Nimrod Shabtay - Curriculum Vitae

Personal website — nimrod.shabtay@gmail.com — +972(0)-54-4828863

Education

- **2022 - Present: PhD (Direct Track) in Electrical Engineering, Tel-Aviv University, Israel**
  Specializing in Multimodal Foundation Models: Evaluation, Adaptation, and Improvements, supervised by Prof. Raja Giryes
  Current GPA: 92/100
- **2012 - 2016: B.Sc. in Electrical Engineering, Tel-Aviv University, Israel**
  Expertise in Image Processing, Computers and Bio-Imaging
  Final GPA: 85/100
  Final Project: Real Time algorithm for rectification of stereo images, 98

Publications

- **CARES: Context-Aware Resolution Selector for VLMs**
  Moshe Kimhi*, **Nimrod Shabtay***, Raja Giryes, Chaim Baskin, Eli Schwartz. Under Review. ***link***
- **Spoken question answering for visual queries**
  Nimrod Shabtay*, Zvi Kons*, Avihu Dekel, Hagai Aronowitz, Ron Hoory, Assaf Arbelle. Interspeech 2025. ***link***
- **Advancing Speech Understanding in Speech-Aware Language Models with GRPO**
  Avishai Elmakies, Hagai Aronowitz, **Nimrod Shabtay**, Eli Schwartz, Ron Hoory, Avihu Dekel. Under Review. ***link***
- **Teaching VLMs to Localize Specific Objects from In-context Examples**
  Sivan Doveh*, **Nimrod Shabtay***, Wei Lin, Eli Schwartz, Hilde Kuehne, Raja Giryes, Rogerio Feris, Leonid Karlinsky, James Glass, Assaf Arbelle, Shimon Ullman, M. Jehanzeb Mirza. ICCV 2025. ***link***
- **LiveXiv – A Multi-Modal Live Benchmark Based on Arxiv Papers Content**
  Nimrod Shabtay, Felipe Maia Polo, Sivan Doveh, Wei Lin, M. Jehanzeb Mirza, Leshem Chosen, Mikhail Yurochkin, Yuekai Sun, Assaf Arbelle, Leonid Karlinsky, Raja Giryes. ICLR 2025. ***link***
- **Deep Phase Coded Image Prior**
  Nimrod Shabtay, Eli Schwartz, Raja Giryes. ICCP 2024. ***link***
- **PIP: Positional-encoding Image Prior**
  Nimrod Shabtay*, Eli Schwartz*, Raja Giryes. Under Review. ***link***

Professional Experience

- **2024 - Present: Research Intern, IBM Research**
  Conducting research on multi-modal foundation models with emphasis on vision-language understanding and speech processing. Work includes developing robust evaluation methodologies, enhancing model capabilities for context-aware visual reasoning, and advancing speech-vision integration. Contributing to efficient model inference techniques and exploring novel training approaches for improved multi-modal performance.
- **2023: Summer Research Intern, IBM Research**
  Enhancing visual reasoning with structured in-context learning:
  - Represent images as structured data using multi-modal models and LLMs
  - Using in-context learning over structured data to perform visual reasoning using Code-LLMs
  - Enhance visual programming using RAG for questions representation

- **2018 - 2023: Senior Computer Vision Algorithm Developer, Nanit**
  - Lead researcher for end-to-end development of deep learning models in a variety of computer vision tasks such as object detection, image classification, and pose estimation
  - Lead developer of a real-time computer vision system monitoring breathing movement - from PoC to code implementation on an edge device
  Media appearances:
    - Guidelines for building an accurate and robust ML/DL model in production
    - Guest speaker at "Gradient Dissent": How Nanit improves and develops models

- **2014 - 2018: Computer Vision and Image Processing Algorithm Developer, CEVA-DSP**
  - Designed and implemented algorithms for various IP/CV tasks, from initial design to hardware specific implementation. Mainly focused on detection, tracking, SLAM, 2D/3D registration, model estimation for applications such as low-light imaging and image stabilization
  - Research and development of Gradient-Boost Tree module for real-time validating registrations
  - Developed an optimized driver for CNN accelerator

Military Service

- **2004 - 2011: Naval Officer, Electronic Department, Israel Defense Force (IDF)**
  Rank: Major (reserve), Regiment: Navy
  - Management of a team of 30 staff members in various fields of electronics, such as radar, communications, optics, electronic warfare, control systems and electricity
  - Promoted to tech lead of the above fields at the fleet HQ

Programming and Software Knowledge

Python, C/C++, Matlab, PyTorch, TensorFlow, Linux and open-source software

Languages

Hebrew (Native), English (Professional)