

Curriculum Vitae

Ivan Laptev

Born on July 3, 1974 in Leningrad

Address: MBZUAI, Masdar City, Abu Dhabi, United Arab Emirates

E-mail: Ivan.Laptev@mbzuai.ac.ae

Web: <http://www.di.ens.fr/~laptev>

Degrees

- Habilitation à diriger des recherches (HDR), École normale supérieureRoyal (ENS), France, 2013, Dissertation: “Modeling and visual recognition of human actions and interactions”,
- PhD in Computer Science, Royal Institute of Technology (KTH), Sweden, 2004, Dissertation: “Local Spatio-Temporal Image Features for Motion Interpretation”,
- MSc in Computer Science, Royal Institute of Technology (KTH), Sweden, 1997.

Academic positions

- Professor of Computer Vision at [MBZUAI](#), Abu Dhabi, UAE, 2023—.
- Head of [INRIA/ENS Willow team](#), Paris, France, 2021-2023.
- Research director, INRIA/ENS Willow team, Paris, France, 2013-2023.
- Research scientist, INRIA/ENS Willow team, Paris, France, 2009-2013.
- Research scientist, INRIA Vista team, Rennes, France, 2005-2009.
- Postdoc, INRIA Vista team, Rennes, France, 2004-2005.
- PhD scholarship, KTH/CVAP, Stockholm, Sweden, 2000-2004.
- Research assistant, TUM, Munich, Germany, 1997-1999.

Entrepreneurship

- Co-founder and supervisory board member of [VisionLabs](#).
- Head of research, VisionLabs, 2016-2025.

Awards and distinctions

- Winner of the REVERIE/Soon Challenge, in conjunction with ICCV 2021.
- Helmholtz prize for the ICCV’03 paper ”Space-Time Interest Points”, 2017.
- 1st place in the Google Cloud & YouTube-8M Video Understanding Challenge, 2017.
- ERC Junior grant, 2012.
- INRIA award for outstanding research “Prime d’excellence scientifique” 2010.
- Outstanding review awards at ECCV 2008, CVPR 2009 and ECCV 2010.
- Honorable mention in PASCAL VOC Challenge, 2007.

Supervision

- 30 PhD students: 9 on-going, 23 graduated and pursue successful careers in academic and industrial research labs: DeepMind, Facebook, etc. 2 students created successful startups.
- 12 postdocs, 50+ Master students, 30+ research engineers.

Professional activities

- | | |
|-------------------------------|--|
| Associate
Editor in Chief | <ul style="list-style-type: none"> • IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2025–. |
| General chair | <ul style="list-style-type: none"> • CVF/IEEE International Conference on Computer Vision (ICCV) 2029. |
| Program chair | <ul style="list-style-type: none"> • Asian Conference on Computer Vision (ACCV) 2024. • CVF/IEEE International Conference on Computer Vision (ICCV) 2023. • CVF/IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2018. |
| Editorial board | <ul style="list-style-type: none"> • IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2014-2019. • International Journal of Computer Vision (IJCV), 2010–2022 • Image and Video Computing (IVC), 2009-2015. |
| Area chair | <ul style="list-style-type: none"> • IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2010, 2013, 2015, 2019, 2020, 2021, 2025, 2026. • IEEE International Conference on Computer Vision (ICCV) 2011, 2025. • European Conference on Computer Vision (ECCV) 2012, 2014, 2018, 2020. • Asian Conference on Computer Vision (ACCV), 2014, 2020. • IEEE International Conference on Automatic Face and Gesture Recognition, 2011. |
| Workshop
co-organizer | <ul style="list-style-type: none"> • ICCV'25 Workshop SLoMO: Story-Level Movie Understanding and Audio Description, Honolulu, US, 2025. • CVPR'25 Workshop GRAIL: Generalization in Robotics Manipulation Workshop and Challenges, Nashville, US, 2025. • CVPR'25 Embodied AI Workshop Challenge: Social Mobile Manipulation Challenge, Nashville, US, 2025. • Machines Can See summits, Dubai, UAE, 2023-2025. • Machines Can See summits, Moscow, Russia, 2017-2022. • ICCV'14 THUMOS: The Second International Workshop on Action Recognition with a Large Number of Classes, Zurich, Switzerland, 2014. • ICCV'13 THUMOS: The First International Workshop on Action Recognition with a Large Number of Classes, Sydney, Australia, 2013. • ECCV'12 Workshop on Action Recognition and Pose Estimation in Still Images, Florence, Italy, 2012. • CVPR'11 Workshop on Gesture Recognition, Colorado Springs, US, 2011. |
| Summer School
co-organizer | <ul style="list-style-type: none"> • Visual Recognition and Machine Learning Summer Schools: Grenoble 2010, Paris 2011, Grenoble 2012, Grenoble 2013. |

- Teaching
- Advanced topics in vision and language, MBZUAI, UAE, 2024, 2025.
 - Object recognition and computer vision, ENS, France, 2008-2024.
 - Visual object recognition and localization, SkolTech, Moscow, January 2014, 30h.
 - Visual Recognition: Objects, Actions and Scenes, University of Trento, Italy, July 2014, 20h.
 - Computer vision course, KTH, Sweden, 2003.
 - Teaching assistance for Computer Vision course, KTH, Sweden, 2001-2004.
 - Teaching assistance for Numerical Methods course, KTH, Sweden, 2003.
 - Computer vision course, TUM, Germany 1998-1999.
- Grants
- Korean AI Research Hub, 2024–2029: 450k USD.
 - PRAIRIE Chair, PI, funding in 2019–2023: 600kE.
 - LV-ENS Chair, 2017-2023, PI, 1,2ME.
 - DGA funded project DRAAF, funding in 2018–2020: 140kE.
 - Google Research Award, Co-PI, 2015. Total funding: 80kE.
 - ERC junior grant ACTIVIA, PI. Total funding for 2013–2018: 1,5ME.
 - Google Research Award, Co-PI, 2012. Total funding: 50kE.
 - QUAERO program, funded by OSEO. Leader of the tasks “Motion Recognition” and “Event recognition”. Total funding during 2008–2013: 658kE.
 - EIT-ICT Labs funded activity “Cross-linking Visual Information and Internet Resources using Mobile Networks”. Funding in 2011-2012: 76kE.
 - MSR-INRIA project “Image and Video Mining for Sciences and Humanities”. Funding during 2009-2011: 300kE.
 - DGA funded project CrowdChecker supporting postdoc of Mikel Rodriguez, funding in 2010–2011: 70kE.
 - Technicolor travel and equipment grant during 2009–2011: 15kE.
 - MUSCLE European Network of Excellence 2004–2008.
- Public software and benchmarks
- Software for extracting Space-Time Interest Points (STIP) from video¹. The STIP software has been used in >50 research publications.
 - Software for rapid object detection in images and video.²
 - KTH human action dataset³. The dataset has become a standard benchmark for human action recognition used in >200 research publications.
 - Hollywood-2 human action dataset which is currently considered as one of the most challenging and realistic benchmarks for action recognition.⁴
- Transfer
- Consultant for ProfilTechnology company, Paris, France, 2012.
 - Scientific adviser of VisionLabs, Moscow, Russia 2012—.
 - Scientific adviser of OculusAI start-up, Stockholm, Sweden, 2008-2009.

¹<http://www.irisa.fr/vista/Equipe/People/Laptev/download.html#stip>

²<http://www.irisa.fr/vista/Equipe/People/Laptev/download.html#objectdetection>

³<http://www.nada.kth.se/cvap/actions/>

⁴<http://www.irisa.fr/vista/actions/hollywood2/>

INRIA/ENS

- Member of the postdoc selection committee, INRIA, 2012–2015.
- Member of Conseil de Laboratoire DI, ENS, 2013.

Invited Presentations

Invited
conference
and workshop
presentations

- Korea AI Summit, Seoul, Oct 27, 2025.
- PRCV Keynote, Shanghai, October 15-18, 2025.
- Video AI Symposium, Paris, Sept 16-17, 2025.
- ELLIS Computer Vision Workshop, April 1-2, 2025.
- ICIIP Workshop EMBodied AI: Trends, Challenges, and Opportunities (EMAI), October 28, 2024.
- Korean Conference on Computer Vision (KCCV), August 12-14, 2024.
- CVPR workshop LPVL: Learning from Procedural Videos and Language: What is Next? June 18, 2024.
- CVPR workshop MANGO New Trends in Multimodal Human Action Perception, Understanding and Generation, June 18, 2024.
- YaTalks, Dec. 4, 2023.
- Korea AI Summit Nov 2-3, 2023.
- ICCV Workshop on AI for Creative Video Editing and Understanding, October 2, 2023.
- ELLIS Workshop, May 2023.
- Video Understanding Symposium, Sept. 2022.
- Andrew Zisserman Festschrift, Sept. 2022.
- Journée Visage, Gestes, Actions et Comportement, June 2022.
- Int. Workshop on AI for Visual Computing, Feb. 2022.
- Ecole Navale, Brest, Oct. 2021.
- Int. Conf. on Information Technology and Nanotechnology, Sept. 2021.
- Conseil d'Etat, July 2021.
- RAAI Summer School, July 2021.
- ELLIS Workshop, July 2021.
- CVPR Workshop Long-form Video Understanding, June 2021.
- AI Journay, Moscow, Dec. 2020.
- Compositional and Multimodal Perception ECCV Workshop, Aug. 2020.
- Learning from Unlabeled Videos CVPR Workshop, June 2020.
- AI Video Summit, FAIR, Los Angeles, June 2019.
- ActivityNet CVPR Workshop, Long Beach, June 2019.
- AIST, keynote speaker, Kazan, July 2019.
- ELLIS Workshop, San Sebastian, September 2019.
- BMVA Symposium on Video Understanding, London, Sept. 2019.
- Extreme Vision ICCV Workshop, Seoul, Oct. 2019.
- CoVieW ICCV Workshop, Seoul, Oct. 2019.
- Hands in Action ICCV Workshop, Seoul, Oct. 2019.
- DICTA, keynote speaker, Perth, Dec. 2019.
- ICCVG, keynote speaker, Warsaw, Sept. 2018.

- EPIC Workshop in conjunction with ECCV'18, Munich, Sept. 2018.
- BMVC'18 tutorial speaker, Newcastle, Sept. 2018.
- Workshop on Brave New Ideas for Video Understanding, in conjunction with CVPR'18, Salt Lake City, June, 2018.
- Integrating Vision and Language, Tartu, March 2018.
- 36th Annual Swedish Symposium on Image Analysis, invited speaker, Stockholm, March 2018.
- Chalearn Workshop on Action, Gesture, and Emotion Recognition, Venice, October, 2017.
- The Joint Video and Language Understanding Workshop, Venice, October, 2017.
- ML Day, Pré-GDR IA, Paris, Sept. 2017.
- Frontiers of Video Technology workshop, Adobe, July, 2017.
- Workshop on YouTube-8M Large-Scale Video Understanding, Honolulu, July, 2017.
- Workshop on Visual Understanding Across Modalities, Honolulu, July, 2017.
- Iberian Conference on Pattern Recognition and Image Analysis, Faro, June, 2017.
- Paris ML Meetup Spatio-temporal Series Hackathon, Paris, February, 2017.
- Skolkovo Robotics, Moscow, May 2016.
- Deep Machine Intelligence and its Applications, SkolTech, Moscow, June 2016.
- Workshop on Brave new ideas for motion representations in videos, Amsterdam, October, 2016.
- Open Day AI Innovation Factory, December, 2016.
- Deep Video Workshop, Santa Cruz, USA, Nov. 2015.
- EHES, Paris, France, Nov. 2015.
- NCCV, (Keynote Speaker) Sept. 14-15, Lunteren, The Netherlands, 2015.
- DALI Workshop, La Palma, Spain, April 2015.
- GDR-ISIS, Paris, France, March, 2015.
- CVPR'15 Area Chair Workshop, Boston, MA, USA, March 2015.
- Deep Learning Summit, San Francisco, USA, Jan. 2015.
- ICVGIP, (Plenary Speaker) Dec. 14-17, Bangalore, India, 2014.
- ECCV'14 Area Chair Workshop, Zurich, Switzerland, June 2014.
- Workshop on Perceptual Organization, CVPR'14, Columbus, June 2014.
- Computer Vision Winter Workshop, Křtiny, Czech Republic, Feb. 2014.
- 1st Workshop on Understanding Human Activities, Sydney, Australia, Dec. 2013.
- IEEE Workshop on Decoding Subtle Cues from Social Interactions, Sydney, Australia, Dec. 2013.

- Int. Workshop on Action Similarity in Unconstrained Videos, Portland, Oregon, USA, June 2013.
 - Japanese-French Frontiers of Science Symposium, Kyoto, Japan, Jan. 2013.
 - 3rd IST Austria Symposium on Computer Vision and Machine Learning, Vienna, Austria, Oct. 2012.
 - First Croatian Workshop on Computer Vision, Zagreb, Croatia, Sept. 2012.
 - 10th Workshop on Content Based Multimedia Indexing, Annecy, France June 2012.
 - 3rd AFCV International Workshop on Recent Trends in Computer Vision, Osaka, Japan, Jan. 2012.
 - ICCV'11 International Workshop on Video Event Categorisation, Barcelona, Spain, Nov. 2011.
 - Assemblée generale du GdR ISIS, Saint-Georges-de-Didonne, France, May 2011.
 - GDR-ISIS scientific meeting, Paris, France, December 2010
 - ECCV'10 Workshop on Sign, Gesture and Activity, Grece, Sept. 2010.
 - ICPR'10 Workshop on Analysis and Evaluation of Large-Scale Multimedia Collections, Istanbul, Turkey, Aug. 2010.
 - ICPR'10 Workshop on Human Behaviour Understanding, Istanbul, Turkey, Aug. 2010.
 - International Workshop on Frontiers of Activity Recognition, Los Angeles, USA, June 2010.
 - SIBGRAPI keynote talk, Rio de Janeiro, Brazil, Oct. 2009.
 - Workshop on Trends in Computer Vision, Prague, July 2009.
 - Int. Workshop on Video, Barcelona, May 2009.
 - Int. Workshop on Object Recognition, Lake Como, Italy, May 2008.
 - MUSCLE conference, Cannes, France, Feb. 2008.
 - Techno-Vision meeting keynote speaker, Paris, France, July 2007.
 - The Rank Prize Funds, Windermere, UK, July 2007.
 - Workshop on category-level object recognition, Siracusa, Italy, Sep. 2006.
 - Workshop on Computational Vision, Rosenö, Sweden, July 2003.
- Tutorials and courses
- International Summer School on Machine Vision, Padova, Italy, Sept. 4-8, 2023.
 - Tutorial on Visual object recognition and localization, 9th Russian Summer School in Information Retrieval (RuSSIR), St Peresburg, August 2015.
 - CVPR'14 Tutorial on Emerging Topics in Human Activity Recognition, Columbus, Ohio, USA, June 2014.
 - Tutorial on human action recognition, INRIA Visual Recognition and Machine Learning Summer School, Paris, July 2013.
 - Tutorial on human action recognition, INRIA Visual Recognition and Machine Learning Summer School, Grenoble, July 2012.
 - Tutorial on human action recognition, Human Activity and Vision Summer School, Sophia-Antipolis, France, Oct. 2012.

- Tutorial on human action recognition, AERFAI Summer School on pattern recognition in multimodal human interaction, Vigo, Spain, June 2012.
- Tutorial on human action recognition, Microsoft Computer Vision School, Moscow, Russia, Aug. 2011.
- Tutorial on human action recognition, International Computer Vision Summer School, Sicily, Italy, July 2011.
- Tutorial on human action recognition, INRIA/ENS Visual Recognition and Machine Learning Summer School, Paris, July 2011.
- Tutorial on human action recognition, Computer Vision and Machine Learning Winter School, ENS Lyon, Jan. 2011.
- ECCV'10 Tutorial on Statistical and Structural Recognition of Human Actions, Heraclion, Crete, Grece, Sept. 2010.
- Tutorial on human action recognition, INRIA Visual Recognition and Machine Learning Summer School, Grenoble, July 2010.
- Tutorial on human action recognition, AERFAI Summer School on Pattern Recognition and Machine Learning in Multimedia Systems, Benicássim, Spain, June 2010.
- Short course on human motion analysis, Summer School on Machine Learning, Statistics and Computer Vision, Ezhou, China, July 2008.
- Series of lectures on scale-space theory, matching and recognition of static images and video sequences, INRIA, Rennes, France, 2005.
- POSTECH, Pohang, August 2024.
- TUM Munich AI Lectures, June 2024.
- Seoul National University, November 2023.
- MBZUAI, March 2023.
- INRIA Grenoble, Sept. 2021.
- SkolTech, Moscow, Dec. 2019.
- Qualcomm-UvA Deep Learning Seminars, Amsterdam, Sept. 2019.
- Journee AI, SAFRAN, Paris, June 2018.
- KTH, Stockholm, Sept. 2019.
- INRIA Rennes, December, 2017.
- Georgia Institute of Technology, Atlanta, September, 2016.
- MailRu, Moscow, May 2016.
- University of Central Florida, Orlando, September, 2016.
- Amazon, Seattle, USA, Nov. 2015.
- Institute for Computer Graphics and Vision, TU Graz, Austria, July, 2015.
- Bauman Moscow State Technical University, Moscow, Russia, April 2015.
- INRIA Sophia-Antipolis, Dec. 5, 2014.
- KU Leuven, Belgium, November 24, 2014.
- Steklov Institute of Mathematics, Saint Petersburg, Russia, Nov. 17, 2014.
- Univ. of Trento, Trento, Italy, July. 2014.

Invited
presentations at
universities and
companies

- IDIAP, Martigny, Switzerland, Feb. 2014.
- EPFL, Lausanne, Switzerland, Feb. 2014.
- SkolTech, Moscow, Russia, Jan. 2014.
- KTH, Stockholm, Sweden, Jan. 2014.
- Technicolor, Paris, France, June 2013.
- TU Darmstadt, Darmstadt, Germany, May 2013.
- Xerox Research Centre Europe, Grenoble, France, March 2013.
- University of Washington, Seattle, USA, Feb. 2013.
- CMU, Pittsburgh, USA, Nov. 2012.
- Oxford Univ. Oxford, UK, Sept. 2012.
- KTH, Stockholm, Sweden, Dec. 2011.
- University of Luxembourg, Luxembourg, Nov. 2010.
- Google, Mountain View, USA, Aug. 2010.
- GIPSA-Lab, Grenoble, Nov. 2008.
- Intel, Beijing, China, Dec. 2007.
- LIAMA, Beijing, China, Dec. 2007.
- INRIA/LEAR, Grenoble, France, May 2007.
- Microsoft Research, Beijing, China, Jan. 2007.
- INRIA/ORION, Sophia Antipolis, France, Nov. 2005.
- Microsoft Research, Cambridge, UK, Aug. 2004.
- Technical University of Prague, Nov. 2003.

Publications

I have published over 150 technical papers most of which appeared in international journals and major peer-reviewed conferences. The leading conferences in computer vision (ICCV, ECCV, CVPR) and machine learning (NIPS) have a low acceptance rate typically below 25%, and publications in their proceedings are considered as important as journal publications. The top journals of the field are the International Journal of Computer Vision (IJCV) and the IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI). Overall, my publications have **over 56,600 citations** and my **h-index is 83** (both obtained from [Google Scholar](https://scholar.google.com/)). Most of my publications are available from <http://www.di.ens.fr/~laptev>.

Journal articles

1. T. Soucek, J.-B. Alayrac, A. Miech, I. Laptev and J. Sivic (2024). Multi-task learning of object states and state-modifying actions from web videos, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, (PAMI) 46(7): 5114–5130.
2. Q. Le Lidec, W. Jallet, L. Montaut, I. Laptev, C. Schmid and J. Carpentier (2024). Contact Models in Robotics: A Comparative Analysis, *IEEE Transactions on Robotics*, 40: 3716–3733
3. Q. Le Lidec, F. Schramm, L. Montaut, C. Schmid, I. Laptev and J. Carpentier (2024).

- Leveraging Randomized Smoothing for Optimal Control of Nonsmooth Dynamical Systems, *Nonlinear Analysis: Hybrid Systems*, 52, pp. 101468.
4. Z. Li, J. Sedlar, J. Carpentier, I. Laptev, N. Mansard and J. Sivic (2022). Estimating 3D Motion and Forces of Human-Object Interactions from Internet Videos, *International Journal of Computer Vision (IJCV)*.
 5. G. Varol, I. Laptev, C. Schmid and A. Zisserman (2021). Synthetic Humans for Action Recognition from Unseen Viewpoints, *International Journal of Computer Vision (IJCV)*.
 6. Q. Le Lidec, I. Kalevatykh, I. Laptev, C. Schmid, and J. Carpentier (2021). Differentiable simulation for physical system identification, *IEEE Robotics and Automation Letters (RAL)*.
 7. M. Tapaswi, V. Kumar and I. Laptev (2021). Long term spatio-temporal modeling for action detection, *Computer Vision and Image Understanding (CVIU)*.
 8. Y. Labbé, S. Zagoruyko, I. Kalevatykh, I. Laptev, J. Carpentier, M. Aubry and J. Sivic (2020). Monte-Carlo Tree Search for Efficient Visually Guided Rearrangement Planning, *IEEE Robotics and Automation Letters (RAL)*, Vol. 5, No. 2, April 2020.
 9. J.B. Alayrac, P. Bojanowski, N. Agrawal, J. Sivic, I. Laptev and S. Lacoste-Julien (2018). Learning from Narrated Instruction Videos *IEEE Trans. on Pattern Analysis and Machine Intelligence*, (PAMI) 40(9): 2194–2208.
 10. G. Varol, I. Laptev and C. Schmid (2017). Long-term Temporal Convolutions for Action Recognition *IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI)*, 40(6): 1510—1517.
 11. H. Idrees, A.R. Zamir, Y.-G. Jiang, A. Gorban, I. Laptev, R. Sukthankar and M. Shah (2017). The THUMOS Challenge on Action Recognition for Videos ”in the Wild” *Computer Vision and Image Understanding (CVIU)*, 155, pp.1–23.
 12. G. Seguin, K. Alahari, J. Sivic and I. Laptev (2015). Pose Estimation and Segmentation of Multiple People in Stereoscopic Movies *IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI)*, 37(8):1643–1655.
 13. Fouhey, D., Delaitre, V., Gupta, A., Efros A., Laptev I. and Sivic, J. (2014). People Watching: Human Actions as a Cue for Single View Geometry, *International Journal of Computer Vision (IJCV)*, 110(3):259-274.
 14. Junejo, I., Dexter, E., Laptev, I., Pérez, P. (2010). View-Independent Action Recognition from Temporal Self-Similarities, *IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI)*, 33(1):172–185.
 15. Laptev, I. (2009). Improving Object Detection with Boosted Histograms, *Image and Vision Computing (IVC)*, 27: 535–544.
 16. Laptev, I., Caputo, B., Schüldt, C. and Lindeberg, T. (2007). Local Velocity-Adapted Motion Events for Spatio-Temporal Recognition, *Computer Vision and Image Understanding (CVIU)*, (108): 207–229.

17. Laptev, I. (2005). On Space-Time Interest Points, *International Journal of Computer Vision (IJCV)*, 64(2/3): 107–123.
18. Laptev, I. and Lindeberg, T. (2004). Velocity-adaptation of spatio-temporal receptive fields for direct recognition of activities: An experimental study, *Image and Vision Computing (IVC)*, 22(2): 105–116.
19. Laptev, I. and Lindeberg, T. (2003). A distance measure and a feature likelihood map concept for scale-invariant model matching, *International Journal of Computer Vision (IJCV)*, 52(2/3): 97–120.
20. Hellwich, O., Laptev I., and Mayer H. (2002). Extraction of linear objects from interferometric SAR data, *International Journal of Remote Sensing* 23(3): 461–475.
21. Laptev, I., Mayer, H., Lindeberg, T., Eckstein, W., Steger, C. and Baumgartner, A. (2000). Automatic extraction of roads from aerial images based on scale-space and snakes, *Machine Vision and Applications (MVA)*, (12): 23–31.

Refereed international conferences

1. A. Costanzino, W. Bayliss, J. Sock, M. G. Blanch, D. Horak, I. Laptev, P. Torr and F. Pizzati (2025). Towards Reliable Identification of Diffusion-based Image Manipulations, *Proc. Neural Information Processing Systems (NeurIPS)*.
2. K. Kumar, R. M. Anwer, F. S. Khan, S. Khan, I. Laptev, and H. Cholakkal (2025). DEFT: Decompositional Efficient Fine-Tuning for Text-to-Image Models, *Proc. Neural Information Processing Systems (NeurIPS)*.
3. L. Ma, J. Wen, M. Lin, R. Xu, X. Liang, B. Lin, J. Ma, Y. Wang, Z. Wei, H. Lin, M. Han, M. Cao, B. Chen, I. Laptev and X. Liang (2025). PhyBlock: A Progressive Benchmark for Physical Understanding and Planning via 3D Block Assembly, *Proc. Neural Information Processing Systems, Dataset track (NeurIPS)*.
4. H. Singh, R.J. Das, M. Han, P. Nakov and I. Laptev (2025). MALMM: Multi-agent large language models for zero-shot robotics manipulation, *Proc. International Conference on Intelligent Robots and Systems (IROS)*.
5. S.R. Motwani, C. Smith, R.J. Das, R. Rafailov, I. Laptev, P.H.S. Torr, F. Pizzati, R. Clark and C.S. de Witt (2025). MALT: Improving reasoning with multi-agent LLM training, *Proc. Conference on Language Modeling (COLM)*.
6. M.E.A. Boudjoghra, I. Laptev and A. Dai (2025). ScanEdit: Hierarchically-Guided Functional 3D Scan Editing, *Proc. CVF/IEEE International Conference on Computer Vision (ICCV)*.
7. M. Han, L. Ma, K. Zhumakhanova, E. Radionova, J. Zhang, X. Chang, X. Liang and I. Laptev (2025). Roomtour3D: Geometry-aware video-instruction tuning for embodied navigation, *Proc. CVF/IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*.

8. T. Soucek, P. Gatti, M. Wray, I. Laptev, D. Damen and J. Sivic (2025). ShowHowTo: Generating scene-conditioned step-by-step visual instructions, *Proc. CVF/IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*.
9. A. Vayani et al. (2025). All languages matter: Evaluating LLMs on culturally diverse 100 languages, *Proc. CVF/IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*.
10. R. Akizhanov, V. Dhédin, M. Khadiv and I. Laptev (2025). Learning feasible transitions for efficient contact planning, *Proc. Learning for Dynamics and Control Conference (L4DC)*.
11. Z. Chen, S. Chen, E. Arlaud, I. Laptev and C. Schmid (2025). ViViDex: Learning Vision-based Dexterous Manipulation from Human Videos, *Proc. IEEE International Conference on Robotics and Automation (ICRA)*.
12. Y. Xing, Y. Li, I. Laptev and S. Lu (2024). Mitigating Object Hallucination via Concentric Causal Attention, *Proc. Neural Information Processing Systems (NeurIPS)*.
13. T. Soucek, D. Damen, M. Wray, I. Laptev and J. Sivic (2024). GenHowTo: Learning to generate actions and state transformations from instructional videos, *Proc. CVF/IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*.
14. A. Ali, G. Gaikov, D. Rybalchenko, A. Chigorin, I. Laptev and S. Zagoruyko (2024). PairD-ETR: Joint Detection and Association of Human Bodies and Faces, *Proc. CVF/IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*.
15. S. Chen, R. Garcia, I. Laptev and C. Schmid (2024). SUGAR: Pre-training 3D Visual Representations for Robotics, *Proc. CVF/IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*.
16. A. Yang, A. Nagrani, I. Laptev, J. Sivic and C. Schmid (2023). Vidchapters-7m: Video chapters at scale Authors, *Proc. Neural Information Processing Systems (NerIPS)*.
17. S. Chen, T. Chabal, I. Laptev and C. Schmid (2023). Object Goal Navigation with Recursive Implicit Maps, *Proc. International Conference on Intelligent Robots and Systems (IROS)*.
18. R. Garcia, R. Strudel, S. Chen, E. Arlaud, I. Laptev and C. Schmid (2023). Robust visual sim-to-real transfer for robotic manipulation, *Proc. International Conference on Intelligent Robots and Systems (IROS)*.
19. M. Futeral, C. Schmid, I. Laptev, B. Sagot, R. Bawden (2023). Tackling ambiguity with images: Improved multimodal machine translation and contrastive evaluation, *Proc. Association for Computational Linguistics (ACL)*.
20. A. Yang, A. Nagrani, P.H. Seo, A. Miech, J. Pont-Tuset, I. Laptev, J. Sivic and C. Schmid (2023). Vid2Seq: Large-scale pretraining of a visual language model for dense video captioning, *Proc. CVF/IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*.
21. Z. Chen, S. Chen, C. Schmid and I. Laptev (2023). gSDF: Geometry-Driven Signed Distance Functions for 3D Hand-Object Reconstruction, *Proc. CVF/IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*.

22. E. Chane-Sane, C. Schmid and I. Laptev (2023). Learning Video-Conditioned Policies for Unseen Manipulation Tasks, *Proc. IEEE International Conference on Robotics and Automation (ICRA)*.
23. Q. Le Lidec, W. Jallet, I. Laptev, C. Schmid and J. Carpentier (2023). Enforcing the consensus between Trajectory Optimization and Policy Learning for precise robot control, *Proc. IEEE International Conference on Robotics and Automation (ICRA)*.
24. S. Chen, P.-L. Guhur, M. Tapaswi, C. Schmid and I. Laptev (2022). Language Conditioned Spatial Relation Reasoning for 3D Object Grounding, *Proc. Neural Information Processing Systems (NerIPS)*.
25. A. Yang, A. Miech, J. Sivic, I. Laptev and C. Schmid (2022). Zero-Shot Video Question Answering via Frozen Bidirectional Language Models, *Proc. Neural Information Processing Systems (NerIPS)*.
26. P.-L. Guhur, S. Chen, R. Garcia, M. Tapaswi, I. Laptev and C. Schmid (2022). Instruction-driven history-aware policies for robotic manipulations, *Proc. Conference on Robot Learning (CoRL)*.
27. Z. Chen, Y. Hasson, C. Schmid and I. Laptev (2022). AlignSDF: Pose-Aligned Signed Distance Fields for Hand-Object Reconstruction *Proc. European Conference on Computer Vision (ECCV)*.
28. S. Chen, P.-L. Guhur, M. Tapaswi, C. Schmid and I. Laptev (2022). Learning from Unlabeled 3D Environments for Vision-and-Language Navigation *Proc. European Conference on Computer Vision (ECCV)*.
29. S. Chen, P.-L. Guhur, M. Tapaswi, C. Schmid and I. Laptev (2022). Think Global, Act Local: Dual-scale Graph Transformer for Vision-and-Language Navigation, *Proc. CVF/IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*.
30. A. Yang, A. Miech, J. Sivic, I. Laptev and C. Schmid (2022). TubeDETR: Spatio-Temporal Video Grounding with Transformers, *Proc. CVF/IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*.
31. T. Souček, J.-B. Alayrac, A. Miech, I. Laptev and J. Sivic (2022). Look for the Change: Learning Object States and State-Modifying Actions from Untrimmed Web Videos, *Proc. CVF/IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*.
32. Y. Hasson, G. Varol, I. Laptev and C. Schmid (2021). Towards unconstrained joint hand-object reconstruction from RGB videos, *International Conference on 3D Vision (3DV)*.
33. S. Chen, P.-L. Guhur, C. Schmid and I. Laptev (2021). History Aware Multimodal Transformer for Vision-and-Language Navigation, *Proc. Neural Information Processing Systems (NerIPS)*.
34. Q. Le Lidec, I. Laptev, C. Schmid and J. Carpentier (2021). Differentiable rendering with perturbed optimizers, *Proc. Neural Information Processing Systems (NerIPS)*.

35. A. El-Nouby, H. Touvron, M. Caron, P. Bojanowski, M. Douze, A. Joulin, I. Laptev, N. Neverova, G. Synnaeve, J. Verbeek and H. Jégou (2021). XCiT: Cross-Covariance Image Transformers, *Proc. Neural Information Processing Systems* (NerIPS).
36. R. Strudel, R. Garcia, I. Laptev and C. Schmid (2021) Segmenter: Transformer for Semantic Segmentation, *Proc. CVF/IEEE International Conference on Computer Vision* (ICCV).
37. P.-L. Guhur, M. Tapaswi, S. Chen, I. Laptev and C. Schmid (2021). Airbert: In-domain Pretraining for Vision-and-Language Navigation, *Proc. CVF/IEEE International Conference on Computer Vision* (ICCV).
38. A. Yang, A. Miech, J. Sivic, I. Laptev and C. Schmid (2021). Just Ask: Learning to Answer Questions from Millions of Narrated Videos, *Proc. CVF/IEEE International Conference on Computer Vision* (ICCV).
39. E. Chane-Sane, C. Schmid and I. Laptev (2021). Goal-Conditioned Reinforcement Learning with Imagined Subgoals, *International Conference on Machine Learning* (ICML).
40. A. Miech, J.-B. Alayrac, I. Laptev, J. Sivic and A. Zisserman (2021). Thinking Fast and Slow: Efficient Text-to-Visual Retrieval with Transformers, *Proc. CVF/IEEE Conf. on Computer Vision and Pattern Recognition* (CVPR).
41. R. Strudel, R. Garcia, J. Carpentier, J.-P. Laumond, I. Laptev and C. Schmid (2020). Learning Obstacle Representations for Neural Motion Planning, *Proc. Conference on Robot Learning* (CoRL).
42. V. Petřík, M. Tapaswi, I. Laptev and J. Sivic (2020). Learning Object Manipulation Skills via Approximate State Estimation from Real Videos, *Proc. Conference on Robot Learning* (CoRL).
43. Learning visual policies for building 3D shape categories (2020), A. Pashevich*, I. Kaleytykh*, I. Laptev and C. Schmid; *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems* (IROS), Las Vegas, NV, USA.
44. D. Zhukov, J.-B. Alayrac, I. Laptev and J. Sivic (2020). Learning Actionness via Long-range Temporal Order Verification, *Proc. European Conference on Computer Vision* (ECCV), Glasgow, UK.
45. A. Miech*, J.-B. Alayrac*, L. Smaira, I. Laptev, J. Sivic and A. Zisserman (2020). End-to-End Learning of Visual Representations from Uncurated Instructional Videos, *Proc. IEEE Conf. on Computer Vision and Pattern Recognition* (CVPR), Seattle, WA, USA.
46. Y. Hasson, B. Tekin, F. Bogo, I. Laptev, M. Pollefeys and C. Schmid (2020). Leveraging Photometric Consistency over Time for Sparsely Supervised Hand-Object Reconstruction, *Proc. IEEE Conf. on Computer Vision and Pattern Recognition* (CVPR), Seattle, WA, USA.
47. A. Kukleva, M. Tapaswi and I. Laptev (2020). Learning Interactions and Relationships between Movie Characters, *Proc. IEEE Conf. on Computer Vision and Pattern Recognition* (CVPR), Seattle, WA, USA.

48. H. Doughty, I. Laptev, W. Mayol-Cuevas and D. Damen (2020). Action Modifiers: Learning from Adverbs in Instructional Videos, *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA.
49. R. Strudel*, A. Pashevich*, I. Kalevatykh, I. Laptev, J. Sivic and C. Schmid; Learning to combine primitive skills: A step towards versatile robotic manipulation (2020), *Proc. IEEE International Conference on Robotics and Automation (ICRA)*, Paris, France. (* indicates equal contribution).
50. A. Miech, D. Zhukov, J.-B. Alayrac, M. Tapaswi, I. Laptev and J. Sivic (2019). HowTo100M: Learning a Text-Video Embedding by Watching Hundred Million Narrated Video Clips, *Proc. IEEE International Conference on Computer Vision (ICCV)*, Seoul, South Korea.
51. J. Peyre, J. Sivic, I. Laptev and C. Schmid (2019). Detecting unseen visual relations using analogies, *Proc. IEEE International Conference on Computer Vision (ICCV)*, Seoul, South Korea.
52. A. Pashevich, R. Strudel, I. Kalevatykh, I. Laptev and C. Schmid (2019). Learning to Augment Synthetic Images for Sim2Real Policy Transfer, *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Macau, China.
53. Y. Hasson, G. Varol, D. Tzionas, I. Kalevatykh, M. Black, I. Laptev and C. Schmid (2019). Learning joint reconstruction of hands and manipulated objects (2019). *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA.
54. D. Zhukov, J.-B. Alayrac, R.G. Cinbis, D. Fouhey, I. Laptev and J. Sivic (2019). Cross-task weakly supervised learning from instructional videos, *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA.
55. Z. Li, J. Sedlar, J. Carpentier, I. Laptev, N. Mansard and J. Sivic (2019). Estimating 3D Motion and Forces of Person-Object Interactions from Monocular Video, *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA.
56. S. Kim, M. Seo, I. Laptev, M. Cho and S. Kwak (2019). Deep Metric Learning Beyond Binary Supervision, *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA.
57. G. Cheron*, J.-B. Alayrac*, I. Laptev and C. Schmid (2018). A Flexible Model for Training Action Localization with Varying Levels of Supervision, *Proc. Neural Information Processing Systems (NIPS)*, Montreal, Canada. (* indicates equal contribution).
58. G. Varol, D. Ceylan, B. Russell, J. Yang, E. Yumer, I. Laptev and C. Schmid (2018). BodyNet: Volumetric Inference of 3D Human Body Shapes, *Proc. European Conference on Computer Vision (ECCV)*, Munich, Germany.
59. N. Chinaev, A. Chigorin and I. Laptev (2018). MobileFace: 3D Face Reconstruction with Efficient CNN Regression, *Proc. PEOPLECAP Workshop in conjunction with ECCV'18*, Munich, Germany.

60. J.-B. Alayrac, J. Sivic, I. Laptev and S. Lacoste-Julien (2017). Joint Discovery of Object States and Manipulation Actions, *Proc. IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy.
61. J. Peyre, J. Sivic, I. Laptev and C. Schmid (2017). Weakly-Supervised Learning of Visual Relations, *Proc. IEEE International Conference on Computer Vision*, Venice, Italy.
62. A. Miech, J.-B. Alayrac, P. Bojanowski, I. Laptev and J. Sivic; Learning From Video and Text via Large-Scale Discriminative Clustering, *Proc. IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy.
63. G. Varol, J. Romero, X. Martin, N. Mahmood, M.J. Black, I. Laptev and C. Schmid (2017). Learning from Synthetic Humans, *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, Hawaii, USA.
64. V. Kantorov, M. Oquab, M. Cho and I. Laptev (2016). ContextLocNet: Context-aware deep network models for weakly supervised localization, *Proc. European Conference on Computer Vision (ECCV)*, Amsterdam, The Netherlands.
65. G. Sigurdsson, G. Varol, X. Wang, A. Farhadi, I. Laptev and A. Gupta (2016). Hollywood in homes: Crowdsourcing data collection for activity understanding, *Proc. European Conference on Computer Vision*, Amsterdam, The Netherlands.
66. G. Sigurdsson, O. Russakovsky, A. Farhadi, I. Laptev and A. Gupta (2016). Much ado about time: Exhaustive annotation of temporal data, *Proc. AAAI Conference on Human Computation and Crowdsourcing*, Austin, TX, USA.
67. J.-B. Alayrac, P. Bojanowski, N. Agrawal, J. Sivic, I. Laptev and S. Lacoste-Julien (2016). Unsupervised learning from narrated instruction videos, *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, MA, USA.
68. G. Seguin, P. Bojanowski, R. Lajugie and I. Laptev (2016). Instance-level video segmentation from object tracks, *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, MA, USA.
69. S. Kwak, M. Cho and I. Laptev (2016). Thin-slicing for pose: Learning to understand pose without explicit pose estimation, *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, MA, USA.
70. P. Bojanowski, R. Lajugie, E. Grave, F. Bach, I. Laptev, J. Ponce and C. Schmid (2015). Weakly-Supervised Alignment of Video with Text, *Proc. IEEE Int. Conference on Computer Vision (ICCV)*, Santiago, Chile.
71. T.-H. Vu, A. Osokin and I. Laptev (2015). Context-aware CNNs for person head detection, *Proc. IEEE Int. Conference on Computer Vision (ICCV)*, Santiago, Chile.
72. G. Cheron, I. Laptev and C. Schmid (2015). P-CNN: Pose-based CNN Features for Action Recognition, *Proc. IEEE Int. Conference on Computer Vision (ICCV)*, Santiago, Chile.

73. S. Kwak, M. Cho, I. Laptev, J. Ponce and C. Schmid (2015). Unsupervised Object Discovery and Tracking in Video Collections, *Proc. IEEE Int. Conference on Computer Vision*, Santiago, Chile. **(45 citations)**
74. M. Oquab, L. Bottou, I. Laptev and J. Sivic (2015). Weakly Supervised Object Recognition with Convolutional Neural Networks, *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA.
75. V. Chari, S. Lacoste-Julien, I. Laptev and J. Sivic (2015). On Pairwise Costs for Network Flow Multi-Object Tracking, *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA.
76. T.-H. Vu, C. Olsson, I. Laptev, A. Oliva and J. Sivic (2014). Predicting Actions from Static Scenes, *Proc. European Conference on Computer Vision (CVPR)*, Zurich, Switzerland.
77. P. Bojanowski, R. Lajugie, F. Bach, I. Laptev, J. Ponce, C. Schmid and J. Sivic (2014). Weakly supervised action labeling in videos under ordering constraints, *Proc. European Conference on Computer Vision*, Zurich, Switzerland.
78. M. Oquab, L. Bottou, I. Laptev and J. Sivic (2014). Learning and Transferring Mid-Level Image Representations using Convolutional Neural Networks, *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Columbus, Ohio, USA.
79. V. Kantorov and I. Laptev (2014). Efficient feature extraction, encoding and classification for action recognition, *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Columbus, Ohio, USA.
80. P. Bojanowski, F. Bach, I. Laptev, J. Ponce, C. Schmid and J. Sivic (2013). Finding Actors and Actions in Movies, *Proc. IEEE International Conference on Computer Vision (ICCV)*, Sydney, Australia.
81. G. Seguin K. Alahari, J. Sivic and I. Laptev (2013). Pose Estimation and Segmentation of People in 3D Movies, *Proc. IEEE International Conference on Computer Vision (ICCV)*, Sydney, Australia.
82. Delaitre, V., Fouhey, D.F., Laptev, I., Sivic, J., Gupta, A. and Efros, A. (2012). Scene semantics from long-term observation of people, *Proc. European Conference on Computer Vision (ECCV)*, Florence, Italy.
83. Fouhey, D.F., Delaitre, V., Gupta, A., Efros, A., Laptev, I. and Sivic J. (2012). People watching: Human actions as a cue for single view geometry *Proc. European Conference on Computer Vision (ECCV)*, Florence, Italy.
84. Azizpour, H. and Laptev, I. (2012). Object Detection Using Strongly-Supervised Deformable Part Models, *Proc. European Conference on Computer Vision (ECCV)*, Florence, Italy.
85. Ullah, M.M. and Laptev, I. (2012). Actlets: A novel local representation for human action recognition in video *Proc. IEEE International Conference on Image Processing (ICIP)*, Orlando, Florida, USA.

86. Delaitre, V. and Sivic, J. and Laptev, I. (2011). Learning person-object interactions for action recognition in still images *Proc. Neural Information Processing Systems (NIPS)*, Granada, Spain.
87. Rodriguez, M., Laptev, I., Sivic, J. and Audibert, J.-Y. (2011). Density-aware person detection and tracking in crowds *Proc. IEEE International Conference on Computer Vision (ICCV)*, Barcelona, Spain.
88. Rodriguez, M., Sivic, J., Laptev I. and Audibert, J.-Y. (2011). Data-driven Crowd Analysis in Videos *Proc. IEEE International Conference on Computer Vision (ICCV)*, Barcelona, Spain.
89. Lezama, J., Alahari, K., Sivic, J. and Laptev, I. (2011). Track to the Future: Spatio-temporal Video Segmentation with Long-range Motion Cues *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Colorado Springs, USA.
90. Raja, K., Laptev, I., Pérez, P. and Oisel, L. (2011). Joint pose estimation and action recognition in image graphs *Proc. IEEE International Conference on Image Processing (ICIP)*, Brussels, Belgium.
91. Delaitre, V., Laptev, I. and Sivic, J. (2010). Recognizing human actions in still images: a study of bag-of-features and part-based representations *Proc. British Machine Vision Conference (BMVC)*, Aberystwyth, UK.
92. Ullah, M.M., Parizi S.N. and Laptev, I. (2010). Improving bag-of-features action recognition with non-local cues *Proc. British Machine Vision Conference (BMVC)*, Aberystwyth, UK
93. Duchenne, O., Laptev, I., Sivic, J., Bach, F., and Ponce, J. (2009). Automatic Annotation of Human Actions in Video, *Proc. International Conference on Computer Vision (ICCV)*, Kyoto, Japan.
94. Parizi, S.N., Laptev, I. and Targhi, A.T. (2009). Modeling Image Context using Object Centered Grid *Proc. Digital Image Computing: Techniques and Applications (DICTA)*, pp. 476–483, Melbourne, Australia.
95. Dexter, E., Pérez, P. and Laptev, I. (2009). Multi-view synchronization of human actions and dynamic scenes *Proc. British Machine Vision Conference (BMVC)*, London, UK.
96. Dexter, E., Pérez, P., Laptev, I. and Junejo I.N. (2009). View-independent Video Synchronization from Temporal Self-similarities *Proc. International Conference on Computer Vision Theory and Applications*, pp. 383–391, Lisboa, Portugal.
97. Wang, H., Ullah, M.M., Kläser, A., Laptev I., and Schmid C. (2009). Evaluation of local spatio-temporal features for action recognition, *Proc. British Machine Vision Conference (BMVC)*, London, UK.
98. Marszałek, M., Laptev, I., Schmid, C. (2009). Actions in context, *Proc. Computer Vision and Pattern Recognition (CVPR)*, Miami, US

99. Laptev, I., Marszałek, M., Schmid, C., and Rozenfeld, B. (2008). Learning realistic human actions from movies, *Proc. Computer Vision and Pattern Recognition (CVPR)*, Anchorage, US.
100. Junejo, I., Dexter, E., Laptev, I., Pérez, P. (2008) Cross-View Action Recognition from Temporal Self-Similarities, *Proc. European Conference on Computer Vision (ECCV)*, Marseille, France.
101. Laptev, I. and Pérez, P. (2007). Retrieving actions in movies, *Proc. IEEE International Conference on Computer Vision (ICCV)*, Rio de Janeiro, Brazil.
102. Law-To, J., Chen, L., Joly, A., Laptev, I., Buisson, O., Gouet-Brunet, V., Boujemaa, N. and Stentiford, F. (2007). Video copy detection: a comparative study, *Proc. ACM International Conference on Image and video Retrieval*, Amsterdam, The Netherlands, pp. 371–378.
103. Laptev, I. (2006). Improvements of Object Detection Using Boosted Histograms, *Proc. British Machine Vision Conference (BMVC)*, Edinburgh, UK, pp. III:949–958.
104. Laptev, I., Belongie, S., Pérez, P. and Wills, J. (2005). Periodic Motion Detection and Segmentation via Approximate Sequence Alignment, *Proc. International Conference on Computer Vision (ICCV)*, Beijing, China, pp. I:816–823.
105. Laptev, I. and Lindeberg, T. (2004). Velocity adaptation of space-time interest points, *Proc. International Conference for Pattern Recognition (ICPR)*, Cambridge, UK, pp. I:52–56.
106. Schüldt, C., Laptev, I. and Caputo, B. (2004). Recognizing human actions: a local SVM approach, *Proc. International Conference for Pattern Recognition (ICPR)*, Cambridge, UK, pp. III:32:36.
107. Lindeberg, T., Akbarzadeh, A. and Laptev, I. (2004). Galilean-corrected spatio-temporal interest operators, *Proc. International Conference for Pattern Recognition (ICPR)*, Cambridge, UK, pp. I:57–62.
108. Laptev, I. and Lindeberg, T. (2003). Space-Time Interest Points, *Proc. Ninth International Conference on Computer Vision (ICCV)*, Nice, France, pp. 432–439.
109. Laptev, I. and Lindeberg, T. (2003). Interest point detection and scale selection in space-time, in L. Griffin and M. Lillholm (eds), *Scale-Space'03*, Vol. 2695 of *Lecture Notes in Computer Science*, Springer Verlag, Berlin, pp. 372–387.
110. Bretzner, L., Laptev I., and Lindeberg, T. (2002). Hand gesture recognition using multi-scale colour features, hierarchical models and particle filtering in *Proc. 5th IEEE International Conference on Automatic Face and Gesture Recognition*, Washington D.C., May, pp. 423–428.
111. Laptev, I. and Lindeberg, T. (2001). A multi-scale feature likelihood map for direct evaluation of object hypotheses, in M. Kerckhove (ed.), *Scale-Space'01*, Vol. 2106 of *LNCS*, Springer, pp. 98–110.

112. Laptev, I. and Lindeberg, T. (2001). Tracking of multi-state hand models using particle filtering and a hierarchy of multi-scale image features, *in* M. Kerckhove (ed.), *Scale-Space'01*, Vol. 2106 of *LNCS*, Springer, pp. 63–74.
113. Mayer, H., Laptev, I. and Baumgartner, A. (1998). Multi-Scale and Snakes for Automatic Road Extraction, *Proc. European Conference on Computer Vision (ECCV)*, Vol. I of *Lecture Notes in Computer Science*, Springer Verlag, Berlin, Freiburg, Germany, pp. 720–733.
114. Mayer, H., Laptev, I., Baumgartner, A. and Steger, C. (1997). Automatic road extraction based on multi-scale modeling, context, and snakes *Proc. International Archives of Photogrammetry and Remote Sensing*, 32 (Part 3), pp.106–113.

Other publications

1. A. Miech, I. Laptev and J. Sivic (2017) Learnable pooling with Context Gating for video classification, arXiv preprint arXiv:1706.06905.
2. Klaser, A., Marszałek, M., Laptev, I. and Schmid, C. (2010). Will person detection help bag-of-features action recognition? *INRIA Tech. Report RR-7373*.
3. Cherniavsky, N., Laptev, I., Sivic, J. and Zisserman, A. (2010). Semi-supervised learning of facial attributes in video *First International Workshop on Parts and Attributes* in conjunction with ECCV, Hersonissos, Greece.
4. Benmokhtar, R. and Laptev, I. (2010). INRIA-WILLOW at TRECVID 2010: Surveillance Event Detection *Proc. TREC Video retrieval evaluation*, Gaithersburg, USA.
5. Laptev, I. and Lindeberg, T. (2004). Local descriptors for spatio-temporal recognition, *in Proc. ECCV Workshop on Spatial Coherence for Visual Motion Analysis (SCVMA)*.
6. Laptev, I. and Lindeberg, T. (2002). Velocity-adaptation of spatio-temporal receptive fields for direct recognition of activities: An experimental study, *in Proc. ECCV Workshop on Statistical Methods in Video Processing*, pp. 61–66.
7. Bretzner, L., Laptev, I., Lindeberg, T., Lenman, S. and Sundblad, Y. (2001). A prototype system for computer vision based human computer interaction, *KTH Technical Report ISRN KTH/NA/P-01/09-SE*.
8. Klupsch, M., Lückenhaus, M., Zierl, C., Laptev, I., Bandlow, T., Grimme, M., Kellerer, K., and Schwarzer, F. (1999). Agilo robocuppers: Robocup team description. *in RoboCup-98: Robot Soccer World Cup II*, volume 1604 of *Lecture Notes in Computer Science*. Springer, pp. 446–451.

Co-editor

- Radig, B., Niemann, H., Zhuravlev, Y., Gourevitch, I. and Laptev, I. (Eds.) (1999). Pattern Recognition and Image Understanding, *Proc. of 5th Open German-Russian Workshop*, Sankt Augustin, Infix.

HDR thesis

- Laptev, I. (2013). Modeling and visual recognition of human actions and interactions, Département d'Informatique, Ecole Normale Supérieure (ENS), Paris, France.

PhD thesis

- Laptev, I. (2004). Local Spatio-Temporal Image Features for Motion Interpretation, Department of Numerical Analysis and Computing Science (NADA), Royal Institute of Technology (KTH), S-100 44, Stockholm, Sweden. ISBN 91-7283-793-4

MSc thesis

- Laptev, I. (1997). Road Extraction Based on Snakes and Sophisticated Line Extraction, Department of Numerical Analysis and Computing Science (NADA), Royal Institute of Technology (KTH), S-100 44, Stockholm, Sweden.

Abu Dhabi

November 16, 2025