PhD in Information Technology and Electrical Engineering

Università degli Studi di Napoli Federico II

PhD Student: Stefano Marrone

XXXII Cycle

Training and Research Activities Report – Third Year

Tutor: Prof. Carlo Sansone

1. Information

I graduated cum Laude in Computer Engineering in April 2016 at University of Naples Federico II.
In February 2017 I started my first year of the PhD in Information Technology and Electrical Engineering (ITEE) XXXII Cycle at the University of Naples Federico II, under the supervision of Prof. Carlo Sansone.
Since then I hold a fellowship from the Consorzio Interuniversitario Nazionale per l'Informatica (CINI).

2. Study and Training Activities

	Credits year 1								Credits year 2								Credits year 3								Total	Check
	Estimated	1 bimonth	2 bimonth	3 bimonth	4 bimonth	5 bimonth	6 bimonth	Summary	Estimated	1 bimonth	2 bimonth	3 bimonth	4 bimonth	5 bimonth	6 bimonth	Summary	Estimated	1 bimonth	2 bimonth	3 bimonth	4 bimonth	5 bimonth	6 bimonth	Summary		
Modules	26	0	3	0	3	3	11	**39,7**	20	0	2,4	4,7	4	7	0	**18,1**	5	0	0	4.8	0	0	0	**4.8**	**62.6**	30-70
Seminars	10	3,7	3,1	0,4	0,7	0,4	1,5	**9,8**	10	2,4	1,1	0,2	0,4	2	1	**7,1**	5	0	1.2	0	0.5	0.6	0.8	**3.1**	**20**	10-30
Research	20	2,3	2,2	1,8	1,5	1,5	1,2	**10,5**	30	3,5	4	4,5	5,3	2,5	15	**34,8**	50	2	6	8	10.6	13	12.5	**52.1**	**97.4**	80-140
	60	18,3	8,9	8,3	2,2	1,9	20	**60,0**	60	5,9	7,5	9,4	9,7	11,5	16	**60,0**	60	2	7.2	13	11.1	14	13.3	**60,0**	**180,0**	180

Modules attended abroad (Imperial College London)	Lecturer	Start	End	H	CFU
English Class – "Lecture Listening"	Prof. A. Hamid	30/05/19	18/06/19	8	0.8
English Class – "Stand up and speak"	Prof. L. Chiu	30/05/20	20/06/20	8	0.8
English Class – "Mouth Mechanics"	Prof. L. Chiu	29/05/21	19/06/21	8	0.8
English Class – "Presentations"	Prof. G. Simon	30/05/22	20/06/22	8	0.8
English Class – "Grammar for Speaking"	Prof. G. Simon	30/05/23	21/06/23	8	0.8
English Class – "Rhythm and Linking"	Prof. A. Hamid	28/05/24	18/06/24	8	0.8

Seminars attended abroad (Imperial College London)	Lecturer	Host	Date	H	CFU
Advances in Statistical Learning	Several	Prof. D. Mandic	05/09/19	6	1.2
Reconciling Deep Learning with Symbolic AI	Prof. M. Shanahan	Prof. Yike Guo	26/09/19	2.5	0.5

Seminars	Lecturer	Host	Date	H	CFU
Accelerated Computing with CUDA C/C++: Architecture, Programming, and Tools	Prof. L. Troiano	Prof. A. Cilardo	25/11/19	3	0.6
Intelligenza Artificiale ed Etica: La ricerca in IA alla prova delle sfide etiche	Several	Prof. G. Tamburrini	12/06/19	4	0.8

3. Research Activity

In recent years the term "**Artificial Intelligence**" (or AI) has become more and more an integral part of the daily life of all of us. We are increasingly dealing with *smart* mobile phones, *intelligent* voice assistants, *robotic* chats, etc. Our interaction with these "intelligent systems" has become so predominant and widespread that even the world of industry has begun to use such AI in factory life and logistics. The term artificial intelligence refers to the ability of a computer (an *artificial* entity) to perform functions resembling the typical reasoning of the human mind (i.e. *intelligence*). Indeed, Marvin Minsky, Alan Turing, Frank Rosenblatt and other AI pioneering studies focused on the development of artificial entities able to autonomously do things usually requiring human intelligence (e.g. the ability to make a decision based on the status of the environment) to be performed.

Although commonly thought to be a child of the last years, the first studies on artificial agents began on the eve of the second world war. From that moment on, amid ups and downs, researches started focusing on the development of theories and mathematical models laying the foundation for the upsurge of artificial intelligence in a wide variety of domains. Nowadays, the term AI is widely abused, and one of the unpleasant effects of this spread with the mass audience is the confusion made with all its related terms, such as "Pattern Recognition", "Machine Learning", "Deep Learning", etc., too often used interchangeably. Indeed, what usually media refers to with AI is actually Machine Learning (ML), a term used to describe the ability of this kind of AI systems to ***learn from examples***, just as we humans learn from experience. This peculiarity has made it possible to relieve the programmer from the task of writing the sequence of operations necessary to perform a given task (algorithm), allowing them to perform increasingly complex tasks, for which it would have been impossible to code a solution.

Among all machine learning models, Artificial Neural Networks (ANN, often referred simply as Neural Network - NN) are definitely the branch that has been receiving the most media coverage since their parallel layered structure of computing elements (i.e. artificial neurons) is inspired to the human brain complex interconnected structure of biological neurons. More recently, the introduction of General Purpose GPU (GP-GPU) computing, the development of free and easy to use frameworks, the availability of huge labelled dataset and progresses in gradient-based optimization, determined the uprising of ***Deep Neural Networks***. The term "Deep Learning" (DL) refers to a particular subset of ANNs characterized, inter alia, by a very "depth structure" (i.e. made up of several layers). Another key aspect of deep models is their ability to autonomously learn the best or set of features for the task under analysis, to the point of even exceeding human capabilities in some tasks. This characteristic, known as *feature learning*, has played a key role in the recent spread of AI since allowed DL use also in domains lacking effective expert-designed features.

In the last years, the impact of AI, and in particular of deep learning, on the industry has been so disrupting that it gave rise to a new wave of research and applications that goes under the name of ***Industry 4.0***. This term refers to the application of AI and cognitive computing to leverage an effective data exchange and processing in manufacturing technologies, services and transports, laying the foundation of what is commonly known as *the fourth industrial revolution*. As a consequence, today's developing trend is increasingly focusing on AI based data-driven approaches, mainly because leveraging user's data (such as location, action patterns, social information, etc.) can make applications able to adapt to them, enhancing the user experience. To this aim, tools like automatic image tagging (e.g. those based on face recognition), voice control, personalized advertising, etc. process enormous amounts of data (often remotely due to the huge computational effort required) too often rich in sensitive information.

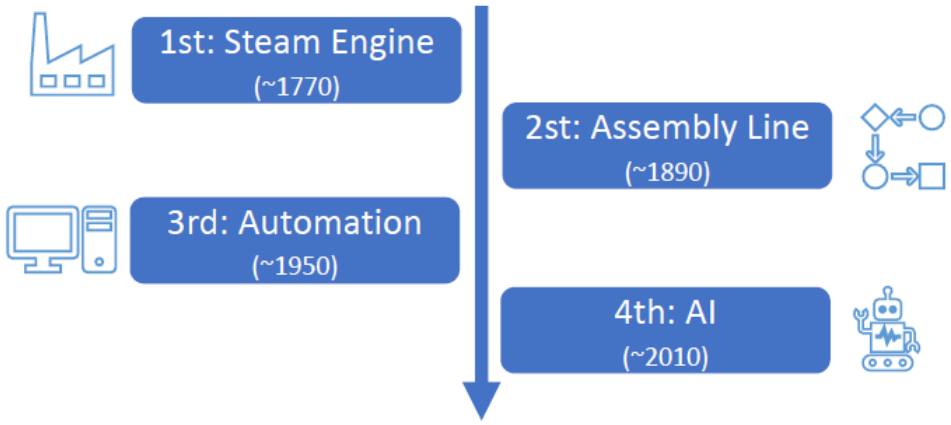

Figure 1: Timeline for the four industrial revolutions: the first, based on the invention of the steam engine; the second, thanks to the development of the assembly line by Henry Ford; the third, with the development of computers and automation; the fourth, supported by artificial intelligence.

Artificial intelligence has thus been proving to be so effective that today it is increasingly been using also in critical domains such as facial recognition, biometric verification (e.g. fingerprints), autonomous driving etc. Although this opens unprecedented scenarios, it is important to note that its misuse (malicious or not) can lead to unintended consequences, such as unethical or unfair use (e.g. discriminating on the basis of ethnicity or gender), or used to harm people's privacy. Indeed, if on one hand, the industry is pushing toward a massive use of artificial intelligence enhanced solution, on the other it is not adequately supporting researches in end-to-end understating of capabilities and vulnerabilities of such systems. The results may be very (negatively) mediatic, especially when regarding borderline domains such those related to subjects privacy or to ethical and fairness, like users profiling, fake news generation, reliability of autonomous driving systems, etc.

We strongly believe that, since being just a (very powerful) tool, AI is not to blame. Nonetheless, we claim that in order to develop a more ethical, fair and secure use of artificial intelligence, all the involved actors (in primis users, developers and legislators) must have a very clear idea about some critical questions, such as *"what is AI?"*, *"what are the ethical implications of its improper usage?"*, *"what are its capabilities and limits?"*, *"is it safe to use AI in critical domains?"*, and so on. Moreover, since AI is very likely to be an important part of our everyday life in the very next future, ***it is crucial to build trustworthy AI systems***.

Therefore, ***the aim of my research activity was to make a first step towards the crucial need for raising awareness about reproducibility, security and fairness threats associated with AI systems, from a technical perspective as well as from the governance and from the ethical point of view.*** Among the several issues that should be faced, in this work we try to face three central points: understanding what "intelligence" means and implies within the context of AI; analyze the limitations and the weaknesses that might affect an AI-based system, independently from the particular adopted technology or technical solutions; assessing the system behaviors in the case of successful attacks and/or in presence of degraded environmental conditions.

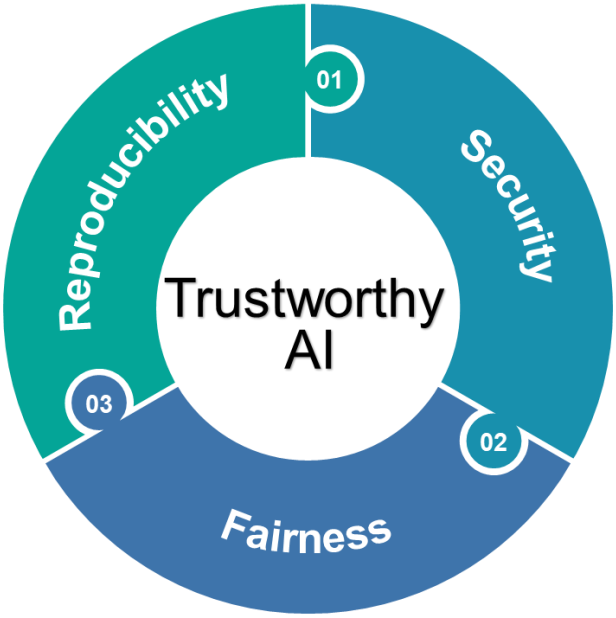

Figure 2: The three aspects that, in our opinion, are crucial to build a trustworthy AI

To this aim, my thesis is divided into three main parts: in part one we introduce the concept of AI, focusing on Deep Learning and on some of its more crucial issues, before moving to ethical implications associated with the notion of "intelligence"; in part two we focus on the perils associated with the reproducibility of results in deep learning, also showing how a proper network design can be used to limit their effects; finally, in part three we address the implications that an AI misuse can cause in a critical domain such as biometrics, proposing some attacks duly designed for the scope.

The cornerstone of the whole thesis are ***adversarial perturbations***, a term referring to the set of techniques intended to deceive AI systems by injecting a small perturbation (noise, often totally imperceptible to a human being) into the data. The key idea is that, although adversarial perturbations are a considerable concern to domain experts, on the other hand, they fuel new possibilities to both favors a fair use of artificial intelligence systems and to better understand the "reasoning" they follow in order to reach the solution of a given problem.

It should be clear that several are the issues that must be faced, such as: designing systems that analyze people data ensuring privacy by default; analyzing the limitations and the weaknesses that might affect an AI-based system, independently from the particular adopted technology or technical solutions; assessing the behaviors in the case of successful attacks and/or in presence of degraded environmental conditions; etc. Indeed, if on one hand, the industry is pushing toward a massive use of artificial intelligence enhanced solution, on the other it is not adequately supporting researches in end-to-end understating of capabilities and vulnerabilities of such systems. The results may be very (negatively) mediatic, especially when regarding borderline domains related to subjects privacy, ethics and fairness, such as users profiling, fake news generation and reliability of autonomous driving systems.

As shown through the thesis, since AI is extremely pervasive in our life, there is a high risk that the choices made by using such models may have a significant impact on society. Therefore, it is becoming more and more crucial to quickly understand how to properly regulate artificial intelligence. But, ***is the legislator able to cope it?*** The solution is not straightforward, not only due to the difficulties arising trying to put in practice AI policies, but also because it is a problem that must be addressed internationally, and not on a local scale.

Unfortunately, this is a very hard matter, since opinions about it are extremely discordant even within the same country. For example, in the USA, on one hand, FBI claims that their AI algorithms are effective and reliable to the point of being usable as scientific evidence, on the other Google is pushing toward the development of a suitable AI regulation. In Europe, the situation appears a little more uniform, mainly thanks to General Data Protection Regulation (**GDPR**), a document through which the European Parliament has proposed, in 2016, a set of rules to regulate the activity of any company operating with data belonging to citizens from any European country.

Article 22. Automated individual decision making, including profiling

1. The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her.
2. Paragraph 1 shall not apply if the decision:
   (a) is necessary for entering into, or performance of, a contract between the data subject and a data controller;
   (b) is authorised by Union or Member State law to which the controller is subject and which also lays down suitable measures to safeguard the data subject's rights and freedoms and legitimate interests; or
   (c) is based on the data subject's explicit consent.
3. In the cases referred to in points (a) and (c) of paragraph 2, the data controller shall implement suitable measures to safeguard the data subject's rights and freedoms and legitimate interests, at least the right to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision.
4. Decisions referred to in paragraph 2 shall not be based on special categories of personal data referred to in Article 9(1), unless point (a) or (g) of Article 9(2) apply and suitable measures to safeguard the data subject's rights and freedoms and legitimate interests are in place.

Figure 3: One of the most important GDPR articles concerning fair artificial intelligence

In conclusion, AI represents without any doubt one of the greatest achievements made by humans. It has the power of really changing our world and to help people, even more than fire and electricity did. However, since ***"with great power comes great responsibility"***, we must learn how to properly use it, developing methods and enacting laws that support its fair, secure and ethical usage for all people around the world.

4. Activity abroad

During the Ph.D., my research focused on theoretical and practical aspects associated with a fair and ethical use of Artificial Intelligence. To deepen my knowledge of the subject, I spent almost 8 months at Imperial College London, hosted by the Computational Privacy Group lead by Dr. Yves-Alexandre De Montjoye.

There, I had the opportunity to learn several new topics, including attack and defense techniques to preserver users' privacy. I used what I have learnt to improve my research activity, resulting in three work submitted to important international journals. Moreover, I took part in very important activity concerning subject privacy (note: these activities are under NDA, thus no information can be given about them). The results will be shortly published in international journals.

5. Products

- Hesham Elhalawani, Timothy A Lin, Stefania Volpe, Abdallah S.R. Mohamed, Aubrey L. White, James Zafereo, Andrew Wong, Joel E. Berends, Shady Abohashem, Bowman Williams, Jeremy M. Aymard, Aasheesh Kanwar, Subha Perni, Crosby D. Rock, Luke Cooksey, Shauna Campbell, Pei Yang, Khanh Nguyen, Rachel Ger, Carlos Eduardo Cardenas, Xenia Fave, Carlo Sansone, Gabriele Piantadosi, **Stefano Marrone**, Rongjie Liu, Chao Huang, Kaixian Yu, Tenfei Li, Yang Yu, Youyi Zhang, Hongtu Zhu, Jeffrey S. Morris, Veerabhadran Baladandayuthapani, John W. Shumway, Alakonanda Ghosh, Andrei Pöhlmann, Hady Ahmady Phoulady, Vibhas Goyal, Guadalupe Canahuate, G. Elisabeta Marai, David Vock, Stephen Y. Lai, Dennis S. Mackin, Laurence E. Court, John Freymann, Keyvan Farahani, Jayashree Kalpathy-Cramer and Clifton D FullerIn "Machine learning applications in head and neck radiation oncology: lessons from open-source Radiomics challenges." Frontiers in Oncology 8 (2018): 294.

- Piantadosi G., **Marrone S.**, Fusco R., Sansone M., Sansone C. "Comprehensive computer-aided diagnosis for breast T1-weighted DCE-MRI through quantitative dynamical features and spatio-temporal local binary patterns." IET Computer Vision 12.7 (2018): 1007-1017

- Amato F., **Marrone S.**, Moscato V., Piantadosi G., Picariello A., Sansone C. "HOLMeS: eHealth in the big data and deep learning era." Information 10.2 (2019): 34

- Galli A., **Marrone S.**, Piantadosi G., Sansone M., Sansone C. “3TP U-Net: Leveraging Tracer Kinetic in DCE-MRI Breast Lesion Segmentation.” In: Artificial Intelligence in Medicine (2019) - (under review)

- Buizzaa C., Quilodran Casas C., Nadler P., Mack J., Titus Z., Le Cornec C., Heylen E., Dur T., Baca Ruiz L., Heaney C., Lopez J.A.D., **Marrone S.**, Arcucci R. “Data Learning: Integrating Data Assimilation with Machine Learning.” In: Journal of Computational Science (2020) - (under review)

- **Marrone S.**, Sansone C. “A Transferable Adversarial Perturbation Attack Against Fingerprint
    - Based Authentication Systems.” In: Computer Vision and Image Understanding (2020) - (under review)

- **Marrone S.**, Papa C., Sansone C., “Effects of Hidden Layer Sizing on CNN Fine-Tuning.” In: Future Generation Computer Systems (2020) - (under review)

- **Marrone S.,** Piantadosi G., Sansone M., Sansone C. "Look-up tables for efficient non-linear parameters estimation." In: International Conference on Optimization and Decision Science, Sorrento (Italy), 4-7 Sept., Springer, Cham (2017) p. 49-57

- **Marrone S.,** Piantadosi G., Fusco R., Petrillo A., Sansone M., Sansone C. "An investigation of deep learning for lesions malignancy classification in breast DCE-MRI." In: International Conference on Image Analysis and Processing, Catania (Italy), 11-15 Sept., Springer, Cham (2017) p. 479-489

- Amato, F., **Marrone, S.,** Moscato, V., Piantadosi, G., Picariello, A., & Sansone, C. "Chatbots Meet eHealth: Automatizing Healthcare." In: *WAIAH @ AI* IA, Bari (Italy), 4 Nov.* (2017) p. 40-49

- **Marrone S.,** Piantadosi G., Sansone M., Sansone C. "On Reproducibility of Deep Convolutional Neural Networks Approaches." In: International Workshop on Reproducible Research in Pattern Recognition, Beijing (China), 20 Aug., Springer, Cham (2018) p. 104-109

- **Marrone S.,** Sansone C., “Approximate Computing for Sizing Hidden Layer in CNN”. In: 4th Workshop on Approximate Computing (AxC) @ DATA, Florence (Italy), 29 Mar., (2019)

- **Marrone S.,** Sansone C., "Adversarial Perturbations Against Fingerprint Based Authentication Systems." In: International Conference on Biometrics (ICB), Crete (Greece), 4-7 Jun., IEEE (2019) p. 1-6

- **Marrone S.,** Sansone C., "An Adversarial Perturbation Approach Against CNN-based Soft Biometrics Detection." In: International Joint Conference on Neural Networks (IJCNN), Budapest (Hungary), 14-19 Jul., IEEE (2019) p. 1-8

- Piantadosi G., **Marrons S.**, Galli A., Sansone M., Sansone C. "DCE-MRI Breast Lesions Segmentation with a 3TP U-Net Deep Convolutional Neural Network." In: IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS), Córdoba (Spain), 5-7 Jun., IEEE (2019) p. 628-633

- **Marrone S.**, Olivieri S., Piantadosi G., Sansone C. "Reproducibility of Deep CNN for Biomedical Image Processing Across Frameworks and Architectures." In: 27th European Signal Processing Conference (EUSIPCO), Coruña (Spain), 2-6 Sept. IEEE (2019) p. 1-5

- Galli A., Gravina M., **Marrone S.**, Piantadosi G., Sansone M., Sansone C. "Evaluating Impacts of Motion Correction on Deep Learning Approaches for Breast DCE-MRI Segmentation and Classification." In: International Conference on Computer Analysis of Images and Patterns, Salerno (Italy), 2-6 Sept., Springer, Cham (2019) p. 294-304

- Gravina M., **Marrone S.**, Piantadosi G., Sansone M., Sansone C. "3TP-CNN: Radiomics and Deep Learning for Lesions Classification in DCE-MRI." In: International Conference on Image Analysis and Processing, Trento (Italy), 9-13 Sept., Springer, Cham (2019) p. 661-671

- Gravina M., Marrone S., Sansone M., Sansone C. (2020) Disentangled 3TP-CNN: Exploiting and Disentangling Contrast Agent Effects Toward Protocol-Independent Lesions Classification in DCR-MRI. In: Pattern Recognition Letters (in preparation)

- Aprea F., **Marrone S.**, Sansone C. (2020) Data Assimilation for Neural Machine Registration in Biomedical Imaging. In: Journal of Computational Science (in preparation)

6. Conferences and Seminars

- Oral presentation to describe my research activity @ Cafè Scientific – Imperial College London, South Kensington Campus, London (UK) – 30/07/2019

- Oral presentation of “BLADeS: Breast Lesions Automatic Detection and Diagnosis System” @ MATAB Expo, Milano/Roma, 25 and 26/06/2019

- Oral presentation on “Exploring Artificial Intelligence Capabilities and Limitations“ @ AIDP Congress, Napoli – 02/02/2019

- Oral presentation on “Exploring Artificial Intelligence Capabilities and Limitations“ @ Rotary Club Milano Digital, webminar – 12/03/2019

- Oral presentation of the paper “An Adversarial Perturbation Approach Against CNN-based Soft Biometrics Detection” @ the 2019 International Joint Conference on Neural Networks (IJCNN), Budapest (Hungary) – 17/07/2019

- Oral presentation of the paper “DCE-MRI Breast Lesions Segmentation with a 3TP U-Net Deep Convolutional Neural Network” @ IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS), Cordoba (Spain) – 07/06/2019
- Oral presentation of the paper “Evaluating Impacts of Motion Correction on Deep Learning Approaches for Breast DCE-MRI Segmentation and Classification” @ International Conference on Computer Analysis of Images and Patterns (CAIP), Salerno (Italy) – 04/09/2019
- Oral presentation of the paper “3TP-CNN: Radiomics and Deep Learning for Lesions Classification in DCE-MRI” @ International Conference on Image Analysis and Processing (ICIAP), Trento (Italy) – 11/09/2019
- Oral presentation on “Exploring Artificial Intelligence Capabilities and Limitations“ @ Data Science Institute, Imperial College London (UK) – 14/05/2019

7. Tutorship and Active Teaching

- Tutor within the DIETI Tutorship program
- Tutor within the IBM Tutorship Nerd program (from February 2018 to April 2018)
- Lecturer for several MATLAB seminars as MATLAB Student Ambassador (ongoing)
- Lecturer for 6 classroom training at Data Mining course (prof. C. Sansone) a.y. 19/20 - (2h each)
- Lecturer for a classroom training at “Architetture dei Sistemi di Elaborazione” course (prof. N. Mazzocca) a.y. 19/20 - (1h)
- Correlator for 5+ Master's Degree Thesis

8. Membership

- Associazione Italiana per la ricerca in Computer Vision, Pattern recognition e machine Learning (CVPL- ex-GIRPR)
- Institute of Electrical and Electronics Engineers (IEEE)
- IEEE Computational Intelligence Society
- IEEE Young Professionals
- IEEE Napoli Student Branch (lead by prof. P. Maresca). Until the 7/01/2019 I was the Chair of the branch

9. Collaborations (selected)

- Istituto Nazionale Tumori IRCCS "Fondazione G. Pascale" (Napoli, IT), for breast cancer analysis (BLADeS project)
- Department of Medicine, Univeristà Vanvitelli (Napoli, IT), for AI based cancer screening (Synergy-Net POR Project)
- Dipartimento di Agraria, Università Federico II (Napoli, IT), for satellite image processing
- Dipartimento di Ingegneria Strutturale, Università Federico II (Napoli, IT), for damaged building analysis with AI

10. Roles Held

- MATLAB Student Ambassador
- Co-Founder of the ARTE University Spin-off
- Guest Editor for the Special Issue ”Intelligent Innovations in Multimedia Data” of MPDI Future Internet Journal
- Chair of the 1st e-BADLE (eHealth in the Big Data and Deep Learning Era) workshop, held in conjunctionwith ICIAP 2019
- Program Committee for CMBS2019 - HealthCare 4.0
- Program Committee for BIBM2018 - Computational methods for Hospital 4.0
- Reviewer for several International Conferences and Journal, including IET Biometrics, IJCNN, ICIAP