Georgios Milis, MEng

PhD Student in Computer Science - Researching AI safety and synthetic media

milis@umd.edu | +1 (240) 305-0146 | Google Scholar | LinkedIn | Github | Website

Education

PhD in Computer Science, *University of Maryland (UMD)* Aug. 2024 - present
Advisor: Prof. Heng Huang

MEng in Electrical & Computer Engineering, *National Technical University of Athens (NTUA)* Sep. 2018 - Apr. 2024
GPA: 9.38/10.0
Thesis: Text-driven Articulate Talking Face Generation, supervised by Prof. Petros Maragos

Skills

Technical - Python, PyTorch, Bash, git, C/C++, Matlab
Research - Machine Learning, Generative Models, Watermarking, Speech & Audio, NLP, Signal Processing
Relevant courses - Machine Learning Theory, AI policy, Advanced Numerical Optimization, Multimodal Foundation Models, LLM Security and Privacy, Computer Vision, Speech and Natural Language Processing

Publications

1. Wu, Y.*, **Milis, G.***, Chen, R.*, Huang, H. "Robust Distortion-Free Watermark for Autoregressive Audio Generation Models" *NeurIPS 2025*
2. Wu, Y.*, **Milis, G.***, Chen, R.*, Guo, J., Huang, H. "A Watermark for Auto-Regressive Speech Generation Models" *Interspeech 2025*
3. **Milis, G.**, Filntisis, P. P., Roussos, A., Maragos, P. "Neural Text to Articulate Talk: Deep Text to Audiovisual Speech Synthesis achieving both Auditory and Photo-realism" arXiv:2312.06613 (revised manuscript under final review for *International Journal of Computer Vision*)

Experience

Graduate Research Assistant, *UMD Institute for Advanced Computer Studies (UMIACS)* Aug. 2024 - present
Researching safety, authenticity, and robustness for AI-generated content. Aiming to contribute to transparency in the post-GenAI world.

Teaching Assistant, *UMD Department of Computer Science* Aug. 2024 - present
CMSC 422: Introduction to Machine Learning (Fall 2024, Spring & Fall 2025)

Junior Language Modeling Engineer, *Cerence Inc. ASR Language Modeling Team* Oct. 2023 - Jun. 2024
Actively maintained production models (both cloud and embedded on vehicles) for 3 top-priority languages, overseeing version upgrades. Patched ASR bugs at the n-gram, dictionary, and phonetic level. Developed automation scripts for testing and release, used across teams.

Projects

Vision Laguage Model Robustness 2025
Ongoing work to enhance VLM robustness to real-world image degradations using supervised finetuning and reinforcement learning (RL).

Audio Watermarking 2025
Introduced audio expertise to my research group and helped develop robust, audio-aware semantic watermarks for speech models. Contributed methodologically (proposed similarity clustering) and ran experiments. Published at Interspeech (I presented orally) and NeurIPS.

Text-driven Talking Face Generation 2023-2024
Conducted research at the intersection of speech synthesis, computer graphics, and generative modeling (transformers, GANs, diffusion models). Built an audiovisual deepfake model combining text-to-speech transformers with 3D head modeling priors and neural rendering. Showcased that bimodality enhances realism compared to the standard cascade of text-to-speech and speech-driven animation.

Notable course projects
Interpretable audio captioning (2024). LLM code security via RL (2024). NLP-based code analysis to prune hardware design space (2023).

Personal

Volunteering
IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2023

Awards
"The Big Moment for Education" Grant, *Eurobank* 2018
Honors in national Mathematics and Physics high school competitions 2017

Languages
Greek (native), English (Cambridge C2, TOEFL 115/120), French (Sorbonne C1), Italian (A2)