Anxhelo Diko

Ph.D. Fellow, Department of Computer Science, Sapienza University of Rome

PERSONAL DETAILS

Birth	February 19, 1997
Address	Via Salaria 113, 00198 Rome, Italy
Phone	(+39) 3497921839
Mail	diko@di.uniroma1.it
Profile	www.linkedin.com/in/anxhelodiko97
Personal Website	https://anxhelodiko.dev

PERSONAL STATEMENT

A hardworking and ambitious individual with a great passion for Computer Vision. Currently, I am enrolled in the second year of the Ph.D. in Computer Science program at the Sapienza University of Rome. As a Ph.D. fellow, my main research focuses on artificial intelligence and computer vision, believing that pixels, the closest representation of the physical world constraints, are the key to next-generation artificial general intelligence. More specifically, I am currently focusing my research on human activity understanding and anticipation from camera sensors, which is a key problem for many autonomous agents relying on perception as their main driving force and for virtual reality (VR). Additionally, another issue that has gotten my interest, directly related to the fantastic world of activity understanding, is the view-invariant representation of learning from videos. A capability that today's agents do not possess but could be beneficial to see and perceive the same thing from different perspectives.

Besides my research interest, I have a strong engineering background cultivated throughout my academic and industry journey. I possess remarkable programming prowess, particularly in Python, my principal tool for rapid prototyping, followed by essential versioning tools and deep/machine learning frameworks. Furthermore, my adeptness extends to C++ and parallel programming using CUDA. This engineering foundation and my theoretical and practical knowledge of the computer science and vision fields enable me to easily adapt my ideas into working prototypes following the best practices.

Finally, I aspire to contribute substantially to the continually evolving landscape of computer vision and deep learning through advanced research and innovative endeavors. Specifically, I intend to significantly contribute to the current state of action anticipation and view-invariant representation learning.

WORK EXPERIENCE

Contracted Machine Learning Researcher *2021-Ongoing*

Sapienza University of Rome (24 months)

- Conduct research in various computer vision problems like anomaly detection, action recognition, action anticipation, object detection and vision-language models.

Machine Learning Engineer *2020-2021*

MedLea Srls, (16 months)

- Designed and deployed machine learning models at scale.
- Designed and implemented a parallel and scalable Ray-Tracing algorithm that runs on GPUs for discretizing 3D mesh representation of geometries into a volumetric representation. The implemented algorithm would cut the computational costs of the services offered by MedLea by 30% in the preparation phase.

Applied machine learning specialist *2018-2018*

PaperClicks, Internship (6 months)

- Responsible for designing and implementing machine learning algorithm for the optimization of core business operations at PaperClicks.

EDUCATION

Ph.D. in Computer Science *2021-2025*
Sapienza University of Rome

- Research Area/s: Computer vision and activity understaning for autonomus agents.
- Advisor: Prof. Luigi Cinque
- Co-advisor: Prof. Danilo Avola

MSc. Computer Science, Sapienza University of Rome *2018-2020*

- Important Courses: Machine Learning, Computer Vision, Applied Artificial Intelligence, Multimodal Interaction, Cloud Computing, Advanced Software Engineering, Distributed Systems.
- Graduated as top 1% of the class.
- Finsl Grade: Cum laude

BSc. Business Computer Science, University of Tirana *2015-2018*

- Important Courses: Algorithms, Data Structures, Lnear Algebra, Calculus, C++, Java, Computer Architecture, Databases, Computer Networks, Information Security, Statistics.
- Thesis: Graduated as top 1% of the class.
- Final Grade: 10/10

HIGHLIGHTED PROJECTS

SEARCHER – Smart unmannEd AeRial vehiCles for Human likE monitoRing *2022-Ongoing*
Italian Ministry of Defense
Senior R&D Engineer

- Study and analysis about state-of-the-art anomaly detection, novelty detection, and UAV attention mechanism algorithms.
- Design and development of novel deep learning algorithms for UAV application in navigation and surveillance.
- Engineering the execution pipeline of the deep learning system on edge.

The project has raised funds from the Ministry of Defense.

Egocentric Action Anticipation With Causality Modeling and Latent Space Prior Injection from Text *2023-Ongoing*
Individual Project, part of PhD research

- Designed and implemented neural networks that could anticipate future human action from RGB videos.
- Introduced two novel attention mechanisms that model the causality between video events happening in different time steps and preserve temporal order.
- Introduced a novel learning approach for action anticipation which bridges the gap between visual interpretation of the future and text.
- Achieved competitive results on three benchmarks, namely EpicKtchens-100/55 and EGTEA++.

View-Invariant Action Anticipation exploiting relative similarities from RGB videos *2023-Ongoing*
Individual Project, part of PhD research

- Align the understanding of actions from multiple different non-synchronized views.
- Exploit the relative similarities between different actions to enable zero-shot learning.

ReViT – Enhancing Vision Transformers With Residual Attention *2022-2023*
Individual Project, part of PhD research

- Designed and implemented a novel residual connection between transformer blocks that propagates and accumulates knowledge from shallow to deeper layers.
- Enhanced Vision Transformers performance by improving their feature diversity in deeper layers.
- Obtained an improvement of +4% on ImageNet1K

Enabling Smart Assistants Communication with Deaf and Mute People Through Sign-Language Detection and Text-To-Speech *2019-2020*

Individual Project

- Designed and implemented a solution that could read sign languages from an RGB camera, translate the sign language into text, transform the text into speech to communicate with a smart assistant, capture the smart assistant answer and convert it into written text.
- The implemented solution could help mute and deaf people communicate with smart assistants controlled over voice commands.

SKILLS

Skills Summary Hard Skills

- **Languages:** Albanian (mother tongue), Italian (proficient), English (proficient).
- **Professional Competences:** Algorithms, computer vision, image processing, video processing, machine learning, deep learning, representation learning, visual recognition, action anticipation, action understanding, object detection, semantic segmentation, homography, programming, parallel computing, transformer neural networks, convolutional neural networks, Unit testing, CI/CD, debugging, Prompting.
- **Programming Languages:** Python 3 (5+ years), C++ (2+ year), CUDA (1+ year), SQL (1+ year).
- **Tools and Frameworks (Expert):** PyTorch, OpenCV, NumPy, SciPy, Scikit-learn, Pandas, Matplotlib, PlotLy, Docker, Bash, git, MPI, OpenMP, AWS, Jenkins.
- **Tools and Frameworks (Proficient):** PyTorch-Lightning, Tensorflow, HuggingFace api, Detectron2, MMDetector, DLlib, CuPy, PlotLy, Scikit-Image, MongoDB, MySQL, VTK, ITK, PoreSpy, MPI, OpenMP, AWS, Jenkins, Kubernetes.
- **Operating System:** Debian-based Linux (Ubuntu, Mint).

Soft Skills

- Communication, teamwork, attention to detail, problem-solving, adaptability, time management, work ethic, perseverance, consistency, and persistence.

PUBLICATIONS

Research articles

1. "Egocentric Action Anticipation With Causality Modeling and Latent Space Prior Injection from Text" - (TPAMI, Under review).
2. "ReViT: Enhancing Vision Transformers with Residual Attention." - (Pattern Recognition, Under Review) `https://github.com/ADiko1997/Vision.ai-PhD/tree/main/ReViT`.
3. "Low-Altitude Aerial Video Surveillance via One-Class SVM Anomaly Detection from Textural Features in UAV Images." - Information 2022, 13, 2. `https://doi.org/10.3390/info13010002`.
4. "A Novel GAN-Based Anomaly Detection and Localization Method for Aerial Video Surveillance at Low-Altitude." - Remote Sens. 2022, 14, 4110. `https://doi.org/10.3390/rs14164110`.
5. "MS-Faster R-CNN: Multi-Stream Backbone for Improved Faster R-CNN Object Detection and Aerial Tracking from UAV Images." - Remote Sens. 2021, 13, 1670. `https://doi.org/10.3390/rs13091670`.

AWARDS AND FELLOWSHIPS

1. Research Fellowship, Sapienza University of Rome, IT (2021).
2. Research Fellowship, Sapienza University of Rome, IT (2022) (Ph.D. grants).
3. Research Fellowship, Sapienza University of Rome, IT (2023) (Ph.D. grants).
4. LazioDisco Scolarship, Rome, IT (2018 - 2020).

CERTIFICATE

1. Advanced C++ Developer (Udemy).