Federico Nocentini

PhD Student
Florence, Italy

(+39) 3384045134 federico.nocentini@unifi.it
Google Scholar GitHub Personal page

Experience

- **Research Scientist Intern** — **Los Angeles, California**
  Flawless AI — *November 2024 - May 2025*
  Multilingual Speech-Driven facial animation.

Education

- **PhD in 3D Computer Vision and Artificial Intelligence** — **Florence, Italy**
  University of Florence - Media Integration and Communication Center — *November 2022 - Present*
  Topics of research: Deep Learning; 3D Computer Vision; 3D Talking Heads Generation.
- **M.Sc in Computer Science and Engineering** — **Florence, Italy**
  University of Florence, Grade: 110/110 (GPA: 4.0/4.0) — *February 2020 - October 2022*
  Thesis: Generating Emotional 3D Talking Heads from Audio Spectrograms.
- **B.Sc in Computer Science and Engineering** — **Florence, Italy**
  University of Florence, Grade: 104/110 (GPA: 3.8/4.0) — *September 2016 - February 2020*
  Thesis: Facial recognition with depth images generated from a deformable model of the face.

Publications

- **Learning Landmarks Motion from Speech for Speaker-Agnostic 3D Talking Heads Generation (ICIAP 2023)**
  A novel approach for generating 3D talking heads from raw audio inputs. This method grounds on the idea that speech related movements can be comprehensively and efficiently described by the motion of a few control points located on the movable parts of the face, i.e., landmarks.
- **ScanTalk: 3D Talking Heads from Unregistered Scans (ECCV 2024)**
  ScanTalk is a novel framework for 3D Talking Heads generation capable of animating any 3D faces in arbitrary topologies including scanned data.
- **EmoVOCA: Speech-Driven Emotional 3D Talking Heads (WACV 2025)**
  EmoVOCA is a new synthetic dataset combining emotional nuances with 3D talking heads. On such data, drawing from state-of-the-art techniques, we designed and trained a deep architecture that generates expressive 3D talking heads using an audio, an emotion, and intensity labels.
- **3D Face Reconstruction Error Decomposed:**
  A Modular Benchmark for Fair and Fast Method Evaluation (FG 2025)
  Current benchmark tools are monolithic, even though there is no consensus on the best way to measure error. We present a modularized codebase, where the fundamental components of error computation are segregated and interchangeable, allowing one to quantify the effect of each.
- **Beyond Fixed Topologies:**
  Unregistered Training and Comprehensive Evaluation Metrics for 3D Talking Heads (arxiv 2025)
  We highlight the limitations of current evaluation metrics and propose new metrics for better lip-syncing evaluation between speech and facial movements. Our extensive evaluation shows our approach performs favorably compared to fixed topology techniques, setting a new benchmark by offering a versatile and high-fidelity solution for 3D talking head generation.

Projects

- **Citation intent classification**

  A model built with Pytorch for citation intent classification in scientific papers using the Scicite dataset. To leverage the connection between the structure of scientific papers and the intent of citations, we studied a multitask framework with two structural scaffolds (auxiliary tasks) related to the structure of scientific documents.

- **Outfit colors classification with CNN**

  A CNN model built with Pytorch for outfit classification based on colors. The classes are taken from Shigenobu Kobayashi's book Color Image Scale.

- **Mean shift clustering with CUDA and OpenMP**

  Sequential C++ and parallel CUDA and OpenMP implementations of the mean shift clustering algorithm. The execution times obtained for datasets of increasing dimension are compared to measure the speedup of the parallel versions.

- **Depth of Field and Motion Blur**

  Implementation of Depth of Field and Motion Blur as postprocessing effects using Three.js

- **Social distancing detector**

  A tool built with Python to monitor people's compliance to social distancing in crowded places in real-time. It uses a neural network based on Keras for people detection and OpenCV for measuring distances in various types of input video streams.

- **K-means clustering with CUDA and OpenMP**

  Sequential Python and parallel CUDA and OpenMP implementations of the k-means clustering algorithm. The execution times obtained for datasets of increasing dimension are compared to measure the speedup of the parallel versions.

Languages

- **Italian**: Mother tongue
- **English**: Full professional proficiency

Skills

Frameworks

PyTorch
Tensorflow
Open3D
PyTorch3D
Trimesh
OpenCV
Numpy, Pandas
Matplotlib

Programming

Python
C/C++
Java
Javascript
Cuda

Other

Most projects are available on my GitHub and publications on Google Scholar.