Machine Learning Genoa center

1 Overview

MaLGa — Machine Learning Genoa center (https://ml.unige.it) is a multidisciplinary research center of the University of Genoa , between mathematics and computer science.

Over the last few years, a growing community of mathematicians and computer scientists from Genoa has increasingly coordinated its activity within a large body of diverse yet connected topics that include Computer vision, Computational Harmonic Analysis, Data Analysis, Statistical Learning and Optimization. The resulting momentum has eventually led to the creation of MaLGa. Its basic goal is to foster fundamental research in each of these areas with the ambition of viewing the various tasks as portions of one single endeavour that sets the interconnection of theoretical understanding and real life problems at center stage. The confluence point of the center, and in fact its main driving force in terms of human and financial resources is Machine Learning, but its most distinctive feature is to be found in the mutual interaction of the broadly diversified research themes and scientific backgrounds. The spectrum of competences, that range from very abstract to concretely applied, is a feature that identifies MaLGa as a bridge between Academia and Industry and thus orients its many educational activities as a combination of scholarly-oriented teaching and professional training.

2 MaLGa vision

Understanding how intelligence works and how it can be emulated by machines is an old dream and arguably one of the biggest challenges in modern science. ML, with its principles and computational implementations, is at the very core of this endeavour. Recently, for the first time, we have been able to develop artificial intelligence systems able to solve complex tasks considered out of reach for decades. ML engines, *trained rather than programmed* to solve a task, enable intelligent technologies such as Siri, Kinect or Google self driving car, to name a few. At the same time, they help deciphering the information in our DNA and make sense of the flood of data gathered on the web, forming the basis of a new ”Science of Data”.

The complexity of these tasks varies significantly, and oftentimes it is necessary to dive deeper into the mathematical underpinnings, including analysis, optimization, probability and statistics; other times, it is important to design efficient algorithms able to function on limited resources, or it is necessary to reach a hands-on understanding of the application scenario that provided the data and motivated the problem, as well as of the specificities of the data themselves. Very often it is crucial to engage an open dialogue with domain experts. For these reasons ML today is a multidisciplinary research field. MaLGa stands at the junction of these different research areas and this is its driving force. In the broad spectrum of ML, MaLGa research focuses specifically on the following topics: Computational and Statistical Learning, Computational Harmonic Analysis, Computer Vision, Machine Learning for Data Science.

A main objective of MaLGa is to grow a new generation of experts in ML related topics, leveraging Italy's strong education system which supports us in raising well-prepared and motivated students and young researchers. We invest in this direction by organising graduate courses, sponsoring talented PhD students and PostDoc, encouraging international exchanges. We believe that education is both a duty and an opportunity of Academia. Every year we organize graduate courses and schools attracting students from Italy and abroad. Some of them are also open to the industrial community, providing a natural way to establish connections and collaborations.

While MaLGa primary objective is basic research in ML, our plan is to impact a variety of different fields and application domains by establishing fruitful collaborations with industry, other research institutions, and the general public. MaLGa provides us a critical mass to interact with the outside world. For this our commitment is strong and diversified in dissemination, continuous education, and technology transfer. The channeling of our research findings toward practical applications provides the opportunity of remaining in constant touch with societal needs and adjusting the focus of our basic research accordingly.

3 MaLGa pillars

From the past to the future. The University of Genoa has a strong tradition in ML and artificial intelligence. The study of machine intelligence was pioneered in Italy by the physicist Antonio Borsellino and the engineer Vincenzo Tagliasco in Genoa. Researchers from this school, like Tomaso Poggio, Vincent Torre, ~~but also~~ Federico Girosi and Alessandro Verri contributed in fundamental ways especially in ML and computer vision. In the last twenty years, the University of Genoa has established itself as a leading group in ML and its applications, in particular to computer vision and computational biology. As a proof of the quality of our research we have been recently awarded an ERC consolidator grant (SLING 819789). The latter has boosted the creation of the MaLGa center, the first research center in Italy whose core is ML. MaLGa allows us to coordinate and expand the ML activity in Genoa. As we argue below, our geographic position, the presence of several engaged institutions and industrial partners, the opportunity to form students by shaping new courses and curricula at the University, are all assets that will allow MaLGa to have long lasting impacts.

From theory to practice: a multidisciplinary effort. MaLGa is primarily a center for fundamental research. A main feature of the research activity at MaLGA is the idea that a fruitful interplay between theoretical and applied questions is the key for long lasting and impactful contributions. More broadly this has always been the founding feature of ML studies at the University of Genoa. Our present and past research has always been equally split between foundational work in ML theory, and applied projects working closely with partners such as hospitals or industries to develop vertical solutions. In particular, the tradition of regularization theory is our natural framework to understand and develop ML algorithms, and we have always been working to channel theoretical advances into practical solutions readily available in a range of different application domains.

MaLGa is a research center between the mathematics and computer science departments in Genoa reflecting our deeply multidisciplinary approach. We next describe how the above general principles are implemented in a consistent research organization.

4 MaLGa structure

Today MaLGa is structured in four groups.

Computational and statistical learning. The goal of this group is to advance the frontiers of learning theory and ML, while building algorithmic tools for the analysis of complex systems and high dimensional data. The approach of this group is to blend well established concepts and methods from computer science, optimization, statistics and signal processing with modern results in (high dimensional) probability, into new computational tools for artificial intelligence and data science. The idea is to channel theory and algorithms into new applications including smarter technologies and sophisticated engines for inference from high-dimensional data. The ultimate objective of the group is a future generation of intelligent algorithms.
Specific projects and representative publications:

- Statistical and computational complexity of ML [20, 22]. The goal is to characterize theoretically the minimal amount of data and computational resources needed for optimal learning under different prior assumptions.

- Efficient algorithms for large scale learning [23, 24]. The goal is to develop large scale ML solutions with optimal prediction accuracy and minimal computational (memory/time/comunication) footprints.

- Learning with structured data [4, 9]. The goal is to learn from data with a variety of structures, including but not limited to multiple related outputs, general geometry (e.g. metric/Banach spaces), multiscale and compositional structure (like audio/video signals) or time structure (time series/dynamical systems).

Computational harmonic analysis and ML. The scientific interests of this group are at the interface between ML, harmonic analysis, inverse problems and partial differential equations. While each individual field has in recent years witnessed significant advances, it appears highly likely that the next major breakthroughs will occur at the intersection of these disciplines. The ultimate goal of this research group is to develop fundamental mathematical tools for the analysis of massive and complex data.
Specific projects and representative publications:

- Foundation of Learning theory [7, 11]. The goal is to investigate the foundation of supervised and unsupervised learning methods developing general frameworks that highlight the connection with neighboring fields, in particular signal processing and harmonic analysis.
- Harmonic analysis [1, 5, 10]. The goal is to investigate a variety of transforms extending classic Fourier analysis, as generalized shearlets and Radon transforms both on manifolds and graphs, tuned for feature extraction and multi-scale data representations.
- Reproducing kernel Hilbert spaces [8, 12]. This is one of the main areas of expertise in MaLGa and it focuses on the vector-valued framework and on its connection with Harmonic Analysis on manifolds and ML data representation.
- Inverse problems and ML [28, 2, 3]. This is one of the trademarks of the MaLGa approach to learning. On the one hand, inverse problems provide a sound framework to understand and solve a variety of learning problems. On the other hand, data-driven approaches based on ML are becoming increasingly popular to solve a variety of inverse problems in imaging and natural sciences.

ML & vision. This group investigates different nuances of visual perception in artificial intelligence systems, where computer vision and ML are combined to obtain robust data-driven methods addressing a variety of problems. In a multidisciplinary approach, the group is interested in mathematically sound algorithms, computational models, and a wide range of applications. In particular, a main goal of the group is to study and develop methods for image processing, scene understanding, motion analysis, and action recognition, with applications

to assisted living, human-machine interaction, robotics, and video-surveillance.
Specific projects and representative publications:

- Multi-resolution image processing [13, 16]. The goal is to explore the potential of multi-resolution representations for 2D and 3D signals (shearlets in particular) to devise image processing and feature detection algorithms invariant and robust to noise.

- Scene and motion understanding [25, 27]. The goal is to design algorithms for understanding semantic and dynamic information in images and videos. We focus primarily on human-centric systems, and study low level as well as high level motion patterns, actions and interactions with objects and with the environment. A relevant application of this research is in assessing physical and emotional well being of elderly and patients with disabilities in assisted living scenarios.

- Vision and robotics [14, 17]. The goal is to incorporate ML and computer vision models within human-robot interaction scenarios. Embodied agents are used as a test bed for general purpose methods as well as a valuable and unique source of visual data.

ML for data science. This group focuses on ML applications in a number of domains heavily relying on data analysis. Here, the approach is to work closely with domain experts to develop predictive models incorporating all the available prior information, but also to devise new algorithms tailored to the need of the applications at hand. A main application domain refers to health and life sciences, including projects in computational biology and the analysis of clinical data. More recently, much effort has been devoted to the analysis of time varying data from biology but also from financial prediction problems.
Specific projects and representative publications:

- Interpretable and explainable ML [26, 21]. The goal is to derive interpretable and transparent ML models to be used as basis of further scientific enquiries. Provable stability and statistical significance of the obtained results is mandatory in this project. The emphasis is on structured data with complex interactions often modelled as networks.

- ML for biological and clinical data [18, 19]. The goal is to provide ML based models to medical doctors and biologists to elaborate and verify different scientific hypotheses and suggest plausible biomarkers.

- Time varying data analysis [6, 15]. The goal is to model and predict a variety of time varying data. This project is largely driven by problems in neuroscience, in collaboration with Gaslini hospital, and by problems in finance in collaboration with several local companies.

5 MaLGa faculties

Giovanni S. Alberti is a tenure-track assistant professor at the Department of Mathematics at the University of Genoa. He received his PhD at the University of Oxford. He held two post-doctoral positions at the École Normale Supérieure in Paris and at ETH Zürich. His research focuses on partial differential equations, applied harmonic analysis, and on their interactions with inverse problems, machine learning and imaging. He was the recipient of the "Gioacchino Iapichino" prize for Mathematical Analysis in 2017 and of the "Eurasian Association on Inverse Problems Young Scientist Award" for distinguished contributions to inverse problems in 2018. He has authored around 30 peer-reviewed papers in international journals and conference proceedings.
Annalisa Barla is associate professor at the DIBRIS Department at the University of Genoa. She received her PhD in Computer Science in at the University of Genoa in 2005 and she then became a post-doctoral researcher in Fondazione Bruno Kessler (Trento, Italy), where she started her research in computational biology. With the FP6 Integrated Project Health-e-Child IST-2004-02774919 she joined (2007) the Statistical Learning and Image Processing group in Genoa, starting a research line in computational biology. She authored more than 70 peer-reviewed papers in international journals and conference proceedings.
Filippo De Mari is full professor at the Department of Mathematics at the University of Genoa. He received his PhD in Mathematics at the Washington University in St. Louis, Missouri in 1987. After a post-doctoral position at the Institut für Dynamische Systeme, at Bremen University (Germany) in 1988–89 and Politecnico di Torino, he became assistant professor at the University of Genoa. His research activity is in the field of (applied) harmonic and signal analysis. He co-authored more than 40 papers on refereed international journals.
Ernesto De Vito is full professor at the Department of Mathematics at the University of Genoa. He received his PhD from the University of Genoa in 1995. He has been assistant professor at the University of Modena, and then associate professor at the University of Genoa. His research activity focuses on machine learning theory and applied harmonic analysis. He co-authored 70 papers in international journals, conference proceedings, and two books. Since 2020 he is the coordinator of *Gruppo UMI: Matematica per l'Intelligenza Artificiale e il Machine Learning* (https://aiml.unich.it/).
Simone Di Marino is associate professor at the Department of Mathematics at the University of Genoa. He received his PhD in Mathematics at the Scuola Normale Superiore di Pisa in 2011 for which he received the "INdAM-UMI-SIMAI" prize. After post-doctoral positions in Université Paris-Sud and Université Paris-Dauphine, he became INdAM permanent researcher, and then associat professor in Geona since 2019. His research activity is in the field of applied optimal transport (evolutionary PDEs, Density Functional Theory, algorithms) as well as analysis in metric spaces. He co-authored more than 20 papers on refereed international journals.

Nicoletta Noceti is a tenure-track assistant professor at the DIBRIS Department at the University of Genoa. She received her PhD in Computer Science (2010) from the University of Genoa. In 2008, she visited the IDIAP Institute (Switzerland) as a research fellow. Over the years, she has explored the theoretical foundations of Computer Vision and Image Processing and applications as robotics, Ambient Assistive Living, and video-surveillance. In such contexts, she has collaborations with universities, research institutes and hospitals. She participated to various national and international research projects (e.g. EU projects SAFE-POST and Healh-e-Child), and technology transfer and development projects with SMEs and large companies. She authored more than 60 peer-reviewed papers in international journals and conference proceedings.
Francesca Odone is full professor at the DIBRIS Department at the University of Genoa. She received the PhD in Computer Science in 2002 from the University of Genoa. In 2002-2005 she was researcher at INFM (CNR), in 2006-2014 she was assistant professor at University of Genoa. She visited Heriot-Watt University, Edinburgh, UK in 1997 and 1999-2000. Her research interests are in the fields of Computer Vision and Machine Learning, including data-driven methods for automatic visual perception, motion analysis, behavior analysis, and pose estimation in video sequences. She explores different application domains, including robotics, assisted living, and video-surveillance. She authored more than 110 peer-reviewed papers in international journals and conference proceedings. She is acting as an Associate Editor of the IEEE Transactions on Emerging Topics in Computing.
Lorenzo Rosasco is full professor at the DIBRIS Department at the University of Genoa, a visiting professor at the Massachusetts Institute of Technology, and external collaborator at the Istituto Italiano di Tecnologia. He received his PhD from the University of Genoa in 2006 and has been visiting student at the Toyota Technological Institute at Chicago and at the Center for Biological and Computational Learning (CBCL) at MIT. He held a research scientist position at MIT between 2006 and 2009. He is the principal investigator of the Laboratory for Computational and Statistical Learning. In 2019 he obtained the ERC consolidator grant "Efficient algorithms for sustainable machine learning". His research focuses on studying theory and algorithms for machine learning. He is known for his foundational work in machine learning as well as the development of sound large scale machine learning algorithms. He authored more than 100 peer-reviewed papers in international journals.
Matteo Santacesaria is a tenure-track assistant professor at the Department of Mathematics of University of Genoa. He received his PhD in Applied Mathematics at École Polytechnique (France) in 2012. He has held post-doctoral positions at Université Joseph Fourier, University of Helsinki and Politecnico di Milano. His main research topics are inverse problems for partial differential equations with applications to medical imaging. His research interests include: Calderòn's inverse conductivity problem, linear and nonlinear reconstruction algorithms, compressed sensing, complex and hypercomplex analysis, regularization and optimization, machine learning. He authored more than 20 peer-reviewed pa-

pers in international journals and conference proceedings.
Alessandro Verri is full professor at the DIBRIS Department at the University of Genoa. He received his PhD from the University of Genoa in 1988. He has been a visiting professor at MIT for several years, then at the Heriot-Watt University (Edinburgh), at the INRIA - IRISA in Rennes and at Berkeley. He supervised more than 30 PhD students, and around 100 master students. His research interests are analytic methods for learning theory, learning algorithms for biomedical data analysis and image understanding, computational methods for the processing of visual information in 2D and 3D. He authored more than 130 peer-reviewed papers in international journals.
Silvia Villa is associate professor at the Department of Mathematics at the University of Genoa. She received her PhD in Mathematics from the University of Florence. She was previously a post-doctoral researcher at the Laboratory for Computational and Statistical learning and an assistant professor at the Department of Mathematics at Politecnico di Milano. She is in the international scientific committee of GdR Mathématiques de l'Optimisation et Applications and in the editorial board of Applied Mathematics and Computation and Journal of Nonsmooth Analysis and Optimization. Her research interests are in convex optimization for inverse problems and machine learning. She authored 38 peer-reviewed papers in international conferences and journals.

References

[1] G. S. Alberti, F. Bartolucci, F. De Mari, and E. De Vito. Unitarization and inversion formulae for the Radon transform between dual pairs. *SIAM J. Math. Anal.*, 51(6):4356–4381, 2019.

[2] G. S. Alberti and M. Santacesaria. Infinite dimensional compressed sensing from anisotropic measurements and applications to inverse problems in PDE. *Appl. Comput. Harmon. Anal.*, 50:105–146, 2021.

[3] G. S. Alberti, E. D. Vito, M. Lassas, L. Ratti, and M. Santacesaria. Learning the optimal regularizer for inverse problems. *https://arxiv.org/abs/2106.06513*, 2021.

[4] M. A. Alvarez, L. Rosasco, N. D. Lawrence, et al. Kernels for vector-valued functions: A review. *Foundations and Trends® in Machine Learning*, 4(3):195–266, 2012.

[5] F. Bartolucci, F. De Mari, E. De Vito, and F. Odone. The Radon transform intertwines wavelets and shearlets. *Appl. Comput. Harmon. Anal.*, 47(3):822–847, 2019.

[6] G. Brichetto, M. M. Bragadin, S. Fiorini, M. A. Battaglia, G. Konrad, M. Ponzio, L. Pedullà, A. Verri, A. Barla, and A. Tacchino. The hidden information in patient-reported outcomes and clinician-assessed outcomes: multiple sclerosis as a proof of concept of a machine learning approach. *Neurological Sciences*, pages 1–4, 2019.

[7] A. Caponnetto and E. De Vito. Optimal rates for the regularized least-squares algorithm. *Foundations of Computational Mathematics*, 7(3):331–368, 2007.

[8] C. Carmeli, E. De Vito, and A. Toigo. Vector valued reproducing kernel Hilbert spaces of integrable functions and mercer theorem. *Analysis and Applications*, 4(04):377–408, 2006.

[9] C. Ciliberto, L. Rosasco, and A. Rudi. A consistent regularization approach for structured prediction. In *Advances in Neural Information Processing Systems*, pages 4412–4420, 2016.

[10] F. De Mari and E. De Vito. Admissible vectors for mock metaplectic representations. *Applied and Computational Harmonic Analysis*, 34(2):163–200, 2013.

[11] C. De Mol, E. De Vito, and L. Rosasco. Elastic-net regularization in learning theory. *Journal of Complexity*, 25(2):201–230, 2009.

[12] E. De Vito, N. Mücke, and L. Rosasco. Reproducing kernel Hilbert spaces on manifolds: Sobolev and diffusion spaces. *Anal. Appl. (Singap.)*, 19(3):363–396, 2021.

[13] M. A. Duval-Poo, F. Odone, and E. D. Vito. Edges and corners with shearlets. *IEEE Trans. Image Processing*, 24(11):3768–3780, 2015.

[14] S. R. Fanello, I. Gori, G. Metta, and F. Odone. Keep it simple and sparse: real-time action recognition. *J. Mach. Learn. Res.*, 14(1):2617–2640, 2013.

[15] S. Fiorini, F. Hajati, A. Barla, and F. Girosi. Predicting diabetes second-line therapy initiation in the Australian population via time span-guided neural attention network. *Plos One*, 14(10), 2019.

[16] D. Malafronte, E. D. Vito, and F. Odone. Space-time signal analysis and the 3D shearlet transform. *Journal of Mathematical Imaging and Vision*, 60(7):1008–1024, 2018.

[17] G. Pasquale, C. Ciliberto, F. Odone, L. Rosasco, and L. Natale. Are we done with object recognition? The iCub perspective. *Robotics and Autonomous Systems*, 112:260–281, 2019.

[18] S. Pesce, M. Greppi, M. Squillario, F. Loiacono, L. Moretta, A. Moretta, S. Sivori, P. Castagnola, A. Barla, S. Candiani, et al. New miRNA signature heralds human NK cell subsets at different maturation steps: involvement of miR-146a-5p in the regulation of KIR expression. *Frontiers in immunology*, 9:2360, 2018.

[19] S. Ravera, M. Podestà, F. Sabatini, M. Dagnino, D. Cilloni, S. Fiorini, A. Barla, and F. Frassoni. Discrete changes in glucose metabolism define aging. *Scientific reports*, 9(1):10347, 2019.

[20] L. Rosasco and S. Villa. Learning with incremental iterative regularization. In *Advances in Neural Information Processing Systems*, pages 1630–1638, 2015.

[21] L. Rosasco, S. Villa, S. Mosci, M. Santoro, and A. Verri. Nonparametric sparsity and regularization. *Journal of Machine Learning Research*, 14:1665–1714, 2013.

[22] A. Rudi, R. Camoriano, and L. Rosasco. Less is more: Nyström computational regularization. In *Advances in Neural Information Processing Systems*, pages 1648–1656, 2015.

[23] A. Rudi, L. Carratino, and L. Rosasco. Falkon: An optimal large scale kernel method. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems 30*, pages 3891–3901. Curran Associates, Inc., 2017.

[24] A. Rudi and L. Rosasco. Generalization properties of learning with random features. *Advances in Neural Information Processing Systems*, 2017.

[25] A. Stagliano, N. Noceti, A. Verri, and F. Odone. Online space-variant background modeling with sparse coding. *IEEE Transactions on Image Processing*, 24(8):2415–2428, 2015.

[26] F. Tomasi, V. Tozzo, A. Verri, and S. Salzo. Forward-backward splitting for time-varying graphical models. In *PGM*, 2018.

[27] A. Vignolo, N. Noceti, F. Rea, A. Sciutti, F. Odone, and G. Sandini. Detecting biological motion for human–robot interaction: A link between perception and action. *Frontiers in Robotics and AI*, 4:14, 2017.

[28] E. D. Vito, L. Rosasco, A. Caponnetto, U. D. Giovannini, and F. Odone. Learning from examples as an inverse problem. In *Journal of Machine Learning Research*, pages 883–904, 2005.