Amit Israeli

Deep Learning Research Engineer

+972 53-432-6597 | Israel | amit154154 | personal website | Amit Israeli | Amit Israeli

Profile

Generative AI Research Engineer specializing in text-to-video and text-to-image diffusion, multimodal models, and computer vision. Experienced in fine-tuning large models (diffusion, autoregressive, and GANs) and building end-to-end pipelines from data curation and training through evaluation to deployment on GPUs and edge devices.

Projects

Evolution Strategy Post-Training for Text-to-Image

- Implemented the **EGGROLL (Evolution Strategies at the Hyperscale)** paper (arXiv: 2511.16652) for **T2I post-training**: low-rank ES updates on **LoRA adapters** with a frozen **Sana** backbone (no diffusion backprop / no PPO).
- Adapted EGGROLL to **one-step Sana Sprint** for near **inference-speed** optimization with **population ES** (e.g., 128 candidates) and stability controls (shared seeds, antithetic sampling, norm/explosion guards).
- Achieved measurable gains on **PartiPrompts** (e.g., improved **text similarity** and **win-rate**) vs. the one-step baseline.

SANA-Video × Tom & Jerry — Text-to-Video LoRA

- Fine-tuned the **2B-parameter SANA-Video T2V diffusion model** with LoRA to generate controllable Tom & Jerry–style clips on a single consumer GPU.
- Built an **automatic scene-labeling pipeline** with Qwen3-VL served via vLLM to generate structured prompts (environment, characters, actions, camera), then used them for **text-conditioned training and evaluation**.
- Open-sourced the code, dataset tooling, and experiment logs as a reproducible project (GitHub / Hugging Face / W&B).

PopYou - VAR Text

- Fine-tuned a Visual AutoRegressive (VAR) model with synthetic and real Funko Pop! custom datasets, adding a new *Funko Pop!* class.
- Designed a two-stage paradigm inspired by *Bridging CLIP and StyleGAN through Latent Alignment for Image Editing* (arXiv:2210.04506): first reconstructing images from SigLIP embeddings, then training a text pathway to enable promptable text-to-image Funko Pop! generation.

Professional Experience

Computer Vision Research Engineer,
Reality Defender
2025 Jan – present
Develop multimodal deepfake and fraud detection models for video+audio, images, and text, with both classification and explanation capabilities, focusing on detecting AI-generated and manipulated media.

Deep Learning Research Engineer, *NLPearl*
2024 Jun – 2025 Jan | Tel Aviv, Israel

- Developed real-time conversational systems for pause detection and response generation using fine-tuned LLMs.
- Explored LoRA-based and multi-stage training setups to boost performance under latency and memory constraints.
- Built a compact language model for multi-task outputs and worked with SOTA audio tokenizers and LLMs for audio-focused tasks.

Computer Vision Research Engineer,
LuckyLab-Freelance
2024 Dec – 2025 Apr

- Built and deployed edge-optimized solutions for segmentation and object detection for production.

Computer Vision and Deep Learning Research Engineer, *Pashoot Robotics*
2023 May – 2024 Jun | Rehovot, Israel

- Improved object detection and segmentation in zero/few-shot settings using foundation models like SAM, YOLO-World, Grounding DINO, and CLIP.
- Worked on object tracking and 6D pose estimation with synthetic data.
- Used Blender and 3D reconstruction techniques (NeRF, Gaussian Splatting, image-to-3D) for simulation tasks.

Education

Ben-Gurion University,
Discontinued BSc in Data Science
2020 – 2023 | Ben-Gurion, Israel
At the start of my final academic year, I transitioned to a professional role as an Algorithms Engineer at Pashoot Robotics, gaining valuable hands-on experience aligned with my career goals.