South Caucasus Conference on Artificial Intelligence

SCCAI2025

Organized by the European Union Horizon Europe grant project GAIN

PROGRAM AND ABSTRACTS

Muskhelishvili Institute of Computational Mathematics
Georgian Technical University

Tbilisi, Georgia

16 - 18 September, 2025


Welcome to SCCAI2025

South Caucasus Conference on Artificial Intelligence (SCCAI2025) will be held in Tbilisi, Georgia from September 16 to September 18 2025. The conference is organized and fully financed by the European Union HORIZON EUROPE Project GAIN.

The conference aims to create an innovative, cross-disciplinary platform where academics, industry experts, and government officials unite to explore new frontiers in theory, methodology, systems, and applications of Artificial Intelligence (AI). This year, the focus is on the latest developments in Computer Vision and Natural Language Processing for Low-Resourced Languages.

Moreover, SCCAI2025 seeks to establish a collaborative research platform for the South Caucasus nations - Armenia, Azerbaijan, and Georgia, promoting their integration into the global AI community.

It is a pleasure to welcome you to the SCCAI2025 and to Tbilisi. We wish you a pleasant stay and that you enjoy the meetings.

Organizers

The organizers are the partner institutions within the GAIN project:

- Muskhelishvili Institute of Computational Mathematics (MICM) of the Georgian Technical University (GTU),
- German Research Center for Artificial Intelligence (DFKI)
- National Institute for Research in Digital Science and Technology in France (INRIA)

Program Committee

- Abzetdin Adamov (ADA University, AZERBAIJAN)
- Tanay Agrawal (INRIA, FRANCE)
- Jan Alexandersson (DFKI, GERMANY)
- Hrachya Astsatryan (IIAP NAS RA, ARMENIA)
- Michal Balazia (INRIA, FRANCE)
- François Brémond (INRIA, FRANCE)
- Eka Chkonia (TSMU, GEORGIA)
- Elena Eyngorn (EXO, GERMANY)
- Giorgi Giorgobiani (MICM, GEORGIA, Program Coordinator)
- Kartlos Kachiashvili (MICM/GTU, GEORGIA)

- Vakhtang Kvaratskhelia (MICM, GEORGIA)
- Tamar Lominadze (GTU, GEORGIA)
- Philipp Müller (DFKI, GERMANY)
- Mikheil Rukhaia (VIAM, Georgia)
- Samir Rustamov (ADA University, AZERBAIJAN)
- Benedikt Emanuel Wirth (DFKI, GERMANY)

Local Organizers - MICM-GAIN Team

Valeri Berikashvili, Giorgi Ghlonti, Tamar Giorgobiani, Tsotne Javakhishvili, Rapael Kalandadze, Ioseb Kachiashvili, Sofi Katamadze, Natia Kukhilava, Marina Menteshashvili, Luka Nadiradze, Zaza Sanikidze, Teimuraz Saghinadze, Zaza Tabagari, Luka Tabagari.

Conference Tracks and Chairs

- **KEYNOTE LECTURES** - Chairs: G. Giorgobiani (MICM), T. Lominadze (GTU), F. Brémond (INRIA), P. Müller (Max Planck Institute for Intelligent Systems), J. Alexandersson (DFKI), E. Eingorn (Exoulanch GmbH).
- **Track A - NLP FOR LOW-RESOURCED LANGUAGES** - Chairs: J. Alexandersson (DFKI), P. Müller (Max Planck Institute for Intelligent Systems), B. Mikaberidze (MICM).
- **Track B - COMPUTER VISION FOR PERCEPTION, INTERACTION & INTELLIGENT SYSTEMS** - Chairs: F. Brémond, M. Balazia (INRIA), T. Saghinadze (MICM).
- **DOCTORAL SYMPOSIUM** - Chairs: T. Agrawal (INRIA), I. Kachiashvili, G. Giorgobiani (MICM).

Doctoral Symposium Experts Panel

- François Brémond - INRIA, France
- Philipp Müller - DFKI, Max Planck Institute for Intelligent Systems, Germany
- Jan Alexandersson - DFKI, Germany
- Michal Balazia - INRIA, France
- Andrey Girenko - DFKI, Germany
- Tanay Agrawal - INRIA, France
- Kartlos Kachiashvili - MICM/GTU, Georgia.

Contents

Conference Schedule

16 September

09:30 - 11:00 Registration

11:00 Welcome

11:15 Key-note Lecture: François Brémond, Action Recognition for People Monitoring

12:00 Doctoral Symposium

12:00 Davit Bitmalkishev Design and Training of a Georgian Vision-to-Text Model Using ViT and ByT5.

12:30 Coffee Break

13:00 Levan Tsinadze, Lattice-Theoretic Formal Concepts for Sparse Autoencoder and Transcoder Feature Analysis in Large Language Models

13:30 Grigol Mikadze, AI-Augmented Privacy Architecture in Digital Access Control Systems: A Secure and Adaptive Design Approach

14:00 Irakli Koberidze, Auditing for Bias in Public Health Communications using Large Language Models as Explainable AI Tools

14:30 Lunch Break

15:30 Luka Nadiradze, Visual and Auditory Cue Modality Comparison in Spatial Short-Term Memory

16:00 Beso Mikaberidze, Parameter-Efficient Transfer Learning for Low-Resource NLP:Towards Advancing Georgian Language Processing

16:30 Irakly Parshutkini, Discrimination q-Rung Picture Linguistic Information TOPSIS Approach in Educational Programs Efficiency Evaluation

17:00 Ketevan Sikharulidze, Collaborative Filtering and TOPSIS Based Talent Scout Recommendation System

17:30 End of the sessions

17 September

10:00 Social Event - Excursion outside Tbilisi

- Uplistsikhe
- Ateni Sioni Church
- Dinner in Ateni at Nika Vacheishvili's Marani[1] and Guest House

[1] Wine cellar

18 September

10:30 Key-note Lecture: Philipp Müller, Multi-modal Behaviour Analysis for Effective Human-machine Collaboration

11:15 Key-note Lecture: Samir Rustamov, AI Listens, Learns, and Responds: Speech Technologies at the Forefront

12:00 Key-note Lecture: Hrachya Astsatryan, High-Performance Artificial Intelligence: Bridging Advanced Computing and AI

12:45 Coffee Break

13:15 Key-note Lecture: Abzatdin Adamov, The Limits of Intelligence in Modern AI: Engineering Consciousness and Awareness for Human-Like Cognition

14:00 Track A. Natural Language Processing for Low-resourced Languages

14:00 Archil Elizbarashvili, D. Gross-amblard, M. Tsintsadze, M. Ducassé, M. Khachidze, Parameter-Efficient Transfer Learning for Low-Resource NLP:Towards Advancing Georgian Language Processing.

14:20 Ali Parsayan, Mariam Razmadze, Next Word Prediction in the Georgian Language: Challenges and Neural Approaches

14:40 Rapael Kalandadze, RAFT for Georgian: A High-Quality Corpus for RAG, Fine-Tuning, and Benchmarking in a Low-Resource Language

15:00 Merium Hazem Anwar Labib Bishara, Lia Kurtanidze, Mikheil Rukhaia, Lali Tibua, Large Unranked Probabilistic Logic and its Applications

15:20 Track B. Computer Vision for Perception, Interaction & Intelligent Systems

15:20 Zaza Tabagari, Zaza Sanikidze, Giorgi Giorgobiani, The Use of Poisson's Equation and Image Processing Methods in Computer Vision Tasks

15:40 Lunch Break

16:40 Tereza Azatyan, Tigran Vandunts, Innovative Approaches to the Use of Computer Vision and Multimodal Systems in Assistive Technologies

17:00 Hakob Tamazyan, Ani Vanyan, Alvard Barseghyan, Anna Khosrovyan, Evan Shelhamer, Hrant Khachatrian, Large GeoCrossBench: Cross-Band Generalization for Remote Sensing

17:20 Ioseb Kachiashvili, Luka Tabagari, Human Detection from Body Movements Using Deep Learning Techniques

17:40 Merihan Hazem Anwar Labib Bishara, Anriette Michel Fouad Bishara, Neurosymbolic Methods in AI

18:00 Ivan Omelchenko, Yuri Gordienko, Vladyslav Taran, Sergey Peresada, Gert Jervan, Mairo Leier, Sergey Lyshevski, Sergii Stirenko, Modified Convolutional Architecture for Camouflaged Object Detection

18:20 Closure of the conference

Conference Abstracts

0.1 Key-note Lectures

Action Recognition for People Monitoring

François Brémond[1]

[1]*STARS - INRIA*
Sophia Antipolis,France
E-mail: francois.bremond@inria.fr

In this talk, we will explore how action recognition can be applied to human monitoring by analyzing video streams. Existing work has either focused on simple activities in real-life scenarios or on recognizing more complex activities (in terms of visual variability) in hand-clipped videos with well-defined temporal boundaries. However, there remains a gap in methods capable of retrieving multiple instances of complex human activity in a continuous (untrimmed) video flow in real-world settings. We will first present various techniques for detecting and tracking individuals in different environments. We will discuss various modalities, such as skeleton, optical flow, and emotion recognition, that can aid in the activity recognition process. We will then review state-of-the-art models for activity recognition and detection, including those using self-attention, transformers, and different pre-training methods. We will also cover specific cases of activity detection, such as video anomaly detection using weakly-supervised methods. Then, we will introduce several new techniques for recognizing Activities of Daily Living (ADLs) using 2D video cameras. The proposed activity monitoring approaches will be illustrated through several home-care application datasets, including Toyota SmartHome, NTU-RGB+D, Charades, and Northwestern UCLA. This comprehensive talk will provide a thorough understanding of the current state and future directions of action recognition for human monitoring in various real-world contexts.

Acknowledgments. The work was supported by the European Union Horizon Europe grant project GAIN.

Keywords: people tracking, behavior understanding, activity monitoring.

Multi-modal Behaviour Analysis for Effective Human-machine Collaboration

Philipp Müller[1,2]

[1]*Deutsches Forschungszentrum für Künstliche Intelligenz GmbH,Germany*
[2]*Max Planck Institute for Intelligent Systems*
E-mail: philipp.mueller@dfki.de

As collaboration between humans and machines becomes closer and spans an increasing number of scenarios, machines are confronted with the task of understanding and interacting with humans on a social level. At present, they still struggle to do so. For example, it is difficult for machines to infer human intention, interpret expressions of emotion, or even to judge whether humans are satisfied with the interaction. In my view, these challenges can only be addressed in an interdisciplinary way

and with multi-modal approaches. In my talk, I will highlight how my research bridges psychology and multi-modal AI in order to create machines that can effectively support humans, in particular in complex social scenarios.

Acknowledgments. This work was partially supported by the European Union under Horizon Europe project GAIN.

Keywords: Social Interaction, Multi-modal Machine Learning, Emotion, Attention.

AI Listens, Learns, and Responds: Speech Technologies at the Forefront

Samir Rustamov[1]

[1] *ADA University, MegaSec LLC*
Baku, Azerbaijan
E-mail: srustamov@ada.edu.az

Speech technologies stand at the intersection of human communication and artificial intelligence, enabling machines not only to recognize and process spoken language but also to respond with increasing intelligence and naturalness. This presentation explores the evolving landscape of speech technologiesŮfrom foundational tasks such as automatic speech recognition, speech synthesis, and speaker verification, to advanced applications in emotion recognition, prosody analysis, and real-time transcription. Drawing from real-world use cases in healthcare, education, customer service, and law enforcement, the talk highlights how AI is transforming voice into a powerful interface across industries. By examining the underlying models, system components, and development challenges, this session sheds light on the path toward building reliable, scalable, and human-centric speech solutions. As we move toward more seamless human-computer interaction, speech technologies are becoming not just tools, but essential components of intelligent, accessible, and responsive digital ecosystems.

Acknowledgments. This research was conducted at the Center for Data Analytics Research (CeDAR) of ADA University and the AI Laboratory of MegaSec Company.

Keywords: Speech Recognition, ASR, TTS, Speech Technologies.

High-Performance Artificial Intelligence: Bridging Advanced Computing and AI

Hrachya Astsatryan[1], **Yuri Shoukourian**[1], **Vladimir Sahakyan**[1], **Levon Aslanyan**[1],

[1] *Institute for Informatics and Automation Problems of the National Academy of Sciences of the Republic of Armenia, Yerevan, Armenia*
E-mail: hrach@sci.am, shouk@sci.am, svlad@sci.am, lasl@sci.am

As Artificial Intelligence (AI) and High-Performance Computing (HPC) rapidly converge, new possibilities emerge for addressing scientific, industrial, and societal challenges at scale. This talk introduces the mission and activities of our lab in the domain of High-Performance Artificial Intelligence (HiPerAI) - where scalable AI algorithms, advanced computing infrastructure, and applied

science meet. We present how AI and HPC technologies are co-designed to address key scalability, latency, heterogeneity, efficiency, energy use, and security challenges.

The presentation is structured on four pillars. Firstly, we describe the computational platform on which our work is based. Our lab uses flexible computing environments, essentially leveraging the strengths of the Armenian National Supercomputing Center.

Second, we highlight our focus on AI optimization of HPC. On this front, we design and deploy ML techniques to increase the performance and efficiency of HPC workloads.

A third central theme is the application of AI to scientific simulation. We present use cases in computational physics, chemistry, biology, and quantum systems, where AI assists in modeling biological membranes, scalable data processing services for Earth Observation data procesing, climate and environment modelling and forecast, or to process and analyze the exponentially increasing amount of astronomical data. AI is not used merely for post-analysis or acceleration, but as an integral component of the modeling and control loop, enhancing fidelity, adaptability, and predictive power.

Finally, we discuss future directions, including secure federated computing, the integration of AI across the edge-cloud-HPC continuum, hybrid quantum-classical AI models, and efforts toward reproducibility, benchmarking, and sustainability in scientific computing. These emerging areas reflect our commitment to building scalable, secure, intelligent computing platforms supporting advanced scientific discovery.

This presentation demonstrates how High-Performance Artificial Intelligence can act as a powerful enabler across domains, bridging the gap between data-driven intelligence and computational science.

Acknowledgments. The work was supported by the state base project funded by the Higher Education and Science Committee of the Ministry of Education, Science, Culture and Sports of the Republic of Armenia

Keywords: High performance computing, Artificial intelligence, Big Data.

References

[1] H. Astsatryan, V. Sahakyan, Yu. Shoukourian, J. Dongarra, P. Cros, M. Dayde, P. Oster, Strengthening compute and data intensive capacities of Armenia. In 2015 14th RoEduNet International Conference-Networking in Education and Research (RoEduNet NER) (pp. 28-33). IEEE, 2015.

The Limits of Intelligence in Modern AI: Engineering Consciousness and Awareness for Human-Like Cognition

Abzatdin Adamov[1]

[1]*School of Information Technologies and Engineering, ADA University*
Baku, Azerbaijan
E-mail: aadamov@ada.edu.az

Despite rapid advances in artificial intelligence, current AI systems remain fundamentally limited in their capacity to replicate human-like intelligence. While large-scale neural networks have achieved impressive results in language technologies, computer vision and some industry applications, they lack key attributes of human cognition Ů namely, consciousness, contextual awareness,

and intentionality. The speech will explore the theoretical and practical boundaries of intelligence in modern AI, critically examining the gap between artificial intelligence and genuine understanding. We investigate emerging approaches aimed at engineering computational models of consciousness and awareness, including architectures that simulate self-reflection, types of memory, and internal state representation. Through this view, we assess whether it is feasible to bridge the gap between task-specific generative intelligence based on statistical ML and general human-like cognition.

Keywords: Artificial Intelligence, Computational Consciousness, Engineering Consciousness and Awareness, Human-Like Cognition.

0.2 Track A. Natural Language Processing for Low-resourced Languages

Large Crowdsourcing as a Complement to Machine Learning for Validating Lexical Data on Georgian Verbs

Archil Elizbarashvili[1], **David Gross-amblard**[2], **Magda Tsintsadze**[1], **Mireille Ducassé**[2], **Manana Khachidze**[1]

[1] *Tbilisi State University, Georgia*

E-mail: archil.elizbarashvili@tsu.ge, magda.tsintsadze@tsu.ge, manana.khachidze@tsu.ge

[2] *UniversitÃl de Rennes, IRISA, France*

E-mail: david.gross-amblard@irisa.fr, mireille.ducasse@irisa.fr

Georgian verbs present significant challenges for learners and computational processing due to their complex and irregular inflectional morphology. Dictionaries often fail to adequately connect inflected forms with their lemmas, complicating morphological analysis and comprehension. To bridge this gap, we developed KARTUVERBS, an extensive online lexical resource based on Semantic Web standards, currently encompassing over five million inflected forms derived from more than 16,000 verbs. To further enrich and validate this dataset, machine learning techniques (specifically decision tree algorithms) were employed to infer missing verbal noun forms. Given variability in confidence scores, robust validation remains essential. Therefore, we integrated a specialized crowdsourcing platform - HEADWORK - to systematically gather and validate linguistic input from experts and informed users. Our second validation campaign aims to refine this dataset further by filtering outdated and invalid forms through detailed participant questionnaires. We anticipate improved accuracy, operational efficiency, and enhanced usability for learners and researchers.

Keywords: crowdsourcing, linked-data, machine learning, Georgian verbs, dictionary lemmas.

Next Word Prediction in the Georgian Language: Challenges and Neural Approaches

Ali Parsayan[1], **Mariam Razmadze**[1,2]

[1] *International Black See University, Tbilisi, Georgia*

E-mail: parsayan@ibsu.edu.ge

[2] *Georgian Technical University, Tbilisi, Georgia*

E-mail: razmadze.m@gtu.ge

Next word prediction is a key task in natural language processing (NLP), underpinning applications such as smart keyboards, text autocompletion, and conversational agents. While substantial progress has been achieved for high-resource languages like English and Chinese, low-resource, morphologically rich languages such as Georgian pose unique challenges. This research investigated next word prediction for Georgian, examining its linguistic complexities and assessing the performance of existing models. This paper proposes a transformer-based neural architecture fine-tuned on a Georgian corpus and compare the results with the prior work on Hidden Markov Model (HMM)-based

next word prediction in Azeri. Experimental results show that contextualized embeddings significantly outperform traditional n-gram models, emphasizing the need for syntactic and morphological awareness when modeling the Georgian language.

Keywords: Georgian Language, Next Word Predictor, Transformer-based Neural Architecture, Linguistic Complexities, Natural Language Processing (NLP).

RAFT for Georgian: A High-Quality Corpus for RAG, Fine-Tuning, and Benchmarking in a Low-Resource Language

Rapael Kalandadze[1,2]

[1] *Tbilisi AI Lab, Tbilisi, Georgia*

[2] *Muskhelishvili Institute of Computational Mathematics, Georgian Technical University, 380159, Tbilisi, Georgia*

E-mail: kalandadze.rapael@gtu.ge

Retrieval-augmented generation (RAG) has shown significant promise in improving factuality and grounding in large language models. However, applying these techniques to low-resource languages like Georgian remains challenging due to the lack of curated, high-quality data. In this project, we present a novel Georgian-language corpus tailored for RAG, built using a RAFT-inspired methodology. We sourced over 2,000 high-quality question-answer pairs from advanced Georgian-language computer science textbooks and academic literature, curated with the support of domain experts from local CS faculties. The dataset underwent multiple stages of filtering and quality assurance, including manual verification by trained human annotators to ensure factual consistency, relevance, and diversity. This corpus represents one of the first structured, RAG-friendly datasets for the Georgian language and is designed to facilitate grounded reasoning, improved retrieval, and better downstream performance of Georgian LLMs. It can be used both as a high-quality dataset for fine-tuning Georgian models and as a benchmark for evaluating retrieval and reasoning performance in Georgian. Our results demonstrate that targeted curation, even at smaller scales, can significantly improve outcomes in specialized domains and under-resourced languages.

Acknowledgments. The work was supported by the European Union Horizon Europe grant project GAIN.

Keywords: Georgian language, Low resource, RAG, RAFT, Synthetic data.

Unranked Probabilistic Logic and its Applications

Merium Hazem Anwar Labib Bishara[1], **Lia Kurtanidze**[2], **Mikheil Rukhaia**[3] **and Lali Tibua**[4]

[1] *International Black Sea University, Tbilisi, Georgia*

[2] *Georgian National University SEU, Tbilisi, Georgia*

[3] *I.Vekua Institute of Applied Mathematics of Tbilisi State University, Tbilisi, Georgia*

E-mail: 22100636@ibsu.edu.ge lia.kurtanidze@gmail.com mrukhaia@gmail.com ltibua@gmail.com

We present an unranked probabilistic logic, which is an extension of a first-order probabilistic logic LFOP_1 [4, 5, 6] with sequence variables and flexible-arity (unranked) function and predicate symbols [1, 2, 3]. The semantics of the logic is defined using Kripke worlds and the strong completeness theorem is proved for it. Such a formalism is interesting as it provides very flexible and expressive platform to model various problems coming from real world applications. We show one of the use cases in natural language processing, namely in ambiguity understanding.

Acknowledgments. This work was supported by Shota Rustaveli National Science Foundation of Georgia under the grant FR-22-4254.

Keywords: Unranked symbols, sequence variables, probabilistic primitives.

References

[1] Besik Dundua, Lia Kurtanidze, and Mikheil Rukhaia. Unranked tableaux calculus for web related applications. In 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON), pages 1181-1184. IEEE, 2017.

[2] Lia Kurtanidze and Mikheil Rukhaia. Skolemization in unranked logics. Bulletin of TICMI, 22(1): 3-10, 2018.

[3] Temur Kutsia and Bruno Buchberger. Predicate logic with sequence variables and sequence function symbols. In International Conference on Mathematical Knowledge Management, pages 205-219. Springer, 2004.

[4] Zoran Ognjanović and Angelina Ilić-Stepić. Logics with probability operators. In Probabilistic Extensions of Various Logical Systems, pages 1-35. Springer, 2020.

[5] Zoran Ognjanović and Miodrag Raśković. Some first-order probability logics. Theoretical Computer Science, 247(1-2): 191-212, 2000.

[6] Zoran Ognjanović, Miodrag Raśković, and Zoran Marković. Probability logics: probability-based formalization of uncertain reasoning. Springer, 2016.

0.3 Track B. Computer Vision for Perception, Interaction & Intelligent Systems

The Use of Poisson's Equation and Image Processing Methods in Computer Vision Tasks

Zaza Tabagari[1], **Zaza Sanikidze**[1], **Giorgi Giorgobiani**[1]

[1] *Muskhelishvili Institute of Computational Mathematics, Georgian Technical University, Tbilisi, Georgia*

E-mail: z.tabagari@gtu.ge; z.sanikidze@gtu.ge; giorgobiani.g@gtu.ge

The reliable functioning of software packages related to computer vision systems is of crucial importance in medical diagnostics, in the management of transport flows and other technical means, as well as in all those domestic spheres where computer vision algorithmic models must operate flawlessly under any conditions, including unforeseen circumstances. A prerequisite for this is the selection of stable analytical-computational schemes for processing relevant images. Among these, combined methods based on the use of mathematical apparatus deserve special attention due to their reliability and efficiency, especially when processing images with more or less complex configurations.

This presented report, along with modern artificial intelligence technologies, discusses the possibility of using certain mathematical methods for the informational analysis, reconstruction, or quality improvement of images, which will subsequently be used in various computer vision tasks.

It is shown that in certain cases, including the reconstruction of fragmentarily given images, the use of gradient analysis and the Laplacian, instead of (or in conjunction with) pixel analysis, offers significant advantages. This implies studying and numerically solving the corresponding Poisson's equation. Specific problems and examples are provided for illustration.

Acknowledgments. The work was supported by the European Union Horizon Europe grant project GAIN.

Keywords: Computer Vision, Image Processing, Poisson's Equation, Gradient Analysis.

Innovative Approaches to the Use of Computer Vision and Multimodal Systems in Assistive Technologies: Educational and Research Potential of a Regional University

Tereza Azatyan[1], **Tigran Vandunts**[2]

[1] *Armenian State Pedagogical University, Yerevan, Armenia*

E-mail: azatyantereza41@aspu.am

[2] *Goris State University, Goris, Armenia, Armenian State Pedagogical University, Yerevan, Armenia*

E-mail: tvandunts@gmail.com

Modern digital society demands inclusive technologies that ensure accessibility for people with disabilities. Computer vision (CV) and multimodal systems are at the core of this transformation, enabling adaptive human-computer interfaces and assistive tools.

This paper explores how CV and multimodal approaches can be implemented within the educational and research environment of a regional universityŮGoris State UniversityŮto foster inclusive technologies and interdisciplinary innovation.

Technical foundations of CV systems (OpenCV, MediaPipe, TensorFlow, PyTorch) and multimodal fusion methods (early, intermediate, late) are examined. CV enables sign language recognition, contactless interaction, and gesture-based control interfaces. Multimodal systems integrate visual, auditory, kinetic, and physiological data to enhance recognition accuracy.

These technologies enable inclusive educational platforms, student engagement tracking, and hybrid learning tools. Regional universities can develop interdisciplinary programs combining pedagogy, psychology, IT, and AI. Goris State University demonstrates the potential of regional institutions to become hubs for assistive technology development through strategic curriculum design, lab infrastructure, and international collaboration.

Acknowledgments. The work was supported by the internal R&D programs of Goris State University and Armenian State Pedagogical University.

Keywords: computer vision, multimodal systems, assistive technologies, inclusive education, regional university, interdisciplinary research.

GeoCrossBench: Cross-Band Generalization for Remote Sensing

Hakob Tamazyan[1,2] **Ani Vanyan**[1,2], **Alvard Barseghyan**[1,2] **Anna Khosrovyan**[1,2], **Evan Shelhamer** [3,4], **Hrant Khachatrian**[1,2]

[1] *YerevaNN, Yerevan, Armenia*

E-mail: {hakob,ani,alla,anna,hrant}@yerevann.com

[2] *Yerevan State University, Yerevan, Armenia*

[3] *Department of Computer Science, University of British Columbia, Vancouver, BC, Canada*

E-mail: shelhamer@cs.ubc.ca

[4] *Vector Institute, Toronto, ON, Canada*

Foundation models are rapidly advancing the field of Earth observation, but a critical challenge remains: their ability to generalize across different spectral bands and sensor types. As the volume and variety of remote sensing (RS) data grow, this cross-band generalization becomes essential for building versatile and practical applications, especially given the cost of retraining large models for every new sensor or data type.

This paper introduces *GeoCrossBench*, a novel benchmark designed to rigorously evaluate this critical cross-band generalization capability in RS foundation models. GeoCrossBench extends the standard GeoBench benchmark [2] by augmenting its datasets with co-registered Sentinel-1 Synthetic Aperture Radar (SAR) data alongside Sentinel-2 optical data. The benchmark covers three canonical RS tasks: scene classification, semantic segmentation, and change detection. The core evaluation protocol involves training models exclusively on common RGB bands and then testing their performance on a variety of unseen 3-channel combinations from both optical (e.g., RGE1, RE1E2, N'S1S2) and radar (VV/VH) modalities.

To establish a strong baseline and push the state-of-the-art, we also develop χViT (ChiViT), a self-supervised extension of the ChannelViT model [1]. This model is pretrained using an iBOT-style self-distillation framework [4] on a large-scale, multi-modal dataset specifically curated for remote sensing. The architecture is designed to handle varying channel inputs, making it well-suited for the cross-band challenge.

The experimental results on GeoCrossBench yield several key insights. First, they reveal that current foundation models, including those specifically designed for remote sensing, exhibit limited

generalization to unseen spectral bands. Performance drops significantly when models trained on RGB are tested on other optical bands or SAR data. Second, in a surprising finding, general-purpose vision models like DINOv2 [3], which were not pretrained on RS data, often outperform specialized RS models in cross-band generalization tasks. While the proposed χViT model demonstrates competitive performance compared to other RS-specific models, the success of general-purpose models suggests that current RS pretraining strategies may not be effectively learning the fundamental relationships between different spectral bands.

In conclusion, this work highlights a gap in the capabilities of current remote sensing foundation models. The introduction of GeoCrossBench provides the community with a standardized tool to measure and drive progress in cross-band generalization. The findings suggest that future research should focus on developing novel pretraining objectives and architectures that explicitly encourage learning robust, transferable features across the spectral bands, leading to more powerful models for comprehensive Earth observation.

Acknowledgments. The research was supported by the Higher Education and Science Committee of MESCS RA (Research project No 24RL-1B049). This work was supported by the Strategic Armenian Science and Technology Investment Community (SASTIC).

Keywords: Remote Sensing, Foundation Model, Vision Transformer, Generalization, Benchmark.

References

[1] Bao, Y., Sivanandan, S., and Karaletsos, T. Channel vision transformers: An image is worth $1 \times 16 \times 16$ words. (2024).

[2] Lacoste, A., Lehmann, N., Rodriguez, P., Sherwin, E., Kerner, H., LÃijtjens, B., Irvin, J., Dao, D., Alemohammad, H., Drouin, A., et al. Geo-bench: Toward foundation models for earth monitoring. *Advances in Neural Information Processing Systems*, 36, 51080–51093 (2023).

[3] Oquab, M., Darcet, T., Moutakanni, T., Vo, H. V., Szafraniec, M., Khalidov, V., Fernandez, P., Haziza, D., Massa, F., El-Nouby, A., et al. Dinov2: Learning robust visual features without supervision. *arXiv preprint arXiv:2304.07193* (2023).

[4] Zhou, J., Wei, C., Wang, H., Shen, W., Xie, C., Yuille, A. L., and Kong, T. Image bert pre-training with online tokenizer. In *International Conference on Learning Representations (ICLR)* (2022).

Human Detection from Body Movements Using Deep Learning Techniques

Ioseb Kachiashvili[1,2] **Luka Tabagari**[3]

[1] *Faculty of Informatics and Control Systems, Georgian Technical University, Tbilisi, Georgia*

[2] *Department of Informatics of the Muskhelishvili Institute of Computational Mathematics, Georgian Technical University, 380159, Tbilisi, Georgia*

E-mail: s.kachiashvili@gtu.ge

[3] *International Black See University, 380159, Tbilisi, Georgia*

E-mail: luka.tabagari@hotmail.com

The problem of human detection from body movements using deep learning techniques (on MPIIEmo Dataset) is considered. For this purpose, we used ByteTrack techniques that provided robust multi-object tracking and ensured the accurate generation of bounding box sequences. For feature extraction MMAction2 with VideoMae v2 method was applied that gave an efficient and scalable solution of the problem. The findings demonstrating the strengths and limitations of examined approaches and highlights, the potential for combining pose estimation, tracking, and efficient feature extraction techniques to advance human detection tasks are offered.

Acknowledgments. The work was supported by the European Union Horizon Europe grant project GAIN.

Keywords: Human Detection, Body Movement, Deep Learning, Convolutional Neural Network, Feature Extraction.

Neurosymbolic Methods in AI

Merihan Hazem Anwar Labib Bishara[1], **Anriette Michel Fouad Bishara**[2]

[1] *International Black Sea University, Tbilisi, Georgia*

[2] *I.Vekua Institute of Applied Mathematics of Tbilisi State University, Tbilisi, Georgia*

E-mail: 22100635@ibsu.edu.ge anriettehazem@gmail.com

Inductive Logic Programming (ILP) is a formalism that operates at the intersection of Machine Learning and Logic Programming [1], where the space of hypotheses is formed by logic programs, typically expressed in a subset of first-order predicate logic [2]. This unique characteristic allows ILP to learn complex relational patterns and hierarchical structures directly from data. This makes ILP a powerful tool for Artificial Intelligence (AI) [4], namely for tasks from natural language processing (NLP) and computer vision (CV).

In this paper, we survey different frameworks that integrates neurosymbolic reasoning [3, 5] and probabilistic ILP. Show some examples, how these frameworks address key AI challenges.

Keywords: probabilistic inductive logic programming, neurosymbolic artificial intelligence.

References

[1] Andrew Cropper and Rolf Morel. Learning programs by learning from failures. Machine Learning, 110(4): 801-856, 2021.

[2] Angelika Kimmig, Guy Van den Broeck, and Luc De Raedt. Algebraic model counting. Journal of Applied Logic, 22: 46-62, 2017.

[3] Artur d'Avila Garcez, Marco Gori, Luis C Lamb, Luciano Serafini, Michael Spranger, and Son N Tran. Neural-symbolic computing: An effective methodology for principled integration of machine learning and reasoning. arXiv preprint arXiv: 1905.06088, 2019.

[4] Céline Hocquette, Andreas Niskanen, Matti Järvisalo, and Andrew Cropper. Learning mdl logic programs from noisy data. In Proceedings of the AAAI Conference on Artificial Intelligence, volume 38, pages 10553-10561, 2024.

[5] Fieke HillerstrÂĺom and Gertjan Burghouts. Towards probabilistic inductive logic programming with neurosymbolic inference and relaxation. Theory and Practice of Logic Programming, 24(4): 628-643, 2024.

Modified Convolutional Architecture for Camouflaged Object Detection

Ivan Omelchenko[1], **Yuri Gordienko**[1], **Vladyslav Taran**[1], **Sergey Peresada**[1], **Gert Jervan**[2], **Mairo Leier**[2], **Sergey Lyshevski**[3], **Sergii Stirenko**[1]

[1] *National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine*
E-mail: *i.omelchenko.ip04@kpi.ua, [gord,stirenko]@comsys.kpi.ua

[2] *Tallinn University of Technology, Tallinn, Estonia*
E-mail: [gert.jervan,mairo.leier]@taltech.ee

[3] *Rochester Institute of Technology, Rochester, New York, USA*
E-mail: sergey.lyshevski@mail.rit.edu

Camouflaged object detection (COD) in natural environments is a challenging and understudied task in the field of computer vision. COD represents a critical frontier in applying artificial intelligence (AI) to environmental monitoring, infrastructure resilience, and conservation. This study presents an overview of modern approaches to COD, including convolutional neural networks (CNNs), transformer-based models, gradient-based methods, and graph-based frameworks and provides quantitative results of various frameworks that belong to these methods. The research outlines the difficulties associated with COD, such as variability in camouflage strategies, scarcity of well-annotated datasets, and computational complexity of existing models. Several benchmark datasets (COD10K, CAMO, NC4K, Military Personnel Data, and MoCA) are examined, each contributing unique properties to model development and evaluation. A modified framework Hybrid-SINet is proposed and evaluated using standard COD metrics, demonstrating improved performance (Hybrid-SINet benchmarking on Kaggle) over the baseline SINet model (Baseline SINet benchmarking on Kaggle) under specific environmental conditions.

Acknowledgments. This research was supported by the Office of Naval Research (ONR) and US National Academy of Science (US NAS) (IMPRESS-U initiative, No. STCU-7125) in the part of exploratory research on the new robust machine learning approaches for object detection and classification; and by the National Research Fund of Ukraine under grant 2025.06/0100 in the part of the development of neural network models for object detection, and by the Ministry of Education and Sciences of Ukraine (MESU), grant 2715r, as part of the search for advanced deep learning methods based on the requirements posed by devices on Edge Computing and Edge Intelligence layers. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the ONR, US NAS, and NRFU.

Keywords: Deep Learning, Object Detection, Camouflaged Object Detection, Convolutional Neural Network, Mean Absolute Error, Infrastructure Resilience Monitoring.

0.4 Doctoral Symposium

Design and Training of a Georgian Vision-to-Text Model Using ViT and ByT5

Davit Bitmalkishev[1]

[1] *University of Georgia,*
Tbilisi, Georgia
E-mail: davit.bitmalkishev@ug.edu.ge

Supervisor: Professor Sergo Tsiramua (University of Georgia)

Creating an artificial neural network capable of generating textual and audio descriptions of graphical images in multiple languages holds significant value from technological, social, economic, and humanitarian perspectives. This paper discusses Bootstrapped Language-Image Pre-training (BLIP), which is designed for simultaneous processing of text and images within a unified framework. Its main limitation is its focus on the English language, which poses a challenge for non-English language tasks. In the present paper is explored several approaches to overcoming this limitation: (1) integrating a translator at the output stage; (2) replacing the encoder/decoder with implementations that support the target language; (3) retraining the existing model using multilingual data and create new one from scratch. Each method comes with specific technical challenges, the analysis of which and corresponding recommendations are presented in the conclusion of the paper. To address this issue, I developed the model with name "Martha". This model is based on the ByT5 language model, which is a variant of the T5 (Text - To - Text Transfer Transformer) model. To extract image features, the model uses ViT-B-16, which is already trained for image feature recognition. Over 200,000 "image/description" pairs for training are already prepared in Georgian, so that the language model can learn how to handle the tensors output by the ViT model, that is, to correctly transform them into verbal interpretations and learn the logical process.

Keywords: Artificial Intelligence, Transformer-based Models, Bootstrapped Language-Image Pretraining.

Lattice-Theoretic Formal Concepts for Sparse Autoencoder and Transcoder Feature Analysis in Large Language Models

Levan Tsinadze[1]

[1] *Department of Computer Science, Faculty of Exact and Natural Sciences, I. Javakhishvili Tbilisi State University,*
Tbilisi, Georgia
E-mail: levan.tsinadze784@ens.tsu.edu.ge

Supervisor: Professor Alexander Gamkrelidze (I. Javakhishvili Tbilisi State University)

In this paper, we apply a specialized lattice-theoretic formal concept construction to tokens and sparse activation pairs derived from residual stream Sparse autoencoders and MLP transcoders in Large Language Models, observing the behaviour of individual token embeddings within sequences and entire sentence representations using lattice operations. We investigate the evolution of each token embedding, the propagation of contextual information from earlier to later layers, and the assembly of sentences from their constituent token embeddings. We also examine the mutual features

that represent how contextual information accumulates in later-layer embeddings. Our framework combines formal concept analysis with mechanistic interpretability, offering theoretical foundations and empirical validation that uncover structural insights into the embedding circuits of LLMs.

Keywords: lattice-theoretic formal concepts, Galois connections,large language models, MLP transcoders, sparse autoencoders, mechanistic interpretability.

AI-Augmented Privacy Architecture in Digital Access Control Systems: A Secure and Adaptive Design Approach

Grigol Mikadze [1]

[1] *I. Javakhishvili Tbilisi State University, Tbilisi, Georgia*

E-mail: grigol.mikadze976@ens.tsu.edu.ge

Supervisor: Professor Manana Khachidze (I. Javakhishvili Tbilisi State University)

Digital access control systems are essential for securing identity-based access in both physical and digital environments. However, their reliance on sensitive credentials and behavioral data creates serious privacy risks if not properly managed. Traditional privacy-by-design approaches provide static safeguards but often lack real-time adaptability to emerging threats. This paper proposes an AI-augmented architectural model that enhances the privacy posture of access control systems. It explores how artificial intelligence can support dynamic policy enforcement, anomaly detection, and intelligent data minimization. By embedding AI into logging, credential governance, and risk evaluation processes, the system becomes more responsive to compliance requirements and contextual privacy risks. The study outlines core architectural principles, integration paths for privacy-enhancing technologies (PETs), and metrics for evaluating effectiveness. The proposed model supports scalable, policy-aware privacy protection while maintaining operational efficiency.

Keywords: Access Control Systems, Data Privacy, Artificial Intelligence, Privacy-by-Design, Secure Architecture, Adaptive Access Control, Privacy-Enhancing Technologies, Anomaly Detection.

Auditing for Bias in Public Health Communications using Large Language Models as Explainable AI Tools

Irakli Koberidze[1]

[1] *Dept. of Computer Science, I. Javakhishvili Tbilisi State University, Tbilisi, Georgia*

E-mail: : irakli.koberidze@tsu.ge

Supervisor: Professor Magda Tsintsadze (I. Javakhishvili Tbilisi State University)

Public health communications are vital for informing the public and building trust, yet they often contain biases that worsen health disparities and erode confidence in health institutions. This research introduces a novel application of Large Language Models (LLMs) as Explainable AI (XAI) auditors. Instead of explaining an AI's own decisions, our framework leverages LLMs to identify and explain demographic biases (e.g., racial, gender, age-related) in human-generated public

health texts. The methodology involves systematic corpus curation, iterative prompt design, and rigorous evaluation of the LLM's identified biases, proposed mitigations, and its inherent biases. Preliminary work suggests the feasibility of this approach, demonstrating the LLM's ability to understand bias concepts and offer initial insights. This work aims to develop a scalable, AI-driven tool for more equitable and trustworthy public health messaging, contributing significantly to health equity, ethical AI development, and public trust.

Keywords: Large Language Models (LLMs), Bias detection, Public health communications, Explainable AI (XAI), Health equity .

Visual and Auditory Cue Modality Comparison in Spatial Short-Term Memory

Luka Nadiradze[1]

[1] *Faculty of Natural Sciences and Medicine, Ilia State University, Tbilisi, Georgia*

E-mail: : luka.nadiradze@iliauni.edu.ge

Supervisors: Dr. Philipp Müller (DFKI/Max Planck Institute for Intelligent Systems), Dr. Benedikt Emanuel Wirth (DFKI), Dr. Timea Folyi (USAAR)

This study tests whether the often-reported visual advantage in spatial short-term memory (STM) reflects genuinely modality-specific memory mechanisms or simply the superior perceptual fidelity of typical visual cues. Twenty-two adults performed a change-detection task in which two four-item sequences of spatial locations had to be judged as same or different. Sequences were delivered either unimodally (visualŰvisual, auditoryŰauditory) or crossmodally (visualŰauditory, auditoryŰvisual) under two difficulty configurations: easy-visuală/ăhard-auditory and easy-auditoryă/ăhard-visual. A 2×4 repeated-measures ANOVA on sensitivity (d') revealed a significant Modalityă×ăDifficulty interaction ($F = 23.69, p < .001$). When visual cues were noise-free, visualŰvisual trials outperformed auditoryŰauditory trials; this pattern reversed when auditory cues were clearer. Crossmodal accuracy was comparable to, or worse than, the weaker unimodal baseline, indicating that inter-modality translation carries a measurable cost. These findings favour an information-reliability account over fixed modality appropriateness, and underscore the need to equate perceptual demands before attributing cognitive advantages to any sensory modality.

Acknowledgments. The work was supported by the European Union Horizon Europe grant project GAIN.

Keywords: spatial short-term memory; crossmodal perception; modality appropriateness; information reliability; cue fidelity; change-detection.

Parameter-Efficient Transfer Learning for Low-Resource NLP:Towards Advancing Georgian Language Processing

Beso Mikaberidze[1]

[1] *Muskhelishvili Institute of Computational Mathematics, Georgian Technical University, Tbilisi, Georgia*

E-mail: beso.mikaberidze@gmail.com

Supervisor: Dr. Philipp Müller (DFKI/Max Planck Institute for Intelligent Systems)

This PhD research explores the design of scalable and parameter-efficient methods for advancing low-resource language processing, with a special focus on the Georgian language. Motivated by the scarcity of training resources and morphological complexity of Georgian, this work aims to leverage cross-lingual transfer through prompt-based and tokenization-aware strategies. We present progress in two areas: (i) a comprehensive evaluation of tokenization methods for contextual embedding quality in Georgian, and (ii) a new prompt-encoder approach for zero-shot multilingual transfer across truly low-resource languages. These components jointly contribute to the long-term goal of building robust and efficient NLP pipelines for Georgian.

Acknowledgments. The work was supported by the European Union Horizon Europe grant project GAIN.

Keywords: NLP, Tokenization, PEFT, Soft Prompt, PETL, Cross-lingual Transfer Learning.

Discrimination q-Rung Picture Linguistic Information TOPSIS Approach in Educational Programs Efficiency Evaluation

Irakly Parshutkini[1]

[1]*Dept. of Computer Science, I. Javakhishvili Tbilisi State University, Tbilisi, Georgia*

E-mail: irakly.parshutkin@tsu.ge

Supervisor: Professor Gia Sirbiladze (I. Javakhishvili Tbilisi State University)

In certain MAGDM models an aggregation of matrices of experts' evaluations/ratings of possible alternatives by their attributes into an etalon decision making matrix is an important task. Based on possible great experience and deep knowledge of the subject, the expert tries to dominate other experts in the decision-making process. Besides, every expert has a possibility of a certain degree to influence on the decision-making process. In such cases experts' evaluations aggregation by the additive or linear aggregation instruments is unacceptable. We have found one approach to solve this task. Connection between experts' pair interaction indexes and possibility levels of their influence on decision making process is constructed. Experts' evaluations are represented by the q-rung orthopair fuzzy numbers. Before aggregation of experts' data into decision making matrix, evaluations are transformed into discrimination q-rung picture linguistic numbers. q-rung picture linguistic numbers contain as well quantitative as qualificative information on experts' reflections on objects. Based on constructed possibility distribution on the experts' group, the ordered weighted averaging (OWA) and Choquet averaging (CA) aggregation operators q-RPLOWA and q-RPLCA under q-rung picture linguistic environment are extended. By the q-RPLOWA and q-RPLCA operators' matrices of experts' evaluations/ratings are condensed into an etalon matrix. Therefore, technique for the TOPSIS approach for the ranking of possible alternatives by aggregating of discrimination q-rung picture linguistic data of etalon matrix is developed. For illustration of the obtained results, Educational Programs Efficiency Evaluation problem is considered. Our constructed MAGDM model complies with the principle of expertise in educational programs evaluation commission.

Keywords: Interactive MAGDM, TOPSIS, associated probabilities, Choquet integral aggregation, fuzzy discrimination measures, q-rung picture linguistic sets.

Collaborative Filtering and TOPSIS Based Talent Scout Recommendation System

Ketevan Sikharulidze[1]

[1]*Dept. of Computer Science, I. Javakhishvili Tbilisi State University, Tbilisi, Georgia*

E-mail: irakly.parshutkin@tsu.ge

Supervisor: Professor Gia Sirbiladze (I. Javakhishvili Tbilisi State University)

The current, personnel selection processes pose a significant challenge for modern companies, especially in the technology sector, where finding high-level specialists is often a time-consuming and inefficient process. This article describes an Android-based application called BR Scout, which uses a two-stage recommendation system: collaborative filtering based on technical, soft skills, and languages, and the TOPSIS algorithm for multi-criteria evaluation. The paper presents the application's architecture, data models, and specific examples of algorithm usage. The aim of the research is to demonstrate how the combination of recommendation algorithms can facilitate the optimal selection of candidates for high-level positions. The article is based on practical implementation and offers an evaluation of the results considering the functional aspects of the application, thereby highlighting BR Scout's potential in the field of HR technologies.

Keywords: Collaborative Filtering, Recommender Systems, TOPSIS Method.

Index