Matteo Testa

Turin, Italy +39 393 066 1012
www.matteotesta.it matteo testa it.mtesta@gmail.com

I am an applied researcher in deep learning with a strong background in signal processing for speech and images, supported by both a solid academic record and industrial experience. I have worked across multiple sectors, including automotive and banking, and my expertise in designing and prototyping deep learning solutions is highly transferable to other domains. Currently, I develop cutting-edge deep learning technologies for Text-to-Speech (TTS) models.

Research Interests & Expertise

- Generative modeling: neural TTS, autoregressive transformers, diffusion and flow-matching models, latent-space modeling.
- LLMs & multimodal reasoning: LLM-powered speech systems, prosody control, cross-lingual voice conversion.
- Self-supervised speech representations: WavLM SSL models and efficient speaker identity removal.

Experience

Cerence | *Principal Research Scientist (Text to Speech)* Jan. 2023 – present

- Mentor and supervise PhD student and junior researcher, fostering knowledge sharing and best practices in speech synthesis research.
- Develop novel models for next generation TTS by improving expressiveness, stability, and product robustness, balancing research innovation with deployment constraints. Patent [P2]
- Conduct research and prototyping on Large Language Models (LLMs) for speech-related tasks, with focus on improved naturalness, prosody control and stability.
- Investigate and implement diffusion and flow-matching models for speech synthesis, aiming at improving naturalness and speaker adaptation.
- Lead small cross-functional research projects from concept to prototype, coordinating timelines and deliverables with other researchers and engineers.
- Continuously review recent literature to identify promising directions and integrate cutting-edge findings into ongoing projects.

Cerence | *Senior Research Engineer (Text to Speech)* Nov. 2020 – Jan. 2023

- Designed and implemented novel approaches to advance Token Conversion (TC) models for speech synthesis.
- Focused on improving model stability and robustness, ensuring consistent speech output across diverse input conditions.
- Developed techniques to detect and mitigate instabilities in speech synthesis models, enabling more reliable deployment in production settings. Patent [P3]

Unicredit R&D | *Senior AI Engineer* Jul. 2019 – Nov. 2020

- Developed AI solutions to automate banking processes through statistical methods, computer vision, and deep learning.

- Worked on Natural Language Processing (NLP) and Computer Vision (CV) tasks to extract and process information from structured and unstructured documents.
- Designed and implemented models for signature verification and matching in check processing, improving fraud detection capabilities.
- Built pipelines for information extraction from scanned documents (e.g., invoices, forms), enabling automatic digitization and classification of bank documents.

Polytechnic University of Turin | *PostDoctoral researcher* March 2016 – June 2019

- Conducted research on deep learning models for computer vision and biometric authentication, with a focus on face recognition, in collaboration with SONY EuTec (funding partner).
- Explored cryptographic aspects of random matrices to enhance security and privacy in biometric systems.
- Developed novel approaches to improve facial authentication performance, leading to a patent [P1].
- Mentored PhD students and contributed to joint publications, fostering knowledge transfer within the research group.
- Actively coordinated with SONY researchers to align academic research with industrial objectives, ensuring practical relevance of outcomes.

Education

Polytechnic University of Turin March 2016

PhD Electronics and Telecommunications Engineering

- **Topics**: Bayesian Inference, Dictionary Learning, Multimedia signal processing and Compressed Sensing
- **Awards**: PhD Quality Award - Top 3 XXVIII cycle PhD student

Universidad de Granada (Spain) April- Sept 2015

PhD Internship

- Hosted at VIP lab (supervisor: prof. Rafael Molina). I worked on Image Processing tasks (e.g., denoising and inpainting) through Variational Bayes inference

Polytechnic University of Turin September 2012

MSc Telecommunications Engineering *110/110*

- **Majors**: Image and video coding, Signal processing, Information theory and codes

Polytechnic University of Turin September 2011

BSc Telecommunications Engineering

Patents

- [**P1**]: Matteo, Testa, et al. "Method and Apparatus for Image Recognition." U.S. Patent Application No. 17/376,195.
- [**P2**]: Matteo, Testa, et al. "Interactive Modification of Speaking Style of Synthesized Speech" U.S. Patent filed – 2023
- [**P3**]: Matteo, Testa, et al. "Adaptation and training of neural speech synthesis" U.S. Patent filed –2024

Selected publications

- [**C0**]: G. Ruggiero, M. Testa, J. Van de Walle and L. Di Caro "Eta-WavLM: Efficient speaker identity removal in self-supervised speech representations using a simple linear equation." ACL, 2025
- [**C1**]: G. Ruggiero, M. Testa, J. Van de Walle and L. Di Caro "Enhancing Polyglot Voices by Leveraging Cross-Lingual Fine-Tuning in Any-to-One Voice Conversion" EMNLP, 2024
- [**C2**]: A. Ali, M. Testa, T. Bianchi and E. Magli "BioMetricNet: deep unconstrained face verification through learning of metrics regularized onto Gaussian distributions" European Conference on Computer Vision (ECCV), 2020
- [**C3**]: M. Testa, A. Ali, T. Bianchi and E. Magli "Learning mappings onto regularized latent spaces for biometric authentication" IEEE 29th International Workshop on Machine Learning for Signal Processing (MLSP), 2019, Pittsburgh, PA, USA
- [**C4**]: A. Ali, M. Testa, T. Bianchi and E. Magli "AuthNet: Biometric Authentication through Adversarial Learning" IEEE 21st International Workshop on Multimedia Signal Processing (MMSP), 2019, Kuala Lumpur, Malaysia
- [**J1**]: J. Serra, M. Testa, R. Molina, A. K. Katsaggelos "Bayesian K-SVD Using Fast Variational Inference" IEEE Transactions on Image Processing, 2017
- [**B1**]: M. Testa, D. Valsesia, T. Bianchi, E. Magli "Compressed Sensing for PrivacyPreserving Data Processing" SpringerBriefs in Signal Processing, Springer, 2019

Others

Teaching:

- **Multimedia Signal Processing, MSc course** Topics: Image coding (lossless and lossy schemes) and Neural Networks, Politecnico di Torino

Awards:

- **NVIDIA GPU Grant** Awarded with a NVidia Quadro P6000 to support my Deep Learning research project
- **PhD Quality Award** Top 3 XXVIII cycle PhD student at Department of Electronics and Telecommunication, Politecnico di Torino
- **PhD scholarship** Politecnico di Torino

Skills

Languages and Frameworks: Python, Pytorch, Numpy, Pandas, Docker, LaTeX
Tools: Git/GitHub, Unix Shell