Nikolaos Gkanatsios

ngkanats@andrew.cmu.edu
nickgkan.github.io

Education

2020 – Present
Carnegie Mellon University, Robotics Institute, PhD Student.
- Advisor: Katerina Fragkiadaki, Associate Professor, MLD, CMU
- Research on Continual Learning for Vision and Robotics

2011 – 2017
National Technical University of Athens (NTUA), *ECE*, B.Sc/M.Sc. (5-year diploma).
- GPA 9.22/10, 9th among 328 in class, 1st in Electronics and Systems major
- Thesis on *"Multimodal Classification and Segmentation of Fine-Grained Activities in Videos"*
- Thesis supervisor: Petros Maragos, Professor, ECE, NTUA

2008 – 2011
3rd Lyceum of New Heraklion, Attika, *General Lyceum Graduation*.
- General Graduation Grade 19.7/20
- General Panhellenic Examinations Grade 19.24/20, top 1% nationwide

---

Research Collaborations with non-Academic Institutions

Jan 2022 – Jun 2023
Microsoft Turing Academic Program (MS-TAP).
- CMU-MSTAP project *"Learning instructible visuo-motor agents through multimodal interactive teaching"*

Dec 2021 – Apr 2023
Amazon Alexa Prize SimBot Challenge.
- Developed an embodied conversational agent as member of the CMU team Symbiote. Technical report

---

Industrial Research Experience

Aug 2018 – Jul 2020
Deeplab, *Machine Learning Engineer*, Athens.
- Research on scene graph generation, in collaboration with Prof. Maragos' lab at NTUA.
- Built an interactive demo of scene graph generation on videos.
- Developed the backend of this webpage for real-time inference of visual relationships on wild images.

Mar 2017 – Jul 2018
METIS Cybertechnology S.A., *Artificial Intelligence Engineer*, Athens.
- Designed and developed the whole natural language processing pipeline behind METIS's textbot.
- Developed several applications and APIs for METIS's products.

Research Collaborations with Academic Institutions

Nov 2019 – Jan 2020 and Apr – Jun 2020
Stanford University, Leonidas Guibas Lab, Stanford, CA, USA, (remotely).
- Worked with Panos Achlioptas on fine-grained language grounding and shape deformation. Recommendation can be provided by Panos.

Dec 2019 – Feb 2020
TU Darmstadt, Intelligent Autonomous Systems Group, Germany, (remotely).
- Worked with Dr. Georgia Chalvatzaki on improving tensor representations for grasping point detection. Recommendation can be provided by Georgia.

Aug 2018 – Aug 2020
Computer Vision and Signal Processing Lab, NTUA, Athens, Greece.
- Research on high-level semantic image representation, with application to scene graph generation, in collaboration with Deeplab.

Publications

2024
3D Diffuser Actor: Policy Diffusion with 3D Scene Representations, *In submission*. PDF Website

2024
Diffusion-ES: Gradient-free Planning with Diffusion for Autonomous Driving and Zero-Shot Instruction Following, *CVPR 2024*. PDF Website

2024 **ODIN: A Single Model for 2D and 3D Perception**, Jain A., Katara P., Gkanatsios N., Harley A, Sarch G., Aggarwal K., Chaudhary V., Fragkiadaki K., *CVPR 2024.* PDF Website

2023 **Analogy-Forming Transformers for Few-Shot 3D Parsing**. Gkanatsios N.*, Singh M.*, Fang Z., Tulsiani S., Fragkiadaki K., *International Conference on Learning Representations, ICLR 2023.* PDF Website

2023 **Energy-based Models are Zero-Shot Planners for Compositional Scene Rearrangement**. Gkanatsios N.*, Jain A*, Zhou X., Zhang Y., Atkeson C., Fragkiadaki K., *Robotics Science and Systems, RSS 2023.* PDF Website

2023 **Unifying Diffusion Models with Action Detection Transformers for Multi-task Robotic Manipulation**. Zhou X.*, Gkanatsios N.*, Gervet T.*, Fragkiadaki K., *Conference on Robot Learning, CoRL 2023.* PDF Website

2023 **Act3D: Infinite Resolution Action Detection Transformer for Robot Manipulation**. Gervet T.*, Zhou X.*, Gkanatsios N., Fragkiadaki K., *Conference on Robot Learning, CoRL 2023.* PDF Website

2022 **Bottom Up Top Down Detection Transformers for Language Grounding in Images and Point Clouds**. Jain A.*, Gkanatsios N.*, Mediratta I., Fragkiadaki K., *European Conference on Computer Vision, ECCV 2022.* PDF Code

2021 **Grounding Consistency: Distilling Spatial Common Sense for Precise Visual Relationship Detection**. Diomataris M., Gkanatsios N., Pitsikalis V., Maragos P., *IEEE International Conference on Computer Vision, ICCV 2021.* PDF Code

2020 **From Saturation to Zero-Shot Visual Relationship Detection Using Local Context**. Gkanatsios N., Pitsikalis V., Maragos P., *BMVA British Machine Vision Conference, BMVC 2020.* PDF Code

2020 **Revisiting Grasp Map Representation with a Focus on Orientation in Grasp Synthesis**. Gkanatsios N., Chalvatzaki G., Maragos P., Peters J., *Visual Learning and Reasoning for Robotic Manipulation, RSS Workshop.* PDF Code

2019 **Attention-Translation-Relation Network for Scalable Scene Graph Generation**. Gkanatsios N., Pitsikalis V., Koutras P., Maragos P., *IEEE International Conference on Computer Vision Workshops, ICCVW 2019.* PDF Code

2019 **Showcasing Deeply Supervised Multimodal Attentional Translation Embeddings: A Demo for Visual Relationship Detection**. Gkanatsios N., Pitsikalis V., Koutras P., Zlatintsi A., Maragos P., *IEEE International Conference on Image Processing, ICIP 2019.* Paper Demo

2019 **Deeply Supervised Multimodal Attentional Translation Embeddings for Visual Relationship Detection**. Gkanatsios N., Pitsikalis V., Koutras P., Zlatintsi A., Maragos P., *IEEE International Conference on Image Processing, ICIP 2019.* PDF Code

2018 **Metis: A Scalable Natural-Language-Based Intelligent Personal Assistant for Maritime Services**. Gkanatsios N., Mermikli K., Katsikas S., *24th International Conference on Information and Software Technologies (ICIST 2018).* Paper

―――― Talks / Oral Presentations

July 2023 **RSS 2023**, Daegu, South Korea: oral presentation of the paper "Energy-based Models are Zero-Shot Planners for Compositional Scene Rearrangement".

Sep 2019 **ICIP 2019**, Taipei, Taiwan: oral presentation of the paper "Deeply Supervised Multimodal Attentional Translation Embeddings for Visual Relationship Detection".

Oct 2018 **ICIST 2018**, Vilnius, Lithuania: oral presentation of the paper "Metis: A Scalable Natural-Language-Based Intelligent Personal Assistant for Maritime Services".

Jun 2018 **VOXXEDDAYS Athens**, Greece: A Scalable Maritime Platform Providing Services Through an NLP-Based Intelligent Personal Assistant

Teaching Experience

Jan 2023 – May 2023 **CMU 16-726 Teaching Assistant**, Pittsburgh.
- ○ Learning-based Image Synthesis, with Prof. Jun-Yan Zhu.

Sep 2022 – Dec 2022 **CMU 16-824 Teaching Assistant**, Pittsburgh.
- ○ Visual Learning and Recognition, with Prof. Jun-Yan Zhu.

Sep 2019 – Sep 2020 **CVSP Teaching Assistant**, Athens.
- ○ I co-supervised an undergraduate student's diploma thesis on visual relationship detection, together with Prof. Petros Maragos.

Honors and Awards

2018 **Georgios Kontaxis Award** for being the 1st in 2017 class in Electronics and Systems major, ECE NTUA.

2018 **Christos Chalkias Award** for being the 1st in 2017 class in Electronics and Systems major, ECE NTUA.

2017 **EESTECH Challenge Athens 1st Prize** for winning the Machine Learning Hackathon, hold by EESTEC Athens and ECE NTUA.

2012 **Hellenic Physical Society** Award for making the top 10 in the final round of the Panhellenic Physics Competition for high-school students on a nationwide level.

2010 – 2011 **Awards from Hellenic Police** for excellence in high-school classes.

Research Interests

Computer Vision, Robotics, Natural Language Processing, Embodied AI

Continual Learning, In-context Learning

Programming

(Experienced) Python, Matlab / (Basic) C, Javascript (NodeJS), SQL / PyTorch

Languages

Greek (Native), English (Michigan Certificate of Proficiency in English), French (Delf B1)

Personal

Basketball, drums (self-taught), bass guitar (self-taught), ancient Greek literature, music transcription, Greek folk dancing.