference on Mobile Systems, Applications, and Services, MobiSys*, pages 13-26, 2013.
Acceptance rate = 15.7%.
http://dx.doi.org//10.1145/2462456.2465428

[C21] A. Bergamo, S. Sinha, and L. Torresani. "Leveraging Structure from Motion to Learn Discriminative Codebooks for Scalable Landmark Classification," *IEEE Conference On Computer Vision and Pattern Recognition, CVPR*, pages 763-770, 2013.
Acceptance rate = 25.2%.
http://dx.doi.org/10.1109/CVPR.2013.104

[C20] C. Fang, and L. Torresani. "Measuring Image Distances via Embedding in a Semantic Manifold," *European Conference On Computer Vision, ECCV*, pages 402-415, 2012.
Acceptance rate = 27%.
http://dx.doi.org/10.1007/978-3-642-33765-9_29

[C19] A. Bergamo, and L. Torresani. "Meta-Class Features for Large-Scale Object Categorization on a Budget," *IEEE Conference On Computer Vision and Pattern Recognition, CVPR*, 3085-3092, 2012.
Acceptance rate = 24%.
http://dx.doi.org/10.1109/CVPR.2012.6248040

[C18] T. Wang, G. Cardone, A. Corradi, L. Torresani,Êand A.T. Campbell, "WalkSafe: A Pedestrian Safety App for Mobile Phone Users Who Walk and Talk While Crossing Roads," *Twelfth Workshop on Mobile Computing Systems and Applications, HotMobile*, (5):1-6, 2012.
Acceptance rate = 20.6%.
http://dx.doi.org/10.1145/2162081.2162089

[C17] A. Bergamo, L. Torresani, and A. Fitzgibbon. "PiCoDes: Learning a Compact Code for Novel-Category Recognition," *Advances in Neural Information Processing Systems, NIPS*, pages 2088-2096, 2011.
Acceptance rate = 21.8%.
http://books.nips.cc/papers/files/nips24/NIPS2011_1175.pdf

[C16] M. Rastegari, C. Fang, and L. Torresani. "Scalable Object-Class Retrieval with Approximate and Top-k Ranking," *IEEE International Conference on Computer Vision, ICCV*, pages 2659-2666, 2011.
Acceptance rate = 24%.
http://dx.doi.org/10.1109/ICCV.2011.6126556

[C15] A. Bergamo, and L. Torresani. "Exploiting weakly-labeled Web images to improve object classification: a domain adaptation approach," *Adv. in Neural Inform. Processing Systems, NIPS*, 181-189, 2010.
Acceptance rate = 24%.
http://books.nips.cc/papers/files/nips23/NIPS2010_0093.pdf

[C14] L. Torresani, M. Szummer, and A. Fitzgibbon. "Efficient Object Category Recognition Using Classemes," *European Conference On Computer Vision, ECCV*, (1):776-789, 2010.

Acceptance rate = 26.2%.
http://dx.doi.org/10.1007/978-3-642-15549-9_56

[C13] A. Shaji, A. Varol, L. Torresani, and P. Fua. "Simultaneous Point Matching and 3D Deformable Surface Reconstruction," *IEEE Conference On Computer Vision and Pattern Recognition, CVPR*, pages 1221-1228, 2010.
Acceptance rate = 22.3%.
http://dx.doi.org/10.1109/CVPR.2010.5539827

[C12] M.H. Nguyen, L. Torresani, F. de la Torre, and C. Rother. "Weakly Supervised Discriminative Localization and Classification: a Joint Learning Process," *IEEE International Conference on Computer Vision, ICCV*, pages 1925-1932, 2009.
Acceptance rate = 23%.
http://dx.doi.org/10.1109/ICCV.2009.5459426

[C11] L. Torresani, M. Szummer, and A. Fitzgibbon. "Learning query-dependent prefilters for scalable image retrieval," *IEEE Conference On Computer Vision and Pattern Recognition, CVPR*, pages 2615-2622, 2009.
Acceptance rate = 26.2%.
http://dx.doi.org/10.1109/CVPR.2009.5206582

[C10] W. Pan, and L. Torresani. "Unsupervised Hierarchical Modeling of Locomotion Styles," *International Conference on Machine Learning, ICML*, (8):785-792, 2009.
Acceptance rate = 25%.
http://dx.doi.org/10.1145/1553374.1553475

[C9] L. Torresani, V. Kolmogorov, and C. Rother. "Feature Correspondence via Graph Matching: Models and Global Optimization," *European Conference On Computer Vision, ECCV*, (2):596-609, 2008.
Acceptance rate = 27.9%.
http://dx.doi.org/10.1007/978-3-540-88688-4_44

[C8] L. Torresani, P. Hackney, and C. Bregler. "Learning Motion Styles Synthesis from Perceptual Observations," *Advances in Neural Information Processing Systems, NIPS*, (19):1393-1400, 2007.
Acceptance rate for oral presentation = 3%.
http://books.nips.cc/papers/files/nips19/NIPS2006_0458.pdf

[C7] L. Torresani, and K.C. Lee. "Large Margin Component Analysis," *Advances in Neural Information Processing Systems, NIPS*, (19):1385-1392, 2007.
Acceptance rate = 24%.
http://books.nips.cc/papers/files/nips19/NIPS2006_0791.pdf

[C6] L. Torresani, and A. Hertzmann. "Automatic Non-Rigid 3D Modeling from Video," *European Conference On Computer Vision, ECCV*, pages 299-312, 2004.
Acceptance rate = 34.2%.
http://dx.doi.org/10.1007/978-3-540-24671-8_24

[C5] L. Torresani, A. Hertzmann, and C. Bregler. "Learning Non-Rigid 3D Shape from 2D Motion," *Advances in Neural Information Processing Systems, NIPS*, 2004.
Acceptance rate = 27.6%.
http://books.nips.cc/papers/files/nips16/NIPS2003_VM09.pdf

[C4] L. Torresani, and C. Bregler. "Space-Time Tracking," *European Conference On Computer Vision, ECCV*, pages 801-812, 2002.
Acceptance rate for oral presentation = 7.5%.
http://dx.doi.org/10.1007/3-540-47969-4_53

[C3] L. Torresani, D. Yang, E. Alexander, and C. Bregler. "Tracking and Modeling Non-Rigid Objects with Rank Constraints," *IEEE Conference On Computer Vision and Pattern Recognition, CVPR*, (1):493-500, 2001.

Best student paper award.
Acceptance rate for oral presentation = 8%.
`http://dx.doi.org/10.1109/CVPR.2001.990515`

[C2] T. Coianiz, and L. Torresani. "Analysis and Encoding of Lip Movements," *Audio and Video Based Biometric Person Authentication*, 1997.

[C1] T. Coianiz, L. Torresani, and B. Caprile, "2D Deformable Models for Visual Speech Analysis," *NATO Advanced Study Institute: Speechreading by Man and Machine*, 1995.

Refereed workshop papers

[W8] S. Agarwal, D. Tran, L. Torresani, and H. Farid, "Deciphering Severely Degraded License Plates," *SPIE Symposium on Electronic Imaging*, 2017.
`http://www.cs.dartmouth.edu/~lorenzo/Papers/spie17.pdf`

[W7] D. Tran, L. Bourdev, R. Fergus, L. Torresani, and M. Paluri, "Deep End2End Voxel2Voxel Prediction," *IEEE Conference on Computer Vision and Pattern Recognition Workshops, CVPR Workshops*, 2016.
`http://dx.doi.org/10.1109/CVPRW.2016.57`

[W6] S. Rodríguez-Vaamonde, L. Torresani, K. Espinosa, and E. Garrote. "Improving tag transfer for image annotation using visual and semantic information," *International Workshop on Content-Based Multimedia Indexing, CBMI*, pages 1-4, 2014.
`http://dx.doi.org/10.1109/CBMI.2014.6849846`

[W5] J. Thompson, M. Casey, and L. Torresani. "Audio stimulus reconstruction from brain activity using multi-source semantic embedding," *3rd NIPS Workshop on Machine Learning and Interpretation in NeuroImaging*, 2013.

[W4] J. Thompson, M. Casey, and L. Torresani. "Musical neurosemantic decoding using online weighted approximate-rank pairwise loss optimization in a joint semantic space," *2nd NIPS Workshop on Machine Learning and Interpretation in NeuroImaging*, 2012.
Acceptance rate for oral presentation = 25%.

[W3] A. Chandrashekar, and L. Torresani. "Learning What Is Where from Unlabeled Images," *NIPS Workshop on Learning Semantics*, 2011.
Acceptance rate for oral presentation = 8.7%.

[W2] A. Bergamo, and L. Torresani. "Learning Image Representations for Efficient Recognition of Novel Classes," *The Learning Workshop*, 2011.
Acceptance rate for oral presentation = 31.7%.

[W1] L. Torresani, P. Hackney, and C. Bregler. "Learning to Synthesize Motion Styles," *Snowbird Learning Workshop*, 2006.
Acceptance rate for oral presentation = 29.9%.

Refereed conference abstract

[S3] L. Tilton, T. Arnold, T. Smits, M. Wevers, M. Williams, L. Torresani, M. Bolonkin, J. Bell, and D. Latsis. "Computer Vision in DH," *Digital Humanities, DH*, 2018.

[S2] J. Thompson, M. Casey, and L. Torresani. "Musical neurosemantic decoding using probabilistic online weighted approximate-rank loss optimization in a joint semantic space," *NIPS Workshop for Women in Machine Learning*, 2012.

[S1] J. Thompson, M. Casey, and L. Torresani. "Predicting Crowdsourced Musical Tags from Brain Activity," *Workshop on Cognitively Based Music Informatics Research*, 2012.

Book chapters

[B2] L. Torresani. "Weakly Supervised Learning," in *Encyclopedia of Computer Vision*, Springer, in press.

[B1] L. Torresani, M. Szummer, and A. Fitzgibbon. "Classemes: a Compact Image Descriptor for Efficient Novel-Class Recognition and Search," in *Registration and Recognition in Images and Videos*, eds. R. Cipolla, S. Battiato, and G. M. Farinella. Springer Berlin Heidelberg. Volume 532, pages 95-111, January 2014.
http://dx.doi.org/10.1007/978-3-642-44907-9_5

Patents

Granted patents

[P8] K. He, H. Wang, M.D. Feiszli, L. Torresani. "Convolutional neural network based on groupwise convolution for efficient video analysis." U.S. Patent 10,984,245. April 20, 2021.

[P7] S.B. Gokturk, B. Sumengen, D. Vu, N. Dalal, D. Yang, X. Lin, A. Khan, M. Shah, D. Anguelov, L. Torresani, and V. Vanhoucke. "System and method for using image analysis and search in E-commerce." U.S. Patent 8,732,030. May 20, 2014.

[P6] S.B. Gokturk, D. Anguelov, L. Torresani, V. Vanhoucke, M. Shah, D. Vu, K.C. Lee. "Techniques for enabling or establishing the use of face recognition algorithms." U.S. Patent 8,385,633. February 26, 2013.

[P5] M. Szummer, A. Fitzgibbon, and L. Torresani. "Content-based information retrieval." U.S. Patent 8,346,800. January 1, 2013.

[P4] S.B. Gokturk, L. Torresani, K.C. Lee, and D. Anguelov. "Computer implemented technique for analyzing images." U.S. Patent 8,233,702. July 31, 2012.

[P3] S.B. Gokturk, B. Sumengen, D. Vu, N. Dalal, D. Yang, X. Lin, A. Khan, M. Shah, D. Anguelov, L. Torresani, and V. Vanhoucke. "System and method for enabling image searching using manual enrichment, classification, and/or segmentation." U.S. Patent 7,660,468. February 9, 2010.

[P2] S.B. Gokturk, B. Sumengen, D. Vu, N. Dalal, D. Yang, X. Lin, A. Khan, M. Shah, D. Anguelov, L. Torresani, and V. Vanhoucke. "System and method for search portions of objects in images and features thereof." U.S. Patent 7,657,126. February 2, 2010.

[P1] S.B. Gokturk, B. Sumengen, D. Vu, N. Dalal, D. Yang, X. Lin, A. Khan, M. Shah, D. Anguelov, L. Torresani, and V. Vanhoucke. "System and method for enabling image recognition and searching of images." U.S. Patent 7,657,100. February 2, 2010.

Invited Talks

- *Learning to Understand Video Through Language.*
  – International Computer Vision Summer School, Sicily, Italy, July 2023.
- *Computer Vision with Sight, Sound, and Speech.*
  – ICCV, Tutorial on Multi-Modality Learning from Videos and Beyond, October 2021.
- *Video Understanding with Language Models.*
  – CVPR, The Eight International Workshop on Egocentric Perception, Interaction and Computing, Epic-KITCHENS, June 2021.
- *Vision using Sight... but also Sound and Speech.*
  – CVPR, The Fourth Multimodal Learning and Applications Workshop, MULA, June 2021.
- *Space-Time Models for Segmentation, Tracking and Recognition in Video.*
  – CVPR, Robust Video Scene Understanding Workshop, RVSU, June 2021.
- *Learning Models for Video Understanding from Sound and Speech.*
  – Google AI, October 2020.

- *Learning Semantic Models from Video with Zero or Few Labels.*
  - MIT-IBM Watson AI Lab, August 2020.
  - MIT CSAIL Vision Seminar, September 2020.
- *Self-Supervised Video Models from Sound and Speech.*
  - CVPR Workshop on Sight and Sound, June 2020.
  - ICML Workshop on Self-Supervision in Audio and Speech, July 2020.
- *Video understanding with Limited Labeled Data.*
  - ICCV Tutorial on Visual Learning with Limited Labeled Data, Seoul, Korea, November 2019.
- *Audio-Visual Learning for Video Understanding.*
  - ICCV Workshop on Multimodal Video Analysis and Moments in Time, Seoul, Korea, November 2019.
  - BMVA Symposium on Video Understanding, London, September 2019.
- *Channel-Separated 3D Networks.*
  - Facebook AI Research, Computer Vision Meeting, Boston, September 2019.
- *Revisiting Spatiotemporal Convolutions for Action Recognition & Self-Supervised Learning for Video Understanding.*
  - CVPR Tutorial on Action Recognition, Long Beach, June 2019.
- *Audio-Visual Learning for Reduced Supervision and Improved Efficiency.*
  - CVPR Workshop on Sight and Sound, Long Beach, June 2019.
- *Action Recognition with Channel Separation and Salient-Clip Sampling.*
  - AI Video Summit, Los Angeles, June 2019.
- *Make your pretext task hard!*
  - Facebook AI Research Conference, New York City, April 2019.
- *Brave New Ideas for Video Understanding.*
  - CVPR Workshop on Video Understanding, Salt Lake City, June 2018.
- *Machine Learning for Development.*
  - US Embassy, Accra, Ghana, May 2018.
- *Deep Learning for Computational Video Understanding.*
  - Ashesi University, Ghana, April 2018.
- *Inpainting for Image Compression & On Connectivity Learning in Multi-Branch Networks.*
  - Google Brain, Montreal, Canada, December 2017.
  - Facebook AI Research (FAIR), Menlo Park, CA, October 2017.
  - Google Brain, Mountain View, CA, September 2017.
- *Deep Spatiotemporal Models for Computational Video Understanding.*
  - Center for Cognitive Neuroscience, Dartmouth College, Workshop on Action Understanding: from Kinematics to Mind, Hanover, NH, August 2017.
- *VideoMCC: a New Benchmark for High-Level Video Understanding & On Spatiotemporal Visual Attention.*
  - Adobe Research, Workshop on the Frontiers of Video Technology, San Jose, CA, July 2017.
- *How can we get deep learning to work (well) on video data?*
  - Systems & Technology Research (STR), Woburn, MA, June 2017.
- *Deep Learning for Image & Video Understanding.*
  - Center for Technology and Behavioral Health, Dartmouth College, Hanover, NH, November 2016.
- *Computer Vision with Big Weakly-Labeled Data.*
  - Mount Holyoke College, South Hadley, MA, February 2016.
  - Dartmouth SIAM (Society of Industrial and Applied Math), Hanover, NH, February 2016.
  - Thayer School of Engineering at Dartmouth, Jones Seminar on Science, Technology, and Society, Hanover, NH, May 2015.
- *Coupled Depth Learning & On 3D ConvNet Features for Video Understanding.*
  - Amazon, Seattle, WA, September 2015.
- *Challenges and Opportunities in Visual Recognition with Big Image Data.*

- University of Utah, School of Computing, March 2014.
- University of Central Florida, Dept. of Electrical Engineering & Computer Science, March 2014.
- Tecnalia, Zamudio, Spain, July 2013.
- University of Southern California, Department of Computer Science, April 2013.
- University of Colorado, Colorado Springs, Computer Science Dept., March 2013.
- *Real-Time Object Class Recognition and Search in Large Databases.*
  - ICCV Workshop, The 1st IEEE Workshop on Large Scale Visual Commerce, Sydney, Australia, December 2013.
- *Search by Visual Content in Gigantic Image Collections.*
  - The Geisel School of Medicine at Dartmouth, Department of Pathology, October 2013.
- *Visual Recognition and Search in Gigantic Image Collections.*
  - NEC Laboratories America, Cupertino, CA, September 2013.
  - Yahoo! Labs, Sunnyvale, CA, September 2013.
  - Xerox Research Center Europe, Grenoble, France, July 2013.
  - Dartmouth SIAM (Society of Industrial and Applied Math), Hanover, NH, May 2013.
- *Learning a Compact Image Code for Efficient Recognition of Novel Classes.*
  - University College London, Computer Science Department, December 2012.
  - NIPS Workshop, Big Data Meets Computer Vision: First International Workshop on Large Scale Visual Recognition and Retrieval, Lake Tahoe, NV, 2012.
  - IBM T. J. Watson Research Center, September 2012.
  - MIT, Department of Brain and Cognitive Sciences, April 2012.
  - Caltech, Computer Science Department, February 2012.
  - University of California San Diego (UCSD), Department of Computer Science and Engineering, February 2012.
  - Google, Los Angeles, February 2012.
  - Boston University, Computer Science Department, February 2012.
  - MIT, Computer Science and Artificial Intelligence Laboratory (CSAIL), January 2012.
- *Efficient Novel-Class Recognition and Search.*
  - International Computer Vision Summer School, Sicily, Italy, July 2011.
- *Learning Image Descriptors for Efficient Recognition of Novel Classes.*
  - Microsoft Research, Redmond, March 2011.
- *Learning Models for Scalable Object Class Recognition and Retrieval.*
  - École Centrale de Paris, Department of Applied Mathematics, May 2011.
  - Carnegie Mellon University, Robotics Institute, February 2011.
  - Microsoft Research New England, January 2011.
  - University of Oxford, Dept. of Engineering Science, United Kingdom, August 2010.
- *Learning models for scalable content-based image search.*
  - New York University, Computer Science Department, February 2010.
  - University of Verona, Department of Computer Science, Italy, September 2009.
- *Text-searching unlabeled photos.*
  - Microsoft Research, Cambridge, United Kingdom, September 2009.
- *Graph Matching for Visual Correspondence.*
  - Ecole Polytechnique Federale de Lausanne, School of Computer Science, Switzerland, Dec. 2008.
  - International Symposium on Markov Random Fields in Computer Vision, Cambridge, United Kingdom, October 2008.
- *Learning Models from Visual Data for 3D Tracking, Recognition, and Animation.*
  - Boston College, Computer Science Department, April 2007.
  - Dartmouth College, Computer Science Department, March 2007.
  - Microsoft Research Cambridge, United Kingdom, March 2007.
  - University of Rochester, Department of Computer Science, March 2007.
  - Brown University, Department of Computer Science, March 2007.

- Duke University, Department of Computer Science, March 2007.
  - University of New Mexico, Department of Computer Science, February 2007.
- *Learning Motion Style Synthesis from Perceptual Observations.*
  - Snowbird Learning Workshop, Snowbird, Utah, April 2006.
- *Learning Models of Human Movement from Video.*
  - Stanford University, Computer Science Dept, April 2005.
  - Honda Research Institute, Mountain View, CA, May 2005.
  - Siemens Corporate Research, Princeton, NJ, April 2005.
  - New York University, Computer Science Department, New York, NY, March 2005.
  - New York Institute of Technology, School of Engineering, New York, March 2005.
  - Sarnoff Corporation, Princeton, NJ, February 2005.
  - Caltech, Computational Vision Group, Pasadena, CA, February 2005.
  - NASA JPL, Machine Vision Group, Pasadena, CA, February 2005.
- *Capturing Motion Models for Animation.*
  - International Conference on Intelligent Software for Dance Notation, Columbus, OH, April 2004.
- *Rank Constraints for Space-Time Tracking and Non-Rigid Structure from Motion.*
  - U.C. Berkeley, Computer Vision Group, Berkeley, CA, April, 2002.
  - VisionLunch, Stanford University, Computer Science Dept., January 2002.

Grants

Awarded funding support

- NIH, NLM Express Research Grants in Biomedical Informatics (R01), *Improving Colorectal Cancer Screening and Risk Assessment through Deep Learning on Medical Images and Records*, co-I. February 2019.
- Facebook, Faculty Research Grant, *Learning Deep Architectures*, PI. April 2018.
- J. William Fulbright Foreign Scholarship Board, Fulbright U.S. Scholar Award, *Machine Learning and Perception for Development*, PI. March 2017.
- Neukom Institute, CompX Faculty Grant, *Deep Learning for Histopathological Characterization of Colorectal Polyps to Improve Colon Cancer Screening*, co-PI. March 2017.
- Intel, Faculty Research Grant, *Deep Image Compression*, PI. November 2015 - January 2017.
- Dartmouth College, Cancer Center Developmental Funds (Pilot Projects), *Automatic detection of small pulmonary nodules on chest CT scans through deep learning*, co-PI. December 2016.
- Facebook, Faculty Research Grant, *ViCom: A New Benchmark for Video Comprehension*, PI. December 2015 - October 2017.
- Knight Foundation, News Challenge Prototype Grant, *Unlocking Film Libraries for Discovery and Search*, PI. June 2016.
- Dartmouth College, Scholarly Innovation and Advancement Award, *Learning Deep Video Understanding from the Human Brain*, PI. March 2016.
- National Science Foundation, *BD Hubs: Collaborative Proposal: NORTHEAST: The Northeast Big Data Innovation Hub*, , Senior Personnel/Steering Advisory Committee/Education Connector Member. October 2015.
- Facebook, Faculty Research Grant, *Deep End2End Voxel2Voxel Prediction*, PI. June 2015.
- Google, Faculty Research Award, *Deep Posejectories for Video Understanding*, PI. August 2014.
- Neukom Institute, CompX Faculty Grant, *Self-taught Object Localization with Deep Networks*, PI. February 2014.
- National Science Foundation, II-EN Computing Research Infrastructure, *GridIron*, co-PI. June 2012.
- Microsoft Research, Gift, PI. January 2011.
- Dartmouth College, Burke Research Initiation Award, *Web-powered Photo Search*, PI. June 2010.
- National Science Foundation, CAREER Award, *Learning Models for Scalable Content-Based Image Retrieval*, PI. April 2010.

- Microsoft Research, Gift, PI. September 2009.

Other awarded grants

- Facebook, PI. November 2016. Gift of (8) "bobcat" servers (HP SL270s Gen8 SE), (32) NVIDIA Tesla K40 GPUs, 1 storage server (HP DL380p G8).
- NVIDIA Corporation, *Deep Posejectories for Video Understanding*, PI. May 2015. Gift of (1) NVIDIA Tesla K40 GPU.
- NVIDIA Corporation, *Learning to Localize Objects from Whole-Image Classifiers*, PI. January 2014. Gift of (2) NVIDIA Tesla K40 GPUs.
- XSEDE, *Visual search in large image collections*, PI. September 2013. Computing access to the Intel Sandy Bridge Cluster (Gordon Computer Cluster).

Mentoring

Postdoctoral research supervised as primary advisor

- Facebook AI, Jue Wang, Postdoctoral Researcher, 2020-2022.
- Facebook AI, Gediminas Bertasius, Postdoctoral Researcher, 2019-2021.
- Facebook AI, Tae-Hyun Oh, Postdoctoral Researcher, 2019-2020.
- Dartmouth College, Loris Bazzani, Postdoctoral Researcher, 2014-2015.
- Dartmouth College, Sravana Reddy, Neukom Postdoctoral Fellow, 2012-2015. Co-advised with Ioana Chitoran, Linguistics.

Graduate thesis research supervised as primary advisor

- Dartmouth College, Maxwell Aladago, Ph.D. student, 2019-present.
- Dartmouth College, Joseph DiPalma, Ph.D. student, 2019-2021. Co-advised with Saeed Hassanpour, Biomedical Data Science.
- Dartmouth College, Yiren Jian, Ph.D. student, 2018-2021.
- Dartmouth College, Jun Han, Ph.D. student, 2017-2019.
- Dartmouth College, Karim Sayed Ahmed, Ph.D. student, 2014-2019.
- Dartmouth College, Mohammad Haris Baig, Ph.D. student, 2012-2017.
- Dartmouth College, Du Tran, Ph.D. student, 2012-2016.
- Dartmouth College, Alessandro Bergamo, Ph.D. student, 2010-2014.
- Dartmouth College, Srivamshi Pittala, MS student, 2013-2014.

Undergraduate thesis research supervised as primary advisor

- Dartmouth College, Jason Wei, Senior Honors Thesis (earned High Honors), 2019-2020.
- Dartmouth College, Bruno Korbar, Senior Honors Thesis (earned High Honors), 2017-2018.
- Dartmouth College, Gediminas Bertasius, Senior Honors Thesis (earned High Honors), 2013-2014.
- Dartmouth College, Piotr Teterwak, Senior Honors Thesis (earned High Honors), 2013-2014.
- Dartmouth College, Cole Ott, Senior Honors Thesis (earned High Honors), 2012.
- Dartmouth College, Robert Eastman, Senior Honors Thesis (earned High Honors), 2010.
- Dartmouth College, Jason Victor, Senior Thesis, 2009.

Graduate research supervised not as primary advisor

- Columbia University, Xudong Lin, Intern Ph.D. student at Facebook AI, 2020.
- University of Texas at Austin, Ruohan Gao, Intern Ph.D. student at Facebook AI, 2019.
- Stony Brook University, Yang Wang, Intern Ph.D. student at Facebook AI, 2019.
- Max Planck Institute for Informatics, Yongqin Xian, Intern Ph.D. student at Facebook AI, 2019.

- University of Pennsylvania, Gediminas Bertasius, Intern Ph.D. student at Facebook AI, 2018.
- Skolkovo Institute of Science and Technology (Skoltech), Russia, Mikhail Shvets, Visiting M.S. student at Dartmouth College, 2016.
- University of Milan, Italy, Claudio Ceruti, Visiting Ph.D. student at Dartmouth College, 2015.
- University of the Basque Country, Spain, Sergio Rodriguez Vaamonde, Visiting Ph.D. student at Dartmouth College, 2012.
- University of Milan, Italy, Alessandro Bergamo, Visiting M.S. student at Dartmouth College, 2010.
- Carnegie Mellon University, Minh Hoai Nguyen, Intern Ph.D. student at Microsoft Research Cambridge, 2008.

Thesis committee membership not as primary advisor

- Stony Brook University, Yang Wang, Ph.D. Thesis, May 2020.
- University of Pennsylvania, Gediminas Bertasius, Ph.D. Thesis, March 2019.
- Boston University, Jianming Zhang, Ph.D. Thesis, August 2016.
- Dartmouth College, Shrirang Mare, Ph.D. Thesis, April 2016.
- Dartmouth College, Li Jiang, MS Thesis, 2015.
- Dartmouth College, Nick Foti, Ph.D. Thesis, 2013.
- Dartmouth College, James Hughes, Ph.D. Thesis, 2013.
- Dartmouth College, Ashok Chandrashekar, Ph.D. Thesis, 2013.
- Dartmouth College, Eric Kee, Ph.D. Thesis, 2013.
- Dartmouth College, Hong Lu, Ph.D. Thesis, 2012.
- Dartmouth College, Xiaobo An, Ph.D. Thesis, 2011.
- Dartmouth College, Emiliano Miluzzo, Ph.D. Thesis, 2011.
- Dartmouth College, William B. Kerr, Ph.D. Thesis, 2011.
- Dartmouth College, Tianyu Wang, M.S. Thesis, 2011.
- Dartmouth College, Joe Cooley, M.S. Thesis, 2011.
- Dartmouth College, Eric Trautman, M.S. Thesis, 2009.
- Dartmouth College, Weihong Wang, Ph.D. Thesis, 2009.
- École Polytechnique Fédérale de Lausanne, Mathieu Salzmann, Ph.D. Thesis, 2008.

Teaching

Course instructor

- Dartmouth College, *Deep Learning* (COSC78/278), Winter 2017, Winter 2019, Winter 2020, Winter 2021.
  I designed and developed this new course on deep learning. The class is designed for both undergraduates and graduate students. Topics include feedforward multilayer neural networks, convolutional neural networks, autoencoders, recurrent neural networks and long short-term memory models. Several lectures are devoted to discuss practical strategies to improve the performance of deep models, such as regularization, data augmentation, pre-training, dropout, multi-task learning and advanced optimization methods. The course also reviews modern applications of deep learning, such as image categorization, speech recognition, natural language processing, and audio analysis.
- Dartmouth College, *Foundations of Applied Computer Science* (COSC11), Spring 2019.
  This course introduces core computational and mathematical techniques for data analysis and physical modeling, foundational to applications including computational biology, computer vision, graphics, machine learning, and robotics. The approaches covered include modeling and optimizing both linear and nonlinear systems, representing and computing with uncertainty, analyzing multi-dimensional data, and sampling from complex domains. The techniques are both grounded in mathematical principles and practically applied to problems from a broad range of areas.
- Ashesi University, Ghana *Machine Learning*, Spring 2018.

As part of my 2018 Fulbright U.S. Scholar Award, I taught Machine Learning to a class of 27 undergraduate students at Ashesi University. This course was not previously offered at Ashesi. Faculty members audited my class in order to learn to teach the class independently in future years.

- Dartmouth College, *Topics in Applied Computer Science - Visual Recognition* (COSC89/189), Winter 2016.
  This is a new course that I created for both undergraduate and graduate students. This course covers visual recognition, i.e., recognition methods applied to images, videos and other forms of visual data. Topics include object categorization, semantic segmentation, boundary detection, face recognition, human pose estimation and action recognition.

- Dartmouth College, *Topics in Applied Computer Science - Deep Learning* (COSC89/189), Spring 2013, Spring 2014, Spring 2015, Spring 2017.
  This is a new course that I created for both undergraduate and graduate students. The objective of this class is to teach students about the recent technical advancements of deep learning, which is a new methodology to train hierarchical machine learning models, such as neural networks. Over the last few months deep learning has produced breakthrough results in many different application areas including speech recognition, image understanding and drug design. In this course students read, present and discuss recently published papers on deep learning. In addition, under my supervision, students work on a term project applying deep learning methods to a problem of their choosing.

- Dartmouth College, *Computer Vision* (COSC83/183), Fall 2010, Fall 2012, Fall 2013, Fall 2015.
  This is a new course that I developed. It reviews computer vision, the art of teaching computers to see. It is intended for both undergraduates and graduate students. Topics include image formation, feature detection, segmentation, 3D reconstruction from multiple views, motion estimation, and object recognition. The course also discusses modern applications of computer vision such as 3D photo tourism, visual search, image stitching, and face recognition.

- Dartmouth College, *Machine Learning and Statistical Data Analysis* (COSC74/174), Spring 2009, Spring 2010, Spring 2011, Spring 2012, Winter 2013, Winter 2014, Fall 2014, Winter 2015, Spring 2016, Fall 2016.
  I designed and developed this new course on statistical modeling and machine learning. The class is designed for both undergraduates and graduate students. Topics include learning theory, supervised and unsupervised machine learning, statistical inference and prediction. The course also reviews modern applications of machine learning such as image categorization, speech recognition, and text processing.

- Dartmouth College, *Numerical and Computational Tools for Applied Science* (COSC70/170), Winter 2010, Winter 2011.
  This course provides a practical and principled coverage of numerical and computational tools of use in many scientific disciplines. The focus is on the analysis and application of numerical methods for linear algebra, optimization, and function approximation. The course also provides an introduction to Matlab, a programming environment for scientific computing. This course is designed for undergraduate and graduate students across the sciences and social sciences.

- Dartmouth College, *Web-Powered Computer Vision* (COSC88/188), Fall 2009.
  This is a course that I created to expose students to the emerging field of data-driven computer vision, including methods that exploit Web image repositories to learn visual recognition models. The course reviews research in this field through paper reading and discussion. Students propose and complete a term project on a topic in this area.

- International Computer Vision Summer School, Sicily, Italy. Efficient Novel-Class Recognition and Search, Summer 2011.
  I was invited to teach a course on efficient object class recognition to a class of about 100 pre-selected attendees, including researchers (both academic and industrial) as well as Ph.D. students and postdocs from top-schools in the world. The course reviews models and techniques in the area of object class recognition, and discusses the challenges in scaling these approaches to operate efficiently in large image collections. The course describes methods that address this requirement by learning compact image descriptors and reviews how data structures and methods from text-retrieval can be adapted to

enable real-time recognition of an object-class in a large image database.

Service

Professional service

- Associate Editor of the IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2019-present.
- Area chair for ICCV (2015, 2017, 2021), CVPR (2015, 2017, 2020, 2022), NeurIPS (2021), ECCV (2020), ACCV (2014).
- Organizer of the Workshop "International Workshop and Challenge on Long-form Video Understanding," IEEE Conference On Computer Vision and Pattern Recognition, CVPR (2021)
- Organizer of the Tutorial "Visual Learning with Limited Labeled Data," IEEE International Conference on Computer Vision, ICCV (2019).
- Organizer of the Workshop "Computer Vision for Global Challenges," IEEE Conference On Computer Vision and Pattern Recognition, CVPR (2019).
- Organizer of the Tutorial "Tutorial on Action Classification and Video Modeling," IEEE Conference On Computer Vision and Pattern Recognition, CVPR (2019).
- Organizer of the "AI Video Summit," Los Angeles, June 2019.
- Organizer of the Workshop "Brave New Ideas for Video Understanding," IEEE Conference On Computer Vision and Pattern Recognition, CVPR (2018).
- Organizer of the Workshop "DeepGlobe: A Challenge for Parsing the Earth through Satellite Images," IEEE Conference On Computer Vision and Pattern Recognition, CVPR (2018).
- Organizer of the Workshop "Multidisciplinary Approaches to Understanding Face Perception," Center for Cognitive Neuroscience, Dartmouth (2018).
- Organizer of the Workshop "Action understanding: from kinematics to mind," Center for Cognitive Neuroscience, Dartmouth (2017).
- Organizer of the Workshop "Large Scale Computer Vision Systems," Neural Information Processing Systems, NIPS (2016).
- Reviewer for IEEE CVPR (2003, 2006, 2009, 2010, 2011, 2012, 2013, 2014, 2016, 2019); IEEE ICCV (2007, 2009, 2011, 2013); ECCV (2008, 2010, 2014, 2018); NIPS (2003, 2010, 2012, 2013, 2016, 2018, 2020); ICML (2015); IEEE TPAMI (2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2016); SIGGRAPH (2003, 2004, 2005, 2010).
- Grant reviewer for National Science Foundation (2009, 2010, 2014, 2015).

University service

- Dartmouth College, Chair of Computer Science Faculty Recruiting Committee (2020-21).
- Dartmouth College, Member of Computer Science Faculty Recruiting Committee (2019-20).
- Dartmouth College, Computer Science Department Ph.D. Admission Committee Member (2009-10, 2010-11, 2011-12, 2012-13, 2018-19).
- Dartmouth College, Chair of Computer Science Faculty Recruiting Committee (2016-17).
- Dartmouth College, Chair of Committee on Senior Fellowships, (2016-17).
- Dartmouth College, Cognitive Science Program Steering Committee Member (2014-2017).
- Dartmouth College, Computer Science PhD Program Adviser, (2015-2017).
- Dartmouth College, Member of Committee on Senior Fellowships, (2015-2016).
- Dartmouth College, Thomas E. Kurtz Faculty Chair Recruiting Committee Member (2014-15).
- Dartmouth College, Department Faculty Recruiting Committee Member (2011-12, 2012-13, 2013-14, 2014-15).
- Dartmouth College, Chair of Kemeny Prize Committee (2012-13, 2013-14, 2014-15).
- Dartmouth College, Chair of Computer Science Cluster Committee (2014-15).
- Dartmouth College, Chair of Computer Science Colloquium (2012-13, 2013-14).

- Dartmouth College, Kemeny Prize Committee Member (2010-11, 2011-12).