



DELYAN BOYCHEV

@ delyan.boychev@insait.ai | [in d-boychev](#) | [github delyan-boychev](#) | [scholar](#)

Sofia, Bulgaria

ABOUT ME

CS undergraduate and INSAIT research scholar working on vision-language models for satellite imagery and 3D scenes, advised by Dr. Danda Paudel. Currently an NLP research intern at Graphwise and an IOAI Lecturer for the Bulgarian IOAI team. Previously a Quant ML research intern at Man Group, and worked on synthetic-image detection, adversarial robustness, and interpretability in computer vision.

EDUCATION

INSAIT Scholar (highly-selective 4-year program), Faculty of Mathematics and Informatics, Sofia University <i>Bachelor of Computer Science</i>	Sofia, Bulgaria Sept. 2024 – Present
High School of Natural Sciences and Mathematics "Vasil Drumev" <i>Software and Hardware Sciences, GPA: 6.00/6.00</i>	Veliko Tarnovo, Bulgaria Sept. 2019 – June 2024

LANGUAGES

- Bulgarian — Native
- English — IELTS Academic, 7.0/9.0

RESEARCH INTERESTS

- Vision-Language Models
- 3D Scene Understanding & Spatial Reasoning
- Remote Sensing & Satellite Imagery
- Representation Learning

TECHNICAL SKILLS

- **Languages:** Python, C++
- **ML/Research:** PyTorch, Hugging Face Transformers
- **Data Science:** NumPy, pandas
- **Tools:** Docker, Git, Linux

EXPERIENCE

NLP Research Intern Graphwise • Conducting NLP research on extracting and structuring information from text.	July 2026 - Present
IOAI Lecturer Bulgarian IOAI Team • Lecturing and mentoring students preparing for the International Olympiad in Artificial Intelligence (IOAI).	2025 - Present
Undergrad Research Scholar INSAIT • Doing research at INSAIT in the area of Computer Vision - mostly working on 3D vision processing and Vision-Language models.	Oct 2024 - Present
Research Intern Man Group • Conducted Quant ML research at Man Group, focusing on developing and evaluating machine-learning models for systematic trading and signal forecasting.	July 2025 - Sept. 2025

HONORS

- Scholarship Recipient, EXPLORER Program, INSAIT (2024)
- European Space Agency Award from EUCYS'24, Katowice, Poland
- International Olympiad Artificial Intelligence 2024, Team Bronze Medal
- Regeneron ISEF'23 Finalist, Dallas, Texas
- John Atanasoff Presidential Award in the category "Debut breakthrough in computer technology" (2023)

PUBLICATIONS

OSMDA: OpenStreetMap-based Domain Adaptation for Remote Sensing VLMs | *arXiv Preprint, 2026*

- Stefan Maria Ailuro, Mario Markov, Mohammad Mahdi, **Delyan Boychev**, Luc Van Gool, Danda Pani Paudel
- A self-contained adaptation framework that uses OpenStreetMap data to generate geographic supervision without manual labels or a stronger external teacher.
- <https://arxiv.org/abs/2603.11804>

ImagiNet: A Multi-Content Benchmark for Synthetic Image Detection | *Workshop paper, 2024*

- **Delyan Boychev**, Radostin Cholakov
- AAAI 2025 Workshop on Datasets and Evaluators of AI Safety.
- A balanced 200K-image benchmark spanning four content types and multiple open and proprietary generators, with strong contrastive-learning baselines.
- <https://arxiv.org/abs/2407.20020>

Interpretable Computer Vision Models through Adversarial Training | *arXiv Preprint, 2023*

- **Delyan Boychev**
- An empirical study of how adversarial training changes learned representations and makes computer-vision model decisions more interpretable.
- <https://arxiv.org/abs/2307.02500>

PROJECTS

LavBench | *ML Competition Platform* 2026 - Present

- A secure, bilingual platform for machine-learning competitions, built for Bulgarian AI Olympiad selection and national contests.
- Runs notebook and Python submissions in hardened Docker sandboxes; provides live leaderboards, double-blind jury review, custom evaluators, and distributed workers.
- <https://github.com/delyan-boychev/lavbench>

Vision-Language Models for Satellite Imagery Understanding | *Research at INSAIT* Dec. 2025 - Present *Advised by [Dr. Danda Paudel](#)*

- Adapting vision-language models to satellite imagery using scalable supervision derived from OpenStreetMap.
- With [Mario Markov](#), [Stefan Maria Ailuro](#) and [Mohammad Mahdi](#).

Spatial Modeling for Point-Cloud Understanding | *Research at INSAIT* June 2025 - Present *Advised by [Dr. Danda Paudel](#)*

- Connecting 3D point clouds with language representations for structured spatial reasoning and scene understanding.
- With [Runyi Yang](#).

Partial Solver | *DSA Course Project (C++)* Nov. 2025

- Implemented an expression solver that parses and evaluates arithmetic formulas with variables and dependency-aware execution.
- Built a tokenizer using polymorphic token classes (e.g., numbers/operators/identifiers) to convert input into structured objects.
- Evaluated infix expressions using the classic *two-stack* algorithm (operators/operands) for correct precedence and associativity.
- Handled inter-expression dependencies via a queue-based scheduler and a variable table, with careful memory management to avoid leaks/double-free.

- URL to the GitHub repository: <https://github.com/delyan-boychev/partial-solver>

Digital Audio Workstation | *OOP Course Project (C++)*

June 2025

- University project for an Object-Oriented Programming course: a simplified digital audio workstation implemented in C++.
- Designed an extensible OOP architecture (tracks/components, command interface, file I/O) with clean abstractions and modular code.
- Implemented core editing workflows such as importing/exporting audio, timeline/clip operations, and basic effects processing.
- URL to the GitHub repository: <https://github.com/delyan-boychev/dig-audio-workstation>

ImagiNet | *Synthetic Image Dataset*

2024

- A 144 GB, 200K-image dataset for synthetic-image detection, balanced across photos, paintings, faces, and uncategorized imagery.
- Supports both real-versus-synthetic classification and generator identification across open and proprietary image generators.
- <https://huggingface.co/datasets/delyanboychev/imagenet>

Gradient Cache Contrastive Learning | *PyTorch Package*

2023 - 2024

- A PyTorch implementation of gradient caching for scaling computer-vision contrastive learning beyond GPU memory limits.
- Reproduces large-batch gradients from smaller chunks; supports SimCLR, SupCon, SelfCon, mixed precision, and distributed training.
- <https://github.com/delyan-boychev/grad-cache-con-learning>

Synthetic Image Detection via Supervised Contrastive Learning | *ML Research* August 2023 - August 2024

- Generative models, such as Diffusion Models (DMs), Variational Auto-Encoders (VAEs), and Generative Adversarial Networks (GANs), produce images with a level of authenticity that makes them nearly indistinguishable from real photos and artwork.
- The work provides a new balanced training and benchmarking dataset containing diverse image content – photos, paintings, faces, and uncategorised images.
- Unlike previous datasets, it includes images from multiple generators – open-source and proprietary – during training.
- To establish a baseline, we train a ResNet-50 model using a self-supervised contrastive (SelfCon) objective for synthetic image detection and model identification.
- The model demonstrates strong performance across previously available benchmarks, achieving an AUC of up to 0.97 and balanced accuracy ranging between 80% and 93%, even under social network conditions involving compression and resizing.
- Shared research with [Radostin Cholakov](#), accepted for presentation at the Workshop on Datasets and Evaluators of AI Safety, AAAI 2025
- URL to the research paper: <https://arxiv.org/abs/2407.20020>

Increasing Computer Vision Models Interpretability | *ML Research*

August 2022 – June 2023

- The work aims to evaluate the effects of adversarial training (conducted via the Projected Gradient Descent algorithm) utilized to produce robust models - less vulnerable to adversarial attacks.
- The results present a correlation between the interpretability and adversarial robustness of computer vision models.
- Adversarially robust models' predictions can be explained through a straightforward procedure.
- URL to the research paper: <https://arxiv.org/abs/2307.02500>
- Presented at Regeneron ISEF'23.