

Lamberto Ballan, Ph.D.

Curriculum Vitae – Updated: January 1, 2017

Personal details

Place of Birth: Firenze, Italy

Citizenship: Italian

Residence: S. Donato in collina, FI, Italy

Spoken Languages: Italian (*native*), English (*fluent*)

Contact

MICC - Università degli Studi di Firenze

Viale Morgagni 65 - 50134 Firenze (FI)

lamberto.ballan@unifi.it

www.lambertoballan.net

<http://twitter.com/lambertoballan>

Research Interests

- Computer Vision, Multimedia, Pattern Recognition, Machine Learning, Image Processing, A.I.

Research Profile

My research interests lie at the boundary of computer vision and multimedia, specifically focused on exploiting big data for visual recognition problems. The main focus of my current research is on designing learning algorithms that make the most effective use of prior and contextual knowledge in presence of sparse and noisy labels.

I am currently a senior postdoctoral researcher at the University of Florence, Italy, and Stanford University working in the Artificial Intelligence Lab. I am supported by a prestigious Marie Curie Fellowship from the European Commission under the FP7-2013-IOF programme. I received my Ph.D. in computer engineering, multimedia and telecommunication in 2011 from the University of Florence. I also spent a short period at the Signal and Image Processing department at Telecom ParisTech (France) in 2010.

I have published around 40 papers in the most prestigious journals and conferences in computer vision, multimedia and image processing. My papers have been cited more than 1100 times (H-index 18) and I received the best paper award at the ACM Workshop on Social Media 2010. I am the lead organizer of the Web-scale Vision and Social Media workshops (ECCV'12, CVPR'14, ICCV'15, ECCV'16) and I gave tutorials at four international conferences (including ICPR'12, ACM-MM'15, CVPR'16). I have actively been participating in various roles in several EU FP6/FP7 research projects and a number of industry-funded projects in Europe and the United States, raising about 290K euro of research funds.

Academic Appointments

- **University of Florence, Italy** – *Media Integration and Communication Center* 2016/10 – present
Marie Curie Fellow / Senior Postdoctoral Researcher
Working on large-scale visual recognition, deep learning and knowledge transfer; Advisor: Prof. A. Del Bimbo.
- **Stanford University, CA, USA** – *AI Lab, Computer Science Department* 2014/10 – 2016/09
Marie Curie Fellow / Senior Postdoctoral Researcher
Working on large-scale visual recognition, deep learning, visual data mining and knowledge transfer; Advisors: Profs. L. Fei-Fei, S. Savarese, B. Girod.
- **University of Florence, Italy** – *Media Integration and Communication Center* 2011/01 – 2014/09
Postdoctoral Researcher – A.R. L.449/1997 (12 months), A.R. L.240/2010 (33 months)
Working on image annotation, social media analysis and image processing; Advisor: Prof. A. Del Bimbo.

- **TELECOM ParisTech**, France – *Signal and Image Processing Department* 2010/04 – 2010/07
Visiting Scholar
Working on Context-Dependent Kernels for visual recognition; Advisor: Dr. Hichem Sahbi.
 - **University of Florence**, Italy – *Media Integration and Communication Center* 2007/03 – 2010/12
Research Assistant – *A.R. L.449/1997 (9 months), Borsa Dottorato (36 months)*
Working on image/video retrieval and multimedia content analysis; Advisor: Prof. A. Del Bimbo.
-

Education

- **University of Florence**, Italy 2011/04
Ph.D. in Computer Engineering, Multimedia and Telecommunication
Thesis: “Object and event recognition in multimedia archives using local visual features”
Advisors: Profs. A. Del Bimbo (primary), M. Bertini (secondary)
 - **University of Florence**, Italy 2006/12
Laurea Degree (5 years degree, M.S. equivalent) in Computer Engineering
-

Honors, Awards and Scholarships

- **Italian Habilitation** as Associate Professor in *Computer Engineering and Information Processing Systems* (09/H1 – ASN 2013).
 - **Marie Curie IOF**, International Outgoing Fellowship, of the European Commission (call FP7-2013).
Marie Curie individual fellowships are highly prestigious and competitive and are meant to support the best, most promising European researchers (accept. rate 12%).
Project: “EAGLE: Exploiting semantic and social knowledge for visual recognition”.
Advisors: Prof. A. Del Bimbo (Univ. of Florence) and Prof. Fei-Fei Li (Stanford University).
Funding: 272K €, 2014/10 – 2017/09 (Grant No. 623930).
 - Young Researcher Support Grant for attending the 15th IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 2015/11; Funding: 1.5K \$.
 - Certificate of Appreciation for participating in the organization of the 12th European Conference on Computer Vision (ECCV), Firenze, Italy, 2012/10.
 - Research Support Grant by the Tuscany Region, Italy, as recognition for reaching the final round of the EU FP7 Marie Curie IOF selections, 2012/04; Funding: 5K €.
 - Selected among the best students who attended one of the first five editions of the International Computer Vision Summer School (ICVSS) to contribute to a book edited by Springer, 2012/06 [B1].
A total of 15 students have been selected; around 150 students from 40+ countries attend ICVSS each year.
 - **Best Paper Award**: 2nd ACM SIGMM International Workshop on Social Media (WSM), in conjunction with ACM Multimedia 2010 [C16].
 - **MIUR Scholarship** for pursuing a Ph.D. in Computer Engineering, Multimedia and Telecommunication at Università degli Studi di Firenze, 2008/01 – 2010/12; Funding: 41K €.
-

Teaching Experience

- **Adjunct Professor**, Ph.D. School in Information Engineering, Univ. of Florence, Italy
 - “Recognition and Classification of Visual Context” (*short course*) a.y. 2013/2014
 - Teaching responsibility: 8 hours (2 CFU), Enrollment: 15 students*
- **Adjunct Professor**, Bachelor in Industrial Design, School of Architecture, Univ. of Florence
 - “Lab. Disegno Industriale III: Multimedia and web design” a.y. 2010/2011
 - Teaching responsibility: 25 hours (3 CFU), Enrollment: 35 students*
- **Guest Lecturer**, Bachelor in Computer Science, Stanford University, USA
 - “Computer Vision: from 3D Reconstruction to Recognition”, Prof. S. Savarese a.y. 2015/2016
 - Lectures: Convolutional Neural Networks, Object Recognition*
- **Guest Lecturer**, Bachelor in Computer Science, Stanford University, USA
 - “Computer Vision: Foundations and Applications”, Prof. Fei-Fei Li a.y. 2014/2015
 - Lectures: Difference of Gaussians, SIFT*
- **Teaching Assistant**, M.Sc. in Computer Engineering, Univ. of Florence, Italy
 - “Visual and Multimedia Recognition”, Prof. A. Del Bimbo a.y. 2013/2014
 - “Multimedia Databases”, Prof. A. Del Bimbo a.y. 2012/2013
 - “Multimedia Databases”, Prof. A. Del Bimbo a.y. 2011/2012
 - “Multimedia Databases”, Prof. A. Del Bimbo a.y. 2010/2011
 - “Multimedia Databases”, Prof. A. Del Bimbo a.y. 2009/2010
 - “Multimedia Databases”, Prof. A. Del Bimbo a.y. 2008/2009
 - “Multimedia Databases”, Prof. A. Del Bimbo a.y. 2007/2008
 - Teaching responsibility: 15 hours, Enrollment: ~30 (per year)*
- **Teaching Assistant**, Bachelor in Computer Engineering, Univ. of Florence, Italy
 - “Fundamentals of Multimedia Design and Production”, Prof. A. Del Bimbo a.y. 2010/2011
 - “Fundamentals of Multimedia Design and Production”, Prof. A. Del Bimbo a.y. 2009/2010
 - Teaching responsibility: 8 hours, Enrollment: ~50 (per year)*
- **Project Mentor and Lecturer**, Master in Multimedia Content Design, Univ. of Florence, Italy
 - “Laboratory on HTML and Web Design”, a.y. 2003/2004
 - “Laboratory on HTML and Web Design”, a.y. 2002/2003
 - Teaching responsibility: 40 hours, Enrollment: ~30 (per year)*

Advising

- **Ph.D. Students:**
 - Jean-Baptiste Boin, 2016–present (PhD, Stanford University) [C1]
 - Pasquale Coscia, 2016–present (PhD, Second Univ. of Naples) [C3, S2]
 - Francesco Castaldo, 2015 (Visiting PhD at Stanford / Second Univ. of Naples) [C2, C3]
 - Tiberio Uricchio, 2012-2016 (PhD, Univ. of Florence) [A1, A3, C6, C8, C10, C11]
- **Graduate Students (M.Sc. Thesis):**
 - Oliver Groth, 2016–present (Visiting Stud at Stanford / Dresden Univ. of Technology)
 - Ansh Kamil, 2016 (Technical Univ. of Munich)
 - Yinghui Xia, 2015-2016 (Stanford University)
 - Alessio Antonielli, 2014 (Univ. of Florence)
 - Simone Macaluso, 2014 (Univ. of Florence)
 - Claudio Tortorici, 2013 (Univ. of Florence) – *next research assistant at Khalifa University, UAE*
 - Alessio Bazzica, 2012 (Univ. of Florence) – *next PhD student at Delft Univ. of Technology, NL*

- Tiberio Uricchio, 2012 (Univ. of Florence) [C8] – *next PhD student at Univ. of Florence, Italy*
 - Andrea M. Serain, 2012 (Univ. of Florence) [C13]
 - Benito Zaccone, 2012 (Univ. of Florence) [C13] – *next PhD student at Univ. of Florence, Italy*
 - Luca Del Tongo, 2011 (Univ. of Florence) [A5]
 - Salvatore Smeraldo, 2010 (Univ. of Florence)
 - Mirco Nencioni, 2010 (Univ. of Florence)
 - Gianluca Franchi, 2010 (Univ. of Florence)
 - Filippo Amendola, 2009 (Univ. of Florence)
 - Lorenzo Seidenari, 2008 (Univ. of Florence) [C18, C19] – *next PhD stud at Univ. of Florence, Italy*
 - Arjun Jain, 2008 (Visiting Scholar, Univ. of Florence) [C24] – *next PhD at Max Planck Inst., Germany*
- **Undergraduate Students** (B.Sc. Thesis / Research projects):
 - Giulia Fontanini, 2012 (Univ. of Florence)
 - Michael Francalanci, 2011 (Univ. of Florence)
 - Alessio Bazzica, 2008 (Univ. of Florence) [C21]

Funded Research Projects

- **EAGLE**: “Exploiting semAntic and social knowledGe for visual rEcognition” 2014/10 – 2017/09
 Funding: European Commission, Marie Curie IOF-2013 FP7 no. 623930 272K €
 Role: **Principal Investigator**; Advisors: Profs. A. Del Bimbo, Fei-Fei Li
- “Webly-Supervised Deep Learning for Large-Scale Activity Recognition” 2016/06 – 2017/07
 Funding: Bavaria California Technology Center (BaCaTeC), Germany 5.5K €
 Role: **Co-Principal Investigator** with Dr. F. Tombari; Advisors: Profs. N. Navab, S. Savarese

Participation in Research Projects

- “Predicting human motion in complex real world environments” 2015/06 – 2016/09
 Funding: Panasonic, Toyota and MURI, Department of Defense, USA
 Role: **WP leader** (human-space interactions for path prediction); P.I.: Prof. S. Savarese
- **PITAGORA**: “Innovative technologies and processes for airport management” 2013/01 – 2014/07
 Funding: Tuscany Region, Italy, and European Regional Development Fund
 Role: **WP leader** (passenger experience); P.I.: Prof. A. Del Bimbo
- **AQUIS-CH**: “Innovative technologies for cultural heritage fruition” 2012/11 – 2014/09
 Funding: Tuscany Region, Italy, and European Social Fund POR FSE 2007-2013
 Role: **WP leader** (knowledge exploitation in social media); P.I.: Prof. A. Del Bimbo
- **euTV**: “euTV, adaptive media channels” 2011/01 – 2012/06
 Funding: European Commission, IST FP7 no. 226248
 Role: researcher; Principal Investigator: Prof. A. Del Bimbo
- **IM3I**: “Immersive Multimedia Interfaces” 2010/07 – 2010/11
 Funding: European Commission, IST FP7 no. 222267
 Role: researcher; Principal Investigator: Prof. A. Del Bimbo
- **FreeSURF**: “FREE Surveillance in a pRivacy-respectFul way” 2008/05 – 2008/07
 Funding: Italian Ministry of Research and Education, MIUR PRIN
 Role: researcher; Principal Investigator: Prof. A. Del Bimbo

- **VIDI-Video:** “Semantic Video Search With a Large Thesaurus of Concepts” 2007/03 – 2009/12
Funding: European Commission, IST FP6 no. 045547
Role: researcher; Principal Investigator: Prof. A. Del Bimbo

Participation in Technology Transfer Projects

- **STAMAT:** “Social Topics and Media Analysis Tool” 2012/01 – 2012/12
Research grant: Università di Firenze – Telecom Italia s.p.a.
Role: researcher and developer
- “Automatic recognition of advertising trademarks in sports videos” 2007/03 – 2010/02
Research grant: Università di Firenze – Sport System Europe s.r.l.
Role: **technical lead of research group** and developer

International Collaborations and Research Visits

- **Technical University of Munich**, Germany 2016/01 – present
Collaborators: F. Tombari (TUM and Univ. Bologna), N. Navab (TUM)
*I have established a collaboration between Stanford and TUM on webly supervised deep learning for large-scale activity recognition. I am the **co-principal investigator** at Stanford (advised by S. Savarese) and am the co-supervisor of a team of five researchers. This project is also supported with a mini-grant from BaCaTeC.*
- **Stanford University**, Stanford, CA, USA 2014/10 – present
Collaborators: Profs. L. Fei-Fei, S. Savarese, B. Girod
I have been appointed as postdoc in Fei-Fei’s VisionLab, who has been the advisor of my Marie Curie project, in October 2014 (see [C4]). Since the spring of 2015, I also established collaborations with S. Savarese - Computer Vision & Geometry lab (see [C2, C3]), and B. Girod - Image, Video and Multimedia Systems group.
- **Second University of Naples**, Caserta, Italy 2015/06 – present
Collaborator: Prof. F. Palmieri
I started collaborating with F. Palmieri while one of his Ph.D. students (F. Castaldo) was visiting Stanford. Since then, I am co-supervising another Ph.D. student (P. Coscia) on a project on trajectory prediction.
- **University of Amsterdam**, NL, and **Renmin University of China**, Beijing 2014/02 – present
Collaborators: C. Snoek (Univ. Amsterdam), X. Li (Renmin Univ. China)
Ongoing collaboration on image tagging, refinement and retrieval. Our joint work is a huge effort to survey and compare the most relevant previous works on a novel large-scale experimental protocol (see our paper on ACM Comp. Surveys [A1]). We have also co-organized tutorials at ACM-MM’15 and CVPR’16.
- **TELECOM ParisTech**, France 2010/04 – 2012/10
Research visit; working with Dr. H. Sahbi (ParisTech)
I spent three months at TELECOM ParisTech in 2010, working on Context-Dependent Kernels for image matching and retrieval. We collaborated until the end of 2012 and we published a paper on IEEE-TIP [A4].

Professional Activities / Academic Service

- **Journal Editor:**
 - **Lead guest editor**, *Computer Vision and Image Understanding (CVIU)* - Elsevier, S.I. on “Computer Vision and the Web” (scheduled for 2017)
 - **Editorial board member**, *Multimedia Tools and Applications (MTAP)* - Springer 2016/09 – present

- **Workshop Organizer / Chair:**
 - **Lead organizer**, 4th Workshop on Web-scale Vision and Social Media (VSM), in conjunction with ECCV 2016, Amsterdam, The Netherlands
 - 2nd ONR Workshop on Structured Learning for Scene Understanding, 2016, Stanford, CA, USA
 - **Lead organizer**, 3rd Workshop on Web-scale Vision and Social Media (VSM), in conjunction with ICCV 2015, Santiago, Chile
 - **Lead organizer**, 2nd Workshop on Web-scale Vision and Social Media (VSM), in conjunction with CVPR 2014, Columbus, OH, USA
 - **Lead organizer**, 1st Workshop on Web-scale Vision and Social Media (VSM), in conjunction with ECCV 2012, Firenze, Italy
- **Local Arrangement Chair:**
 - IEEE Int'l Workshop on Content-Based Multimedia Indexing (CBMI), 2016
- **Conference Organizing Committee Member:**
 - European Conference on Computer Vision (ECCV), 2012
 - ACM Int'l Conference on Multimedia (ACM-MM), 2010
- **Session Chair:**
 - IEEE Int'l Conference on Multimedia and Expo (ICME), 2013
- **Area Chair / Senior Program Committee Member:**
 - ACM International Conference on Multimedia (ACM-MM), 2017
 - IEEE Winter Conference on Applications of Computer Vision (WACV), 2016
- **Program Committee Member:**
 - IEEE Int'l Conference on Computer Vision and Pattern Recognition (CVPR), 2014, 2015, 2016, 2017
 - IEEE International Conference on Computer Vision (ICCV), 2015
 - European Conference on Computer Vision (ECCV), 2012, 2016
 - ACM International Conference on Multimedia (ACM-MM), 2012, 2014, 2015, 2016
 - ACM International Conference on Multimedia Retrieval (ICMR), 2011, 2015, 2016
 - IEEE Int'l Conference on Multimedia and Expo (ICME), 2013, 2014, 2015, 2016, 2017
 - International Conference on Pattern Recognition (ICPR), 2012, 2014, 2016
 - IEEE Winter Conference on Applications of Computer Vision (WACV), 2017
 - Asian Conference on Computer Vision (ACCV), 2016
 - Int'l Conference on Computer Vision Theory and Applications (VISAPP), 2014, 2015
 - IEEE Int'l Conf on Advanced Video and Signal-Based Surveillance (AVSS), 2017
 - IEEE International Conference on Semantic Computing (ICSC), 2012, 2013, 2014, 2015
 - ACM Multimedia Systems Conference (MMSys), 2013
 - IEEE International Symposium on Multimedia (ISM), 2016
 - ACM Int'l Conf Internet Multimedia Computing and Service (ICIMCS), 2013, 2014
 - Workshop on Large Scale Visual Recognition and Retrieval (BigVision), at CVPR 2015, 2016
 - Workshop on Assistive Computer Vision and Robotics (ACVR), at ECCV 2016
 - Workshop on Big Data Computer Vision, at CVPR 2013
 - Workshop on Analysis and Retrieval of Tracked Events and Motion in Imagery Streams, at ICCV 2011
 - Workshop on Deep Learning for Intelligent Multimedia Analytics (DeLIMMA), at ICME 2017
 - Workshop on Wearable and Ego-vision Systems for Augmented Experience (WEsAX), at ICME 2015
- **Reviewer for International Conferences (selection):**

CVPR 2013; ECCV 2014; BMVC 2013; ICPR 2008–2010; ICIP 2008,2010; ICIAP 2007–2011; ACM-MM 2008–2011; ICME 2008–2011; CIVR/MIR 2007–2010; ICMR 2012–2013; WSDM 2014–2015
- **Reviewer for International Journals (selection):**
 - IEEE Trans. on Pattern Analysis and Machine Intelligence (TPAMI), 2012–2016

- IEEE Transactions on Multimedia (**TMM**), 2010, 2012–2016
- IEEE Transactions on Image Processing (**TIP**), 2016
- IEEE Transactions on Information Forensics and Security (**TIFS**), 2012–2013, 2015–2016
- IEEE Trans. on Circuits and Systems for Video Technology (**TCSVT**), 2011–2013, 2015–2016
- Journal of Machine Learning Research (**JMLR**), 2011
- International Journal of Computer Vision (**IJCV**) - Springer, 2015
- Computer Vision and Image Understanding (**CVIU**) - Elsevier, 2013–2016
- Pattern Recognition (**PR**) - Elsevier, 2008–2014, 2016
- Pattern Recognition Letters (PRL) - Elsevier, 2010–2013, 2016
- Image and Vision Computing - Elsevier, 2014
- Neurocomputing - Elsevier, 2012–2013
- Journal of Visual Communication and Image Representation - Elsevier, 2013–2016
- Signal Processing: Image Communication - Elsevier, 2012
- Multimedia Tools and Applications (MTAP) - Springer, 2009–2014, 2016
- IEEE Multimedia, 2016
- Int'l Journal of Multimedia Information Retrieval (MMIR) - Springer, 2012–2014
- Int'l Journal of Machine Learning and Cybernetics - Springer, 2012
- Pattern Analysis and Applications (PAA) - Springer, 2009–2013, 2016
- **Reviewer for Funding Agencies:**
 - National Institute of Standards and Technology (NIST), USA: project proposal for the USA–Israel Binational Industrial Research and Development (BIRD) Foundation, 2015
- **University Service**, Univ. of Florence, Italy:

– Co-founded and organized the Vision and Multimedia Reading Group	2013/04 – 2014/08
– Founded and organized the MICC Computer Science Seminars event	2008/06 – 2009/04
– Student representative on the School of Engineering's board	2003/04 – 2006/09

Professional Memberships and Affiliations

- CVF Member, The Computer Vision Foundation 2013/03 – present
- GIRPR Member, the Italian chapter of IAPR 2008/09 – present
- ACM Member, Association for Computing Machinery 2008/01 – 2014/07
- IEEE Member, Institute of Electrical and Electronics Engineers 2008/01 – 2013/12

Tutorials at International Conferences and Schools

- “Image Tag Assignment, Refinement and Retrieval”

– IEEE Int'l Conf. on Computer Vision and Pattern Recognition (CVPR)	2016/06
– ACM International Conference on Multimedia (ACM-MM)	2015/10
- “Hands on Advanced Bag-of-Words Models for Visual Object Recognition”

– Int'l Conference on Pattern Recognition (ICPR)	2014/08
– Int'l Conference on Image Analysis and Processing (ICIAP)	2013/09

Invited Talks and Seminars

- *“Image understanding in social media by exploiting social-network metadata”*
– University of Catania, Italy – hosted by Prof. G. M. Farinella 2017/1
- *“Sharing knowledge for large-scale visual recognition”*
– University of Parma, Italy – hosted by Prof. A. Prati 2016/12
– Univ. of California, Santa Cruz, USA – hosted by Prof. R. Manduchi 2016/09
– Google Research, Mountain View, USA – hosted by Dr. R. Sukthankar 2016/08
– Aquifi Inc., Palo Alto, USA – hosted by Dr. C. Dal Mutto 2016/06
- **Panelist**, *“Visual recognition in the big data and deep learning era”*
– Int’l Workshop on Web-scale Vision and Social Media (VSM), at ICCV 2015/12
- **Keynote speech**, *“Data-driven tag refinement and localization in web video”*
– Int’l Workshop on Large-Scale Video Search and Mining (LSVSM), at CVPR 2015/06
- *“Love Thy Neighbors: Image Annotation by Exploiting Image Metadata”*
– Univ. of California, Berkeley, USA – Stanford/UC Berkeley Vision Meeting 2015/05
- *“Social media annotation: from images to videos”*
– Stanford University, USA – hosted by Prof. L. Fei-Fei 2013/07
- *“Video event classification using string kernels”*
– University of Southern California, Los Angeles, USA – hosted by Prof. G. Medioni 2011/04
- *“Automatic logo detection and recognition in sport videos”*
– TELECOM ParisTech, France – hosted by Dr. H. Sahbi 2009/02
– University of Florence, Italy – MICC Computer Science Seminars 2008/07

Outreach Activities

- **SAILORS**: I have been involved in the first two editions of the Stanford Artificial Intelligence Laboratory’s OutReach Summer program. SAILORS is a two-week technically rigorous camp for high school girls and it is intended to increase diversity in the field of AI (<http://sailors.stanford.edu>).
– **Project mentor**, *“Using computer vision to monitor hand hygiene in hospitals”* 2016/07
– **Research instructor**, *“Nearest neighbors for image recognition and classification”* 2015/08

Publicly Available Software and Technology Transfer

- **Jingwei**: *“A framework for evaluating image tag assignment, refinement and retrieval”*
This is an open-source testbed for evaluating methods for image tag assignment, tag refinement and tag-based image retrieval. It is developed as part of our survey effort, aiming to provide a timely reflection of the state-of-the-art (ACM-CSUR’16 [A1]). The source code is on GitHub: <http://github.com/li-xirong/jingwei>.
- **SiftForensic**: *“Copy-move forgery detection using SIFT features”*
This package contains the Matlab implementation of the copy-move detection algorithm presented in IEEE-TIFS’11 [A7]. It is widely used as a baseline for copy-move detection (our related papers have been cited around 400 times). The source code is on GitHub: <http://github.com/lambertoballan/sift-forensic>.
- **Handsonbow**: *“Hands on Advanced Bag-of-Words Models for Visual Recognition”*
I am co-author of the hands-on bag of words practicals, a collection of Matlab exercises introducing fundamental concepts in visual recognition. It is developed as part of our ICIAP/ICPR tutorials, but is also used in classes on computer vision. The source code is on GitHub: <http://github.com/lambertoballan/handsonbow>.

- **LogoFinder:** “Logo detection and recognition software for SportSystem Europe s.r.l.”
A key output of the technology transfer project with SportSystem s.r.l. is an innovative software for automatic logo detection and recognition in videos [C26, C25, C24]. This tool has been used by the company for detecting the appearance of logos in hundreds of hours of broadcast videos to derive statistics for marketing analysis.

Publicly Available Datasets

Over the years, I have collected several datasets that are widely used by the academic community.

- **Stanford Drone Dataset:** Aerial RGB videos + ground truth trajectories from multiple targets (pedestrians, bicyclist, skateboarders, carts, etc.) navigating an outdoor campus to learn all interactions between humans and their surrounding. This dataset is available on: http://cvgl.stanford.edu/projects/uav_data.
- **MICC-CMFD Datasets:** We collected the first large-scale datasets for copy-move forgery detection, i.e. MICC-F200, MICC-F2000 and MICC-F600. Our datasets are widely used as benchmarks for CMFD algorithms. These datasets are available on: <http://www.lambertoballan.net/research/image-forensics>.
- **MICC-Flickr 101 Dataset:** We collected a novel image classification dataset from Flickr. It was conceived with the idea of cloning a reference benchmark as Caltech 101, but using realistic pictures from a social media. This dataset is available on: <https://www.micc.unifi.it/resources/datasets/micc-flickr101>.
- **MICC-Logos:** I collected a very challenging dataset for benchmarking logo recognition and trademark retrieval “in the wild”. This dataset has been collected during my visiting period at TELECOM ParisTech, and is available on: <http://www.lambertoballan.net/research/cdk-logo>.

Publications

Bibliometric indices

- *h*-index: 18 (Google Scholar), 14 (Scopus)
- Number of citations: 1225 (Google Scholar), 645 (Scopus)

Selected publications

These are the **key publications** that are selected to **represent my research** activities:

1. X. Li*, T. Uricchio*, **L. Ballan**, M. Bertini, C. Snoek, A. Del Bimbo. Socializing the semantic gap: A comparative survey on image tag assignment, refinement and retrieval. *ACM Computing Surveys*, 49(1):14:1–14:39, Jul 2016 (* equal contribution)
[DOI:10.1145/2906152 – IF:5.243, Class:1 – Cited by:30]
2. **L. Ballan**, F. Castaldo, A. Alahi, F. Palmieri, S. Savarese. Knowledge transfer for scene-specific motion prediction. *Proc. of European Conference on Computer Vision (ECCV)*, Amsterdam, Netherlands, 2016
[DOI:10.1007/978-3-319-46448-0_42 – Class:A*]
Press and media coverage: MIT Tech. Review, TechCrunch, ABC News, and BBC News (May–July 2016)
3. J. Johnson*, **L. Ballan***, L. Fei-Fei. Love Thy Neighbors: Image Annotation by Exploiting Image Metadata. *Proc. of IEEE International Conference on Computer Vision (ICCV)*, Santiago, Chile, 2015 (* equal contribution)
[DOI:10.1109/ICCV.2015.525 – Class:A* – Cited by:12]

4. **L. Ballan**, M. Bertini, G. Serra, A. Del Bimbo. A data-driven approach for tag refinement and localization in web videos. *Computer Vision and Image Understanding*, 140:58–67, Nov 2015 [DOI:10.1016/j.cviu.2015.05.009 – IF:2.134, Class:1 – Cited by:6]
5. H. Sahbi, **L. Ballan**, G. Serra, A. Del Bimbo. Context-dependent logo matching and recognition. *IEEE Transactions on Image Processing*, 22(3):1018–1031, Mar 2013 [DOI:10.1109/TIP.2012.2226046 – IF:3.111, Class:1 – Cited by:48]
6. **L. Ballan**, M. Bertini, A. Del Bimbo, L. Seidenari, G. Serra. Effective codebooks for human action representation and classification in unconstrained videos. *IEEE Transactions on Multimedia*, 14(4):1234–1245, Aug 2012 [DOI:10.1109/TMM.2012.2191268 – IF:2.536, Class:1 – Cited by:29]
7. I. Amerini, **L. Ballan**, R. Caldelli, A. Del Bimbo, G. Serra. A SIFT-based forensic method for copy-move attack detection and transformation recovery. *IEEE Transactions on Information Forensics and Security*, 6(3):1099–1110, Sep 2011 [DOI:10.1109/TIFS.2011.2129512 – IF:2.441, Class:1 – Cited by:315]
This is the 2nd most cited paper in IEEE-TIFS (published between 2011-2015) according to Scholar & Scopus Press and media coverage: “Ecco come si trucca un’immagine: l’esperto spiega il fotoritocco del Pulitzer”, Repubblica, Jan 2014 – “Cosi si trucca l’immagine”, Corriere della Sera, Jan 2014
8. **L. Ballan**, M. Bertini, A. Del Bimbo, L. Seidenari, and G. Serra. Event detection and recognition for semantic annotation of video. *Multimedia Tools and Applications*, 51(1):279–302, Jan 2011 (Special Issue on Multimedia Surveys by World Experts) [DOI:10.1007/s11042-010-0643-7 – IF:1.331, Class:2 – Cited by:112]
This is the 5th most cited paper in MTAP (published between 2011-2015) according to Google Scholar

Full publications list

— Journal Articles —

- [A1] X. Li*, T. Uricchio*, **L. Ballan**, M. Bertini, C. Snoek, and A. Del Bimbo. Socializing the semantic gap: A comparative survey on image tag assignment, refinement and retrieval. *ACM Comp. Surveys*, 49(1):14:1–14:39, Jul 2016. (* equal contrib.) [DOI:10.1145/2906152] [IF:5.243, Class:1] [Cited by:30].
- [A2] **L. Ballan**, M. Bertini, G. Serra, and A. Del Bimbo. A data-driven approach for tag refinement and localization in web videos. *Computer Vision and Image Understanding*, 140:58–67, Nov 2015. [DOI:10.1016/j.cviu.2015.05.009] [IF:2.134, Class:1].
- [A3] **L. Ballan**, M. Bertini, T. Uricchio, and A. Del Bimbo. Data-driven approaches for social image and video tagging. *Multimedia Tools and Applications*, 74(4):1443–1468, Feb 2015. [DOI:10.1007/s11042-014-1976-4] [IF:1.331, Class:2] [Cited by:11].
- [A4] H. Sahbi, **L. Ballan**, G. Serra, and A. Del Bimbo. Context-dependent logo matching and recognition. *IEEE Transactions on Image Processing*, 22(3):1018–1031, Mar 2013. [DOI:10.1109/TIP.2012.2226046] [IF:3.111, Class:1] [Cited by:48].
- [A5] I. Amerini, **L. Ballan**, R. Caldelli, A. Del Bimbo, L. Del Tongo, and G. Serra. Copy-move forgery detection and localization by means of robust clustering with j-linkage. *Signal Processing: Image Communication*, 28(6):659–669, Jul 2013. [DOI:10.1016/j.image.2013.03.006] [IF:1.602, Class:2] [Cited by:81].
- [A6] **L. Ballan**, M. Bertini, A. Del Bimbo, L. Seidenari, and G. Serra. Effective codebooks for human action representation and classification in unconstrained videos. *IEEE Transactions on Multimedia*, 14(4):1234–1245, Aug 2012. [DOI:10.1109/TMM.2012.2191268] [IF:2.536, Class:1] [Cited by:29].

- [A7] I. Amerini, **L. Ballan**, R. Caldelli, A. Del Bimbo, and G. Serra. A SIFT-based forensic method for copy-move attack detection and transformation recovery. *IEEE Trans. on Information Forensics and Security*, 6(3):1099–1110, Sep 2011. [DOI:10.1109/TIFS.2011.2129512] [IF:2.441, Class:1] [Cited by:315].
- [A8] **L. Ballan**, M. Bertini, A. Del Bimbo, L. Seidenari, and G. Serra. Event detection and recognition for semantic annotation of video. *Multimedia Tools and App.*, 51(1):279–302, Jan 2011. (S.I. on Multimedia Surveys by World Experts) [DOI:10.1007/s11042-010-0643-7] [IF:1.331, Class:2] [Cited by:112].
- [A9] **L. Ballan**, M. Bertini, A. Del Bimbo, and G. Serra. Video annotation and retrieval using ontologies and rule learning. *IEEE MultiMedia*, 17(4):80–88, Oct 2010. [DOI:10.1109/MMUL.2010.4] [IF:1.361, Class:2] [Cited by:56].
- [A10] **L. Ballan**, M. Bertini, A. Del Bimbo, and G. Serra. Video event classification using string kernels. *Multimedia Tools and Applications*, 48(1):69–87, May 2010. (Special Issue on Content Based Multimedia Indexing) [DOI:10.1007/s11042-009-0351-3] [IF:1.331, Class:2] [Cited by:46].
- [A11] **L. Ballan**, M. Bertini, A. Del Bimbo, and G. Serra. Semantic annotation of soccer videos by visual instance clustering and spatial/temporal reasoning in ontologies. *Multimedia Tools and Applications*, 48(2):313–337, Jun 2010. [DOI:10.1007/s11042-009-0342-4] [IF:1.331, Class:2] [Cited by:29].

— International Conferences and Workshops —

- [C1] J.-B. Boin, A. Araujo, **L. Ballan**, and B. Girod. Pose-aware fisher vector aggregation for 3d object retrieval. In *Proc. of IEEE Int'l Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, New Orleans, LA, USA, March 2017. [Class:B].
- [C2] **L. Ballan**, F. Castaldo, A. Alahi, F. Palmieri, and S. Savarese. Knowledge transfer for scene-specific motion prediction. In *Proc. of European Conference on Computer Vision (ECCV)*, Amsterdam, Netherlands, Oct 2016. [Class:A*].
- [C3] P. Coscia, F. Castaldo, F. Palmieri, **L. Ballan**, A. Alahi, and S. Savarese. Point-based path prediction from polar histograms. In *Proc. of International Conference on Information Fusion (FUSION)*, Heidelberg, Germany, Jul 2016.
- [C4] J. Johnson*, **L. Ballan***, and L. Fei-Fei. Love thy neighbors: Image annotation by exploiting image metadata. In *Proc. of IEEE International Conference on Computer Vision (ICCV)*, Santiago, Chile, Dec 2015. (* equal contribution) [Class:A*, Cited by:12].
- [C5] X. Li*, T. Uricchio*, **L. Ballan**, M. Bertini, C. G. M. Snoek, and A. Del Bimbo. Image tag assignment, refinement and retrieval. In *Proc. of ACM International Conference on Multimedia (ACM-MM)*, Brisbane, Australia, Oct 2015. (* equal contribution) [Class:A*].
- [C6] **L. Ballan***, T. Uricchio*, L. Seidenari, and A. Del Bimbo. A cross-media model for automatic image annotation. In *Proc. of ACM International Conference on Multimedia Retrieval (ICMR)*, Glasgow, UK, Apr 2014. (* equal contribution) [Class:B, Cited by:24].
- [C7] **L. Ballan** and L. Seidenari. Hands on advanced bag-of-words models for visual object recognition. In *Proc. of Int'l Conference on Pattern Recognition (ICPR)*, Stockholm, Sweden, Aug 2014. [Tutorial].
- [C8] T. Uricchio, **L. Ballan**, M. Bertini, and A. Del Bimbo. An evaluation of nearest-neighbor methods for tag refinement. In *Proc. of IEEE International Conference on Multimedia & Expo (ICME)*, San Jose, CA, USA, Jul 2013. [Class:B, Cited by:12].
- [C9] G. Serra, T. Alisi, M. Bertini, **L. Ballan**, A. Del Bimbo, L. W. Goix, and C. A. Licciardi. Stamat: A framework for social topics and media analysis. In *Proc. of IEEE International Conference on Multimedia & Expo (ICME)*, San Jose, CA, USA, Jul 2013. [Demo].

- [C10] T. Uricchio, **L. Ballan**, M. Bertini, and A. Del Bimbo. Evaluating temporal information for social image annotation and retrieval. In *Proc. of International Conference on Image Analysis and Processing (ICIAP)*, Napoli, Italy, Sep 2013.
- [C11] **L. Ballan**, M. Bertini, T. Uricchio, and A. Del Bimbo. Social media annotation. In *Proc. of IEEE International Workshop on Content-Based Multimedia Indexing (CBMI)*, Veszprem, Hungary, Jun 2013. [Invited].
- [C12] T. Uricchio, M. Bertini, **L. Ballan**, and A. Del Bimbo. MICC-UNIFI at ImageCLEF 2013 Scalable Concept Image Annotation. In *Proc. of CLEF Workshop*, Valencia, Spain, Sep 2013.
- [C13] **L. Ballan**, M. Bertini, A. Del Bimbo, A. M. Serain, G. Serra, and B. F. Zaccone. Combining generative and discriminative models for classifying social images from 101 object categories. In *Proc. of Int'l Conference on Pattern Recognition (ICPR)*, Tsukuba, Japan, Nov 2012. [Class:B].
- [C14] R. Caldelli, I. Amerini, **L. Ballan**, G. Serra, M. Barni, and A. Costanzo. On the effectiveness of local warping against SIFT-based copy-move detection. In *Proc. of Int'l Symposium on Communications, Control and Signal Proc. (ISCCSP)*, Roma, Italy, May 2012. [Invited, Cited by:11].
- [C15] **L. Ballan**, M. Bertini, A. Del Bimbo, and G. Serra. Enriching and localizing semantic tags in internet videos. In *Proc. of ACM International Conference on Multimedia (MM)*, Scottsdale, AZ, USA, Nov 2011. [Class:A*, Cited by:21].
- [C16] **L. Ballan**, M. Bertini, A. Del Bimbo, M. Meoni, and G. Serra. Tag suggestion and localization in user-generated videos based on social knowledge. In *Proc. of ACM Multimedia International Workshop on Social Media (WSM)*, Firenze, Italy, Oct 2010. [Best Paper Award, Cited by:36].
- [C17] I. Amerini, **L. Ballan**, R. Caldelli, A. Del Bimbo, and G. Serra. Geometric tampering estimation by means of a sift-based forensic analysis. In *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Dallas, TX, USA, Mar 2010. [Class:B, Cited by:62].
- [C18] **L. Ballan**, M. Bertini, A. Del Bimbo, L. Seidenari, and G. Serra. Effective codebooks for human action categorization. In *Proc. of ICCV International Workshop on Video-oriented Object and Event Classification (VOEC)*, Kyoto, Japan, Sep 2009. [Cited by:31].
- [C19] **L. Ballan**, M. Bertini, A. Del Bimbo, L. Seidenari, and G. Serra. Recognizing human actions by fusing spatio-temporal appearance and motion descriptors. In *Proc. of IEEE International Conference on Image Processing (ICIP)*, Cairo, Egypt, Nov 2009. [Class:B, Cited by:37].
- [C20] **L. Ballan**, M. Bertini, A. Del Bimbo, and G. Serra. Video event classification using bag of words and string kernels. In *Proc. of International Conference on Image Analysis and Processing (ICIAP)*, Vietri sul Mare (SA), Italy, Sep 2009. [Cited by:26].
- [C21] **L. Ballan**, A. Bazzica, M. Bertini, A. Del Bimbo, and G. Serra. Deep networks for audio event classification in soccer videos. In *Proc. of IEEE International Conference on Multimedia & Expo (ICME)*, New York, NY, USA, Jun 2009. [Class:B, Cited by:15].
- [C22] **L. Ballan**, M. Bertini, A. Del Bimbo, and G. Serra. Action categorization in soccer videos using string kernels. In *Proc. of IEEE International Workshop on Content-Based Multimedia Indexing (CBMI)*, Chania, Crete, Jun 2009.
- [C23] **L. Ballan**, M. Bertini, A. Del Bimbo, L. Seidenari, and G. Serra. Human action recognition and localization using spatio-temporal descriptors and tracking. In *Proc. of Int'l Workshop on Pattern Recog. & A.I. for Human Behaviour Anal. (PRAI*HBA)*, Reggio Emilia, Italy, Dec 2009.

- [C24] **L. Ballan**, M. Bertini, A. Del Bimbo, and A. Jain. Automatic trademark detection and recognition in sport videos. In *Proc. of IEEE International Conference on Multimedia & Expo (ICME)*, Hannover, Germany, Jun 2008. [Class:B, Cited by:12].
- [C25] **L. Ballan**, M. Bertini, and A. Jain. A system for automatic detection and recognition of advertising trademarks in sports videos. In *Proc. of ACM International Conference on Multimedia (MM)*, Vancouver, BC, Canada, Oct 2008. [Demo, Cited by:16].
- [C26] A. D. Bagdanov, **L. Ballan**, M. Bertini, and A. Del Bimbo. Trademark matching and retrieval in sports video databases. In *Proc. of ACM International Conference on Multimedia Information Retrieval (MIR)*, Augsburg, Germany, Sep 2007. [Class:B, Cited by:74].
- [C27] **L. Ballan**, M. Bertini, A. Del Bimbo, and W. Nunziati. Soccer players identification based on visual local features. In *Proc. of ACM International Conference on Image and Video Retrieval (CIVR)*, Amsterdam, Netherlands, Jul 2007. [Class:B, Cited by:20].
- [C28] **L. Ballan**, M. Bertini, A. Del Bimbo, and W. Nunziati. Automatic detection and recognition of players in soccer videos. In *Proc. of International Conference on Visual Information Systems (VISUAL)*, Shanghai, China, Jun 2007.

— Books and Book Chapters —

- [B1] **L. Ballan**, L. Seidenari, G. Serra, M. Bertini, and A. Del Bimbo. Recognizing human actions by using effective codebooks and tracking. In R. Cipolla, S. Battiato, and G. M. Farinella, editors, *Advanced Topics in Computer Vision*. Springer, 2013. [Invited among **the best students** who attended ICVSS].
- [B2] **L. Ballan**, M. Bertini, A. Del Bimbo, W. Nunziati, and G. Serra. Commercials and trademarks recognition. In Y. Kompatsiaris, B. Meriardo, and S. Lian, editors, *TV Content Analysis: Techniques and Applications*. CRC Press, Mar 2012.

— Under Revision —

- [S1] A. Kapil, C. Rupprecht, **L. Ballan**, F. Tombari, and N. Navab. Learning without prejudice: Avoiding bias in webly-supervised action recognition. In *Under review at a top-tier conference, 2017*. [Class:A*].
- [S2] T. Uricchio, **L. Ballan**, L. Seidenari, and A. Del Bimbo. Automatic image annotation via label transfer in the semantic space. *Pattern Recognition, 2016*. (Under major revision) [IF:3.399, Class:1].
- [S3] P. Coscia, F. Castaldo, F. Palmieri, **L. Ballan**, A. Alahi, and S. Savarese. Long-term trajectory prediction using circular distributions. *Pattern Recognition, 2016*. (Submitted) [IF:3.399, Class:1].

Press and Media Coverage

- Our work at Stanford on understanding human motion and trajectory prediction in complex real world environments has been featured by the Financial Times, MIT Technology Review, TechCrunch, ABC News, and BBC News (May–July 2016). I have contributed leading the task on knowledge transfer for scene-specific motion prediction [C2].
- Our work on detecting image forgeries using computer vision [C17, A7, A5] had received a lot of attention from the italian press. This was a collaboration between the MICC, my group at the Univ. of Florence, and the Image and Signal Processing group at CNIT, Florence.
 - “Ecco come si trucca un’immagine: l’esperto spiega il fotoritocco del Pulitzer” – Repubblica, January 2014
 - “Così si trucca l’immagine” – Corriere della Sera, January 2014

References (former supervisors and collaborators)

- Prof. **Alberto Del Bimbo**, *Full Professor* of Computer Engineering at University of Florence
Via Santa Marta 3, 50137 Firenze, FI, Italy
Phone: +39 055 4796-262, E-mail: alberto.delbimbo@unifi.it
- Prof. **Silvio Savarese**, *Assistant Professor* of Computer Science at Stanford University
353 Serra Mall, Stanford, CA 94305-9040, USA
Phone: +1 650 497-0360, E-mail: ssilvio@stanford.edu
- Prof. **Cees G. M. Snoek**, *Associate Professor* of Computer Science at University of Amsterdam
Science Park 904, 1098 GH Amsterdam, The Netherlands
Phone: +31 0 205257528, E-mail: cgmsnoek@uva.nl
- Prof. **Bernd Girod**, *R. L. and A. S. Hancock Professor* of Electrical Engineering at Stanford University
350 Serra Mall, Stanford, CA 94305-9515, USA
Phone: +1 650 724-6354, E-mail: bgirod@stanford.edu
- Dr. **Hichem Sahbi**, *Associate Professor (CR1 CNRS)* at TELECOM ParisTech
46 Rue Barrault, 75013 Paris, France
Phone: +33 1 45817382, E-mail: hichem.sahbi@telecom-paristech.fr