UNIVERSITÀ DEGLI STUDI DI NAPOLI FEDERICO II

DOTTORATO DI RICERCA / PHD PROGRAM IN
INFORMATION TECHNOLOGY AND ELECTRICAL ENGINEERING

Activities and Publications Report

PhD Student: Fabrizio Guillaro

Student DR number: DR995863

PhD Cycle: XXXVII

PhD Cycle Chairman: Prof. Stefano Russo

PhD program student's start date: 01/11/2021
PhD program student's end date: 31/10/2024

Supervisor:

e-mail: verdoliv@unina.it

Co-supervisor:

e-mail: poggi@unina.it

PhD scholarship funding entity:

DARPA, under the SEMAFOR program through the DISCOVER project.

General information

Fabrizio Guillaro received the Master Science degree in Computer Engineering from the University of Napoli Federico II in 2021. He attended a curriculum in **Image Forensics** within the PhD program in Information Technology and Electrical Engineering. He received a grant from DARPA, under the SEMAFOR program through the DISCOVER project.

Study activities

Attended Courses

Year	Course Title	Type	Credits	Lecturer	Organization
1st	Introduction to Deep Learning	Ad hoc course	6	Prof. Giovanni Poggi, Dr. Diego Gragnaniello (University of Salerno)	Scuola Superiore Meridionale
1st	Visione per Sistemi Robotici	MSc course	9	Prof. Giovanni Poggi	UniNa
1st	Image and Video Processing for Autonomous Driving	MSc course	6	Prof. Luisa Verdoliva	UniNa
2nd	How to boost your PhD	Ad hoc course	4	Prof. Antigone Marino	ITEE
2nd	Statistical Multimedia Security and Forensics	Ad hoc course	4	Prof. Fernando Pérez-González (University of Vigo)	University of Trento – IECS Doctoral School
3rd	Strategic Orientation for STEM Research & Writing	Ad hoc course	5	Dr. Chie Shin Fraser	ITEE
3rd	Innovation and Entrepreneurship	Ad hoc course	4	Prof. Pierluigi Rippa	ITEE

Attended PhD Schools

Year	School title	Location	Credits	Dates	Organization
1st	DeepLearn 2022 Summer - 6th Internation Gran Canaria School on Deep Learning	Las Palmas de Gran Canaria, Spain	5	25/07/22 - 29/07/22	University of Las Palmas de Gran Canaria, Spain
2nd	International Computer Vision Summer School (ICVSS)	Scicli (Ragusa), Sicilia	6	9/07/2023 - 15/07/2023	University of Catania
3rd	IEEE-EURASIP Summer School on Signal Processing (S3P-2024)	Capri, Italy	4	23/09/2024 - 29/09/2024	University of Napoli Federico II

Attended Seminars

Year	Seminar Title	Credits	Lecturer	Lecturer affiliation	Organization
1st	Single cell omics leverage Machine Learning to dissect tumor microenvironment and cancer immuno editing	0.4	Dr. Raoul J.P. Bonnal	IFOM - the FIRC Institute of Molecular Oncology	ITEE
1st	Threat Hunting Essentials	0.4	Roman Rezvukhin	Group-IB	ITEE
1st	Physiological forensics	0.6	Chau-Wai Wong, Min Wu	North Carolina State University	IEEE SPS-IFS
1st	Is Machine Learning Security IFS-business as usual?	0.2	Teddy Furon	Inria	IEEE - WIFS
1st	Multimedia data recovery and its related workflows in digital forensics	0.3	Patrick De Smet	Belgian National Institute of Criminalistics and Criminology	IEEE - WIFS
1st	GDPR basics for computer scientists	0.4	Dr. Rigo Wenning	European Research Consortium for Informatics and Mathematics	ITEE
1st	The Deepfake Threat	0.2	Graham Brookie, Emerson Brooking	Atlantic Council's Digital Forensic Research Lab	DARPA
1st	Intelligenza artificiale e sistemi d'arma autonomi	0.3	Fosca Giannotti, Guglielmo Tamburrini	Scuola Normale Superiore, UniNa	CNR-ISTI
1st	Federated Learning in Big Data over Networks	0.2	Alexander Jung	Aalto University	SPS-DSI
1st	Towards Trustworthy AI - Integrating Reasoning and Learning	0.2	Fredrik Heintz	Linköping University	International AI Doctoral Academy (AIDA)
1st	Visual Data Analysis: Why? When? How?	0.2	Pat Hanrahan, Tamara Munzner	Stanford University, University of British Columbia	Association for Computing Machinery (ACM)
1st	Apprendimento Automatico per la Visione Artificiale (1)	0.8	Prof. Vito Roberto	University of Udine	University of Udine
1st	Apprendimento Automatico per la Visione Artificiale (2)	0.8	Prof. Vito Roberto	University of Udine	University of Udine
1st	Apprendimento Automatico per la Visione Artificiale (3)	0.8	Prof. Vito Roberto	University of Udine	University of Udine
1st	Apprendimento Automatico per la Visione Artificiale (4)	0.8	Prof. Vito Roberto	University of Udine	University of Udine

1st	Apprendimento Automatico per la Visione Artificiale (5)	0.8	Prof. Vito Roberto	University of Udine	University of Udine
1st	Apprendimento Automatico per la Visione Artificiale (6)	0.8	Prof. Vito Roberto	University of Udine	University of Udine
1st	Apprendimento Automatico per la Visione Artificiale (7)	0.8	Prof. Vito Roberto	University of Udine	University of Udine
1st	IEEE Authorship and Open Access Symposium: Tips and Best Practices to Get Published from IEEE	0.3	Dr. Derek Abbott, Dr. Paolo Bonato, Eszter Lukács, Judy Brady	University of Adelaide, Harvard University, IEEE, IEEE	IEEE Association
1st	An Introduction to Deep Learning for Natural Language Processing	0.2	Dr. Marco Valentino	University of Manchester	ITEE
1st	Explainable Natural Language Inference	0.3	Dr. Marco Valentino	University of Manchester	ITEE
1st	Deepfake Detection: Humans vs. Machines	0.2	Dr. Pavel Korshunov	Idiap Research Institute	IEEE SPS-IFS
1st	Interactive Art and Pattern Recognition	0.6	Larry O'Gorman	Bell Labs Research	ICPR 2022
1st	Towards Actionable XAI	0.2	Dr. Sebastian Lapuschkin	Fraunhofer Heinrich Hertz Institute	International AI Doctoral Academy (AIDA)
2nd	From Handcrafted to End-to-End Learning, and Back: a Journey for Multi-Object Tracking	0.4	Prof. Dr. Laura Leal-Taixé	NVIDIA, Technical University of Munich	International AI Doctoral Academy (AIDA) and University of Modena and Reggio Emilia
2nd	Digital Forensics	0.4	Artem Artemov	Group-IB	ITEE
2nd	Face Presentation Attack Detection	0.3	Prof. Sébastien Marcel	Idiap Research Institute	IEEE Biometrics Council
2nd	Advances on Multimodal Machine Learning Solutions for Speech Processing Tasks and Emotion Recognition	0.2	Dr. Fei Tao and Dr. Carlos Busso	Beihang University, University of Texas at Dallas	IEEE Signal Processing Society
2nd	The Super Neuron Model - A new generation of ANN-based Machine Learning and Applications	0.2	Moncef Gabbouj	Tampere University	EURASIP
2nd	What's Up with Image and Video Forensics?	0.2	Prof. Fernando Pérez-González	University of Vigo	EURASIP

2nd	Unleashing the Power of LLMs: a Historical perspective on Generative AI	0.2	Prof. Tarry Singh	University of Texas at Dallas	ITEE
2nd	Computational Disinformation Symposium	1.4	Micah Musser, Joe Cheravitch, Sougata Saha...		NYU Tandon School of Engineering
2nd	Prompting in Vision	0.6	Phillip Isola, Ludwig Schmidt, Ziwei Liu, ...		IEEE/CVF CVPR 2023
2nd	Scientific Integrity Verification Through Image Forensics	0.2	Dr. Daniel Moreira	Loyola University Chicago	SPS-IFS

Research activities

Fabrizio Guillaro has conducted research in the field of Image Forensics, focusing on image forgery detection and localization. He proposed and developed a novel approach that can be applied to detect a large variety of image forgeries, from classic cheapfakes to more recent manipulations based on deep learning. The detector is designed to enhance the in-camera and out-camera artifacts even in challenging scenarios, such as circulation on social networks. In addition, it provides a confidence map to reduce false alarms. The research work also includes a study on the vulnerabilities of current forensic detectors to adversarial attacks.

Fabrizio Guillaro spent three months as PhD Student Researcher at Google LLC in Mountain View, California, USA in the Mint (Media Integrity) group under Google Research. He continued the internship remotely for an additional three and a half months from Naples, Italy. During this period, his research focused on AI-generated image detection and worked on out-of-distribution (OOD) generalization methods.

He presented one research contribution at the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) in 2023. Part of his work was done in collaboration with Google Research, with other ongoing projects with Google DeepMind that are currently under preparation.

Tutoring and supplementary teaching activities

Fabrizio Guillaro performed the following tutorship activities (Teaching assistance and Laboratory activities):

- Elaborazione di Segnali Multimediali (49 hours). Tutor: Prof. Luisa Verdoliva
- Teoria dei Segnali (2.5 hours). Tutor: Prof. Giovanni Poggi

Credits summary

PhD Year	Courses	Seminars	Research	Tutoring / Supplementary Teaching
1st	26	10.8	23	1.28
2nd	14	4.1	41.1	0.28
3rd	13	0	47.4	0.5

Research periods in institutions abroad and/or in companies

PhD Year	Institution / Company	Hosting tutor	Period	Activities
2nd	Google LLC (Mountain View, California, USA)	Dr. Avneesh Sud	30/10/2023 – 31/10/2023	First days: orientation and onboarding
3rd	Google LLC (Mountain View, California, USA)	Dr. Avneesh Sud	01/11/2024 – 29/01/2024	Research work: AI-generated image detection and analysis of the out-of-distribution problem.
3rd	Google S.r.l. (Remotely in Naples, Italy)	Dr. Avneesh Sud	30/01/2024 – 10/05/2024	Research work: confidence estimation for AI-generated image detectors.

PhD Thesis

Synthetic media generation has seen tremendous progress in the span of just a few years. Powered by large language models, text-to-image synthesis tools allow the user to create from scratch and modify images at will by means of simple text instructions. While this represents a great opportunity for visual arts applications, it can be also used by malicious actors who aim at deliberately disseminating disinformation. Consequently, *multimedia forensics* has been drawing increasing attention, with a strong demand for effective detectors able to distinguish manipulated images from real ones.

Aim of this thesis is to develop a general and reliable approach for detecting and localizing image manipulations. The approach developed by Fabrizio Guillaro is based on the extraction of a learned camera model fingerprint (Noiseprint++) that is combined with the RGB image to detect and localize image manipulations. Noiseprint++ is designed to improve robustness to post-processing operations, such as resizing and compression, that may attenuate forensic traces in a realistic scenario when images circulate over the web. To this end, during training a proper contrastive learning approach is applied on image patches that underwent a different combination of processing operations (editing history). It is worth noting that the training dataset includes only real images coming from a very large collection of camera models. Hence, forgeries are detected as deviations from the expected regular pattern that characterizes each pristine image. Looking for anomalies makes the approach able to robustly detect a variety of local manipulations, ensuring generalization. The proposed detector, called TruFor, combines the extracted Noiseprint++ with RGB data to provide as output a confidence map that highlights areas where localization predictions may be error prone. This is particularly important in forensic applications to reduce false alarms. Extensive experiments on several publicly available datasets show that Trufor can detect and localize both cheapfakes and deepfakes with superior performance compared to state-of-the-art.

Additionally, in this thesis the robustness of current forensic detectors to adversarial attacks has been explored. The transferability of attacks between different families (convolutional neural networks and vision transformers) has been analyzed, both numerically and with the help of the resulting Fourier-domain patterns. This analysis has shed light on how forensic detectors work and is therefore a valuable tool for developing more effective and robust methods.

Research products

Research results appear in 3 contributions to international conferences and 1 contribution to a chapter book.

List of scientific publications

International conference papers

H. Mareen, D. Vanden Bussche, **F. Guillaro**, D. Cozzolino, G. Van Wallendael, P. Lambert, L. Verdoliva,
"Comprint: Image Forgery Detection and Localization using Compression Fingerprints",
In: Pattern Recognition, Computer Vision, and Image Processing. ICPR 2022 International Workshops and Challenges. Lecture Notes in Computer Science,
vol 13644, pp. 281-299. Springer, Cham. Montréal, QC, Canada, 2022.
DOI:10.1007/978-3-031-37742-6_23

F. Guillaro, D. Cozzolino, A. Sud, N. Dufour, L. Verdoliva,
"TruFor: Leveraging all-round clues for trustworthy image forgery detection and localization",
in IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR),
Vancouver, BC, Canada, 2023, pp. 20606-20615,
DOI: 10.1109/CVPR52729.2023.01974

V. De Rosa, **F. Guillaro**, G. Poggi, D. Cozzolino, L. Verdoliva ,
“Exploring the Adversarial Robustness of CLIP for AI-generated Image Detection”,
IEEE International Workshop on Information Forensics and Security (WIFS),
Rome, Italy, December 2024.

Chapter Books

F. Guillaro, D. Cozzolino, G. Poggi, L. Verdoliva ,
"Uncertainty-driven detection and localization of image forgeries",
Chapter in CNIT Volume. Series: Signal Processing and Learning for Next Generation Multimedia, edited by Riccardo Bernardini, Lucio Marcenaro, Roberto Rinaldo and Pietro Zanuttigh
pp. 145-164, 2024, DOI: 10.57620/CNIT-Report_13

Date 15/10/2024

PhD student signature Fabrizio Guillaro

Supervisor signature Luisa Verdoliva