

MARJAN STOIMCHEV

PhD in AI | Computer Vision & Deep Learning Engineer (R&D)

Ljubljana, Slovenia · Open to Remote or Relocation | +386-69-970-685 | marjanstoimcev@gmail.com
LinkedIn: linkedin.com/in/marjan-stoimchev-phd-636657179/ | GitHub: github.com/marjanstoimchev |
Portfolio: marjanstoimchev.github.io

SUMMARY

PhD Research Scientist and Computer Vision Engineer with 7+ years of experience developing robust, domain-adaptable computer vision systems across Geospatial AI, Earth Observation (EO), medical imaging, and industrial segmentation workflows. My work focuses on label-efficient learning through self-supervised, semi-supervised, and foundation-model adaptation methods, with particular emphasis on representation learning, hierarchical and multi-label prediction, and rigorous experimental evaluation. I have designed and validated novel vision-transformer-based models for land-cover mapping, biomedical image analysis, and low-data domain adaptation, while also translating research advances into reproducible, scalable, and deployment-ready ML pipelines. I bridge deep AI research with practical engineering, building systems that are scientifically grounded, production-oriented, and measurable in real-world impact.

CORE TECHNICAL SKILLS

Geospatial AI & Foundation Models: Vision Transformers (ViT), Masked Autoencoders (MAE), DINO/DINOv3-style pretraining, self-supervised learning, semi-supervised learning, and label-efficient representation learning.

Segmentation & Annotation Systems: YOLOv9, SAM-based pseudo-labeling, pseudo-mask refinement, boundary correction, synthetic image-mask generation, active learning, uncertainty-based annotation triage, CVAT, and annotation workflow automation.

Structured & Multi-Task Learning: Hierarchical and multi-label prediction, explicit label dependency modeling, graph neural networks (GNNs), weak/noisy supervision, and multi-task learning.

Software Engineering & ML Systems: Python; PyTorch & PyTorch Lightning; scalable training and inference pipelines; mixed-precision and multi-GPU training; experiment tracking; HPC/Slurm; Docker; reproducible ML workflows; ONNX/TensorRT-oriented deployment tooling.

Domains: Earth Observation, remote sensing, inventory image analysis, industrial annotation workflows, medical imaging, microscopy, and digital pathology.

CORE PROJECTS

Remote Sensing & Geospatial AI (ongoing)

- Slovenia Data Cube & AI Factory:** Leading development of the national Slovenia Data Cube to provide analysis-ready, high-performance geospatial data for the Slovenian AI Factory initiative.
- GeoAI Foundation Models:** Developing and training foundation models for remote sensing data based on multi-task and self-supervised paradigms to serve as backbones for semantic segmentation and object detection.
- Multi-Source Data Fusion:** Architecting pipelines for temporal and multi-modal fusion of complementary sources, including Sentinel-1 radar and Sentinel-2 optical imagery.

Biomedical Imaging

- Masked Prototype Prediction (MPP):** Designed a self-supervised representation learning framework using a DINO-style student-teacher architecture with MAE-style encoders.
- Zero-shot Segmentation:** Engineered zero-shot SAM-based nuclei detection pipelines for digital pathology to reduce expert-level manual annotation requirements.
- Embedding Analysis:** Developed DINOv2/DINOv3/MAE embedding pipelines for whole-slide analysis, enabling cell-type clustering and survival analysis with optimized multi-GPU training and mixed-precision inference.

Material Science

- Microscopy Drift Correction:** Designed a vision-based autoencoder system for automated drift correction in STEM microscopy, reducing spatial distortion for automated analysis.

PROFESSIONAL EXPERIENCE

Computer Vision / ML Engineering Consultant | Independent Consultancy

Apr 2026 - Present

Delivered production-oriented segmentation and annotation systems for inventory image analysis.

- End-to-End Segmentation Pipeline:** Built a segmentation-assisted annotation pipeline including model training, inference integration, evaluation scripts, reproducible configs, and technical documentation.
- DINOv3-to-YOLOv9 Distillation:** Implemented a distillation workflow to transfer task-specific representation knowledge from a DINOv3-style teacher into a lightweight YOLOv9 student for downstream proposal generation and annotation assistance.
- CVAT Workflow Deployment:** Deployed a self-hosted CVAT annotation environment and designed the workflow for importing, reviewing, correcting, and exporting autosegmentation outputs.
- Active Learning Triage:** Developed uncertainty-aware annotation prioritization using confidence-, entropy-, and disagreement-based scoring to identify images, objects, and regions requiring human review.
- Pseudo-Mask Refinement:** Implemented boundary-aware refinement and correction modules to reduce mask leakage, holes, jagged edges, and under-segmentation in automatically generated masks.

- **Synthetic Image-Mask Generation:** Built a synthetic data pipeline for segmentation training, supporting controlled variation in object placement, scale, occlusion, lighting, scene diversity, and class distribution.
- **Temporal Consistency:** Added cross-image consistency logic to stabilize masks across related frames, burst captures, and near-duplicate images.
- **Deployment Readiness:** Refactored, documented, and containerized the codebase to improve reproducibility, handoff quality, and deployment readiness.

Computer Vision Engineer | MedAI

Nov 2024 - Present

Solving bottlenecks in clinical phenotyping by replacing manual workflows with high-speed, self-supervised AI.

- **Accelerated Prediction Timelines:** Engineered embedding pipelines that compressed prediction wait times from weeks to minutes, enabling rapid iterations for biological research.
- **Solved the Labeled Data Bottleneck:** Reduced dependency on expensive annotated datasets by developing DINOv2/v3 and MAE self-supervised models for multiplex immunofluorescence.
- **Optimized Deployment:** Built production-ready pipelines for spatial clustering and neighborhood analysis, integrating them into clinical decision-support prototypes.
- **Bridge to Stakeholders:** Translated complex model outputs into explainable biological insights, improving cell-type identification accuracy and reducing manual annotation time by approximately 20%.

PhD Research Scientist | Jozef Stefan Institute

Oct 2021 - Dec 2025

Developed novel architectures from the ground up and established reproducible engineering standards.

- **Engineered Novel Architectures:** Designed and implemented custom graph-based ViT+GNN architectures to model complex taxonomic dependencies, surpassing standard implementations in structured prediction.
- **Built Reproducible Pipelines:** Established lab standards for reproducible research by implementing modular multi-GPU pipelines using PyTorch Lightning and Hydra, later adopted by intern teams.
- **Solved Data Efficiency in Remote Sensing:** Built semi-supervised SSL-MAE models that improved benchmark F1 scores by 12% while using 40% less labeled data, directly addressing the cost of data acquisition.
- **Research Leadership:** Led MSc interns in refactoring experimental research code into clean, executable pipelines across remote sensing and material science domains.

SELECTED PUBLICATIONS

Journal Articles

- **Stoimchev, M., Levatic, J., Koccev, D. and Dzeroski, S.:** "SSL-MAE: Adaptive Semisupervised Learning Framework for Multilabel Classification of Remote Sensing Images Using Masked Autoencoders," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 18, pp. 14882-14896, 2025, doi: 10.1109/JSTARS.2025.3578196
- **Stoimchev, M., Levatic, J., Koccev, D. and Dzeroski, S.:** "Semi-Supervised Multi-Label Classification of Land Use/Land Cover in Remote Sensing Images With Predictive Clustering Trees and Ensembles," IEEE Transactions on Geoscience and Remote Sensing, vol. 62, pp. 1-16, 2024, Art no. 4706416, doi: 10.1109/TGRS.2024.3426981.
- **Stoimchev, M., Ivanovska, M. and Shtruc, V.:** "Learning to Combine Local and Global Image Information for Contactless Palmprint Recognition," Sensors 2022, 22, 73. <https://doi.org/10.3390/s22010073>

Conference Papers

- **Stoimchev, M., Levatic, J., Koccev, D. and Dzeroski, S.:** "HELM: Hierarchical and Explicit Label Modeling." EurIPS Workshop on Earth Observation, 2025.
- **Stoimchev, M., Levatic, J., Koccev, D. and Dzeroski, S.:** "Hierarchical and Explicit Label Modeling with Graph Learning." AI & Space @ Discovery Science 2025.
- **Stoimchev, M., Levatic, J., Koccev, D. and Dzeroski, S.:** "Hierarchical Multi-Label Classification of Land Use/Land Cover Images." MACLEAN @ ECML-PKDD 2024.
- **Stoimchev, M., Levatic, J., Koccev, D. and Dzeroski, S.:** "Dimensionality Reduction for Semi-Supervised Learning from Remote Sensing Images." MACLEAN @ ECML-PKDD 2024.

EDUCATION

PhD, Information & Communication Technologies, Jozef Stefan Institute

Oct 2021 - Dec 2025

- **Focus:** Semi-supervised and hierarchical multi-label learning.
- **Impact:** Solved ground-up architectural problems for remote sensing and biomedical imaging applications.

MSc, Electrical Engineering, University of Ljubljana

Sep 2018 - Jan 2021

- **Honors:** Recipient of the Preseren Award for best independent research work during the master's program.

BSc, Electrical Engineering, Ss. Cyril & Methodius University, Skopje

Sep 2013 - May 2018

- **Coursework:** Control theory, robotics, signal processing, and AI systems.

ADDITIONAL INFORMATION

- **Languages:** English (C2), Macedonian (native), Slovenian (C2), Serbian (C2), German (A1), French (A1)
- **Hobbies:** Basketball, swimming, hiking, painting, diving, and cycling