Simone Sorrenti

simone.sorrenti.21@gmail.com | **+39 3899176181** | **GitHub** | **Website** | **LinkedIn**

SKILLS

Programming Languages: Python, C++, Java, JavaScript, PHP, HTML, CSS, SQL
Machine Learning: PyTorch, TensorFlow, Hugging Face, OpenCV, Numpy, Pandas, Matplotlib, Seaborn
Technologies & Tools: CUDA, ROS2, Linux, Git
Languages: English (B2), Italian (Native)

EXPERIENCE

- **R&D AI & Robotics Engineer**, TXT E-Tech — Milan, Italy | Jun 2024 – Present
  - Designed deep learning models for object detection and OCR, enhancing cockpit condition assessment by autonomously detecting components like screens, buttons, and knobs and extracting textual and symbolic information.
  - Developed an audio matching algorithm, improving cockpit audio verification by preprocessing audio (removing noise, silences, pre-emphasis), extracting features (MFCC, CHROMA), and applying DTW and cosine similarity for precise playback detection.
  - Collaborated with clients to define software requirements and developed deep learning models and GAN algorithms, improving anomaly detection accuracy in solar cells of satellite panels.
  - Led a client-facing project to create a dataset for a Kaggle competition, ensuring quality by coordinating frame extraction, cropping, blurring sensitive information, and annotating bounding boxes.
- **AI Researcher**, Polytechnic of Milan, [Thesis] [GitHub] — Milan, Italy | Feb 2023 – Apr 2024
  - Improved illegal waste detection by developing deep learning models (CNN & ViT) using satellite imagery for supervised classification and weakly supervised segmentation, resulting in enhanced detection accuracy for illegal landfills.
  - Enhanced model performance and expanded the dataset by addressing data scarcity and class imbalance through data augmentation techniques, class weights, and oversampling, improving landfill detection accuracy and robustness.
  - Standardized satellite images by developing a preprocessing pipeline, ensuring consistent data quality across datasets.
  - Boosted classification accuracy to 90% by employing weakly supervised learning with Hierarchical Heatmap Generation and Multiple Instance Learning (MIL), enabling area localization and pixel-level probability assignment while overcoming hardware limitations.
- **IT Consultant**, Blue Reply — Milan, Italy | Jun 2020 – Aug 2021
  - Ideated and implemented a back-end algorithm to create customizable document templates, improving document generation efficiency by automating layout instructions from the front-end and ensuring accurate component placement and pagination.
  - Developed and implemented dynamic, user-friendly front-end pages for an insurance service, enhancing user experience by increasing usability and creating API microservices for improved functionality and performance.
- **Application Development Analyst**, Accenture — Milan, Italy | Jan 2020 – Jun 2020
  - Developed interactive web applications for the insurance and banking sectors, enhancing client satisfaction by working closely with cross-functional teams to identify needs and deliver tailored solutions.

EDUCATION

- **M.Sc. in Computer Science and Engineering (Artificial Intelligence)**, Polytechnic of Milan
  Sep 2020 – Apr 2024 | **GPA**: 3.74/4.0, **Italian Grade**: 110/110
  Thesis: PERIVALLON - Detection of Illegal Landfills using Deep Learning: A Weakly Supervised Approach
  Coursework: Machine Learning, Artificial Neural Networks and Deep Learning, Recommender Systems, Robotics, Natural Language Processing, Mathematical models and methods for Image Processing, The rise of Transformers
- **B.Sc. in Informatics**, University of Bari
  Sep 2014 – Oct 2018 | **GPA**: 3.79/4.0, **Italian Grade**: 110L/110
  Thesis: Solar radiation prediction through Machine Learning algorithms
  Coursework: Programming Languages, Discrete Mathematics, Numerical Calculus, Databases, Software Engineering, Software Systems Integration and Testing, Mobile Software Development

ACADEMIC & PERSONAL PROJECTS

- **Contrastive Language-Image Pre-Training:** Explored CLIP models for image captioning, zero-shot classification, image retrieval and clustering tasks across diverse datasets, including medical and fashion data. [Report] [GitHub]
- **Mathematical Models and Methods for Image Processing**: Explored advanced image processing techniques, including sparse coding and robust fitting methods, through mini projects on denoising, inpainting, and anomaly detection. [GitHub]
- **Question Answering System:** Implemented a question answering system with BERT, GPT-2, and T5 on the SQuAD dataset, focusing on answer extraction and response generation. [GitHub]
- **Autonomous Mapping and Navigation:** Created and visualized environmental maps using 2D and 3D laser data. Implemented waypoint-based navigation with autonomous movement and localization for effective pathfinding and obstacle avoidance. [GitHub]
- **Applying Vision Transformers for Brain Tumor Classification:** Implemented the Vision Transformer from scratch for classifying tumors in MRI images. [GitHub]
- **Recommender System:** Developed a recommender system for TV shows utilizing Content-Based Filtering, Collaborative Filtering, Context-Aware, Graph-Based and Hybrid approaches. Focused on user interaction data and show features. [Kaggle]
- **Prediction of Solar Radiation Through Machine Learning Algorithms:** Developed a system for predicting solar radiation, utilizing machine learning algorithms and applied feature selection techniques to identify key meteorological parameters. [Thesis]