GAIN 2nd Summer School on AI Modeling of Human Cognition and Emotion

PROGRAMME AND
PRACTICAL INFORMATION

12 - 14 August 2024
MICM, Tbilisi, Georgia

Welcome!

The strategic goal of the GAIN project is to integrate Georgia, a Widening country, into the European research community focusing on Artificial Intelligence (AI), a technology of today and tomorrow. This will be achieved by strengthening the research and innovation capacities of Muskhelishvili Institute of Computational Mathematics (MICM). The project is supported by two leading European research organizations - the German Research Center for Artificial Intelligence (DFKI, Germany) and the National Institute for Research in Digital Science and Technology (INRIA, France), along with the high-tech company EXOLAUNCH GmbH.

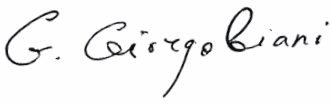
GAIN Summer Schools aim to enhance scientific excellence, build capacity, and train young researchers, improving their networking skills in an international setting.

The GAIN 1st Summer School took place at INRIA, Sophia Antipolis, France in 2023, alongside the 3rd INRIA-DFKI European Summer School on AI, IDESSAI 23.

The GAIN 2nd Summer School is hosted by MICM, Tbilisi, Georgia. The program focuses on AI Modelling of Human Cognition and Emotion. Participants will have the opportunity to attend courses and lectures on various topics of AI.

We extend a warm welcome to all participants attending the Summer School, we hope you enjoy it. Wishing you success.

On behalf of the GAIN consortium,
Coordinator of the project,
Deputy director of MICM
Giorgi Giorgobiani



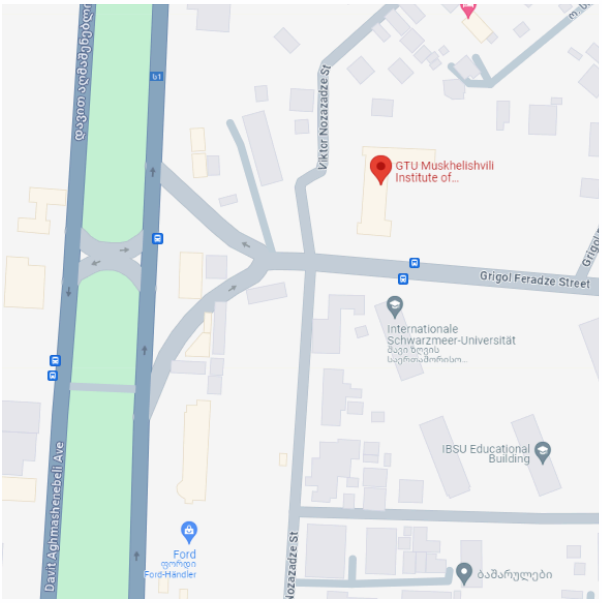
The Project GAIN is funded by the European Union HORIZON EUROPE - WIDERA-2021-ACCESS-03 (Twinning) programme under the grant agreement No. 101078950

Getting to MICM

GTU Muskhelishvili Institute of Computational Mathematics
4 Grigol Feradze Street
0159 Tbilisi, Georgia

micm.edu.ge

Click here to see it on Google Maps



By taxi

- Order via the following Android or IOS Apps:
  - Bolt **(the best choice most of the time)**
  - Yandex
  - Maxim **(you can also call +995 32 2 60 60 60)**
  - Uber Georgia
- Taxis in Tbilisi are typically affordable and convenient, making them a great option for traveling to distant spots like MICM from the city center.

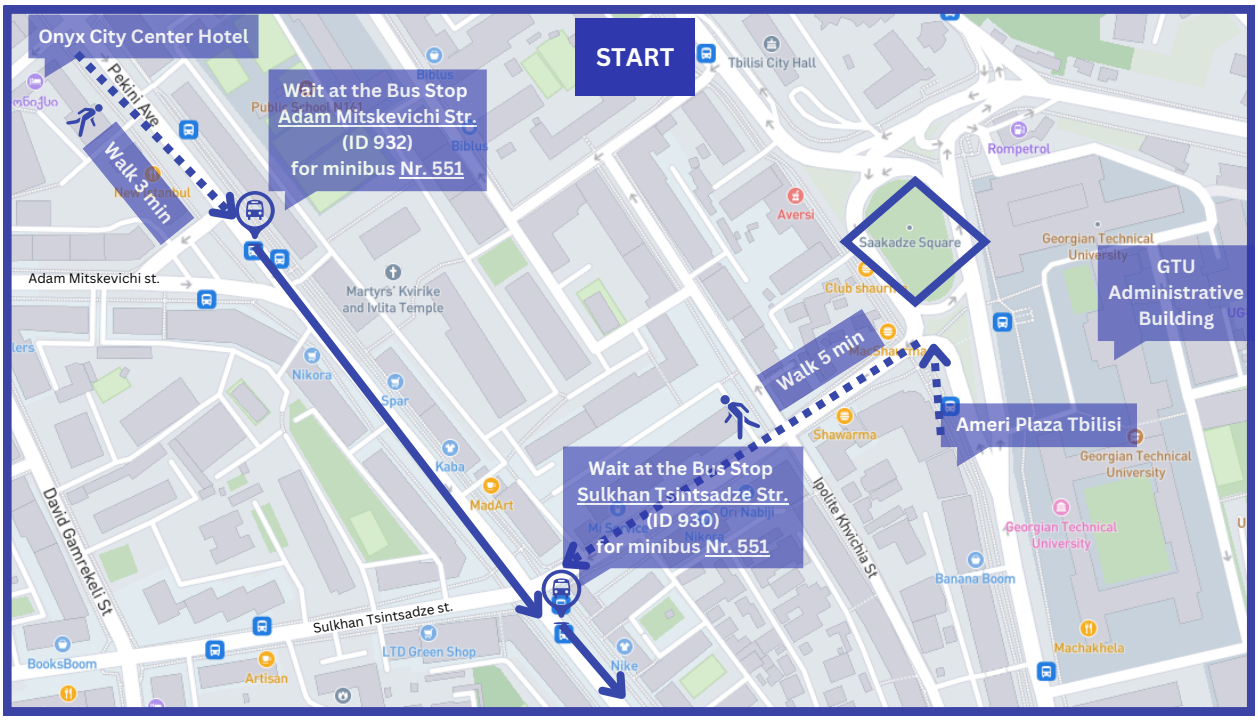
By bus or minibus

- You can use Tbilisi Transport Company App (TTC App) or Google Maps to plan your journey by bus.
- But be **careful**, bus arrival time in Google Maps is usually incorrect, use TTC App instead.
- The fare costs **1 GEL** (Georgian Lari, approx. 0.34 Euro).
- In public transport, **cash payments are not accepted**; instead, you can use your **bank card (e.g., VISA, Mastercard)** to purchase tickets directly on the bus by tapping your card on the ticket machine. Paper tickets are not issued, so during ticket checks, you present your bank card and tap it on the terminal provided by the staff. Additionally, **digital wallet apps** such as Google Pay and Apple Pay are accepted for payments.
- Alternatively, you can purchase a special transport card (Metromoney) at subway stations (2 GEL) or you can buy the ticket online in the TTC App (valid for 90 min.). To top up the balance of either Metromoney card or online ticket use special orange Express Pay machines from Bank of Georgia (BOG) located at nearly every bus stop.

Getting to MICM *(continued)*

By minibus from Pekini ave. (near Saakadze Square)

- There are number of possibilities to get from **Pekini ave.** (near **Saakadze Square, Ameri Plaza and Onyx Hotels**) to **MICM,** but to go there directly, without a transfer, you can take the **minibus Nr. 551** in the direction of Tbilisi Mall. See the diagram below. ⬇
- The **minibus 551** comes every 10 to 15 minutes.
- This trip takes approx. 45-50 minutes.
- As opposed to normal buses, minibus is smaller, dark blue vehicle.
- It also has some peculiarities: you have to wait for it at the bus stop, but need to **wave a hand** to ensure that the driver stops; when you reach your destination you have to **ask the driver to stop** (better tell them beforehand your desired bus stop).
- You can return to **Ameri Plaza Hotel** using the same route and minibus.



The Project GAIN is funded by the European Union HORIZON EUROPE - WIDERA-2021-ACCESS-03 (Twinning) programme under the grant agreement No. 101078950

Programme

12 Aug

11:00 - 11:30 Registration

11:30 - 12:00 Welcome

12:00 - 13:00 Keynote 1
Andrey Girenko, DFKI[1]. "European approach to AI: Policy, regulation, initiatives"

13:00 - 14:00 Lunch

14:00 - 18:00 Course 1
François Brémond, INRIA[2]. "Computer vision for action recognition"

13 Aug

10:00 - 11:00 Keynote 2
Minha Lee, TU/e[3]. "Exploring Moral Implications of Collaborative Mind Perception"

11:00 - 14:00 Course 2 - Part 1
Philipp Müller, DFKI. "AI-based multimodal interaction analysis" (lecture)

14:00 - 15:00 Lunch

15:00 - 15:30 Course 2 - Part 2
Philipp Müller, DFKI. Brainstorming groups

15:30 - 19:00 Presentations of the young GAIN researchers

14 Aug

10:00 - 11:00 Keynote 3
Elena Eyngorn, EXO[4]. "Soft Skills for Engineers" (Online webinar)

11:00 - 13:00 Course 3 - Part 1
Benedikt Emanuel Wirth, DFKI. "Face Recognition in humans and DCNNs"

13:00 - 14:00 Lunch

14:00 - 15:00 Course 3 - Part 2
Benedikt Emanuel Wirth, DFKI. Small Groups Exercises

15:40 - 18:30 Visit to Tbilisi Mental Health Centre *(restricted group)*
Psychiatric Pilot Project at MICM (based on MEPHESTO)

1. German Research Center for Artificial Intelligence, Germany
2. National Institute for Research in Digital Science and Technology, France
3. Eindhoven University of Technology, Netherlands
4. EXOLAUNCH GmbH

Session Abstracts

Keynote 1

European Approach to AI: Policy, Regulation, Initiatives

Andrey Girenko, DFKI

The European AI Strategy aims at making the EU a world-class hub for AI and ensuring that AI is human-centric and trustworthy. Such an objective translates into the European approach to excellence and trust through concrete rules and actions. The presentation will cover the EU AI strategies and Coordinated Plan on AI. A deeper insight will be given to the European AI Act - the world's first regulation concerning the use of AI. The presentation will be concluded with an overview of the R&I related activities on the EU level (ADRA activities in Horizon Europe and Digital Europe).

Course 1

Computer Vision for Action Recognition

François Brémond, INRIA

In this course, 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 tracking, optical flow, gaze detection, 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 course will provide a thorough understanding of the current state and future directions of action recognition for human monitoring in various real-world contexts.

Session Abstracts *(continued)*

Keynote 2

Exploring Moral Implications of Collaborative Mind Perception

Minha Lee, TU/e

We denote other beings to have minds of their own when we perceive them to have cognitive and affective capacities, to different degrees. A chatbot can be perceived to have some level of cognition, but not emotions, for instance. But, this bias can be overturned through how the agent and the environment it operates in are designed. Hence, when envisioning cooperative human-machine dyads and teams, one question can be asked: To what extent should our perception of other minds be calibrated collaboratively, and in what ways? The potential to shape our human perception of non-human "minds" has various ethical considerations to be discussed.

Course 2

Ai-Based Multimodal Interaction Analysis

Philipp Müller, DFKI

Social interactions permeate our lives, be it in the form of family gatherings, study groups, or psychotherapy sessions. If we want to build machines that are able to interact with us and support us in such interactions, these machines need the ability to sense and interpret human social behaviour. In this lecture, we will learn about the foundations of multi-modal social interaction analysis. We will discuss the most relevant modalities including speech, body language, and facial behaviour as well as how these modalities can be combined to infer aspects of human's internal states and interpersonal relations. Students will have the opportunity to develop their own research ideas in brainstorming groups and pitch them to the general audience.

Session Abstracts *(continued)*

Keynote 3

Workshop on Soft Skills for Engineers

Elena Eyngorn, EXO

In today's fast-paced and collaborative engineering environments, technical prowess alone is not sufficient for success. This workshop, "Soft Skills for Engineers," is designed to equip engineers with essential interpersonal skills that enhance teamwork, communication, leadership, and problem-solving capabilities. Participants will learn practical strategies for effective collaboration, clear and persuasive communication, conflict resolution, and emotional intelligence. Through interactive sessions and real-world scenarios, engineers will gain insights into how soft skills can complement their technical expertise, leading to more innovative solutions and successful project outcomes. Join us to bridge the gap between technical skills and interpersonal effectiveness, and take your engineering career to new heights.

Course 3

Face Recognition in Humans and DCNNs

Benedikt Emanuel Wirth, DFKI

AI-based face recognition is increasingly used in applied scenarios such as user identification, passport controls or tracing of wanted criminals on online platforms. In these use cases, errors can have dire consequences such as unauthorized individuals getting access to sensitive data or innocent people being persecuted due to a false match. While Deep Convolutional Neural Networks (DCNNs) recently achieved remarkable face recognition performance, little is known about how DCNNs recognize faces. It has recently been shown that DCNNs recognize objects in rather different ways than human observers. However, face recognition is a more complex task than object recognition because all faces share the same first order relations (i.e., all faces have the same basic spatial configuration of two eyes above the nose, the nose above the mouth, etc.) and similar textures (especially faces from the same ethnic group without beards). Thus, the human visual system relies on specialized structures and processes for face recognition. In the session, the following questions will be discussed: How do human observers recognize faces? Do DCNNs recognizes faces in a similar way as human observers? Are DCNNs good models and substitutes for human face recognition? Are face recognition DCNNs susceptible to specific errors and manipulations?

Session Abstracts

Presentations of the young GAIN researchers

Presentation 1

Evaluation in EEG Emotion Recognition: State of the Art Review and Unified Framework

Electroencephalography based Emotion Recognition (EEG-ER) has become a growing area of research in recent years. We present here a comprehensive literature review of EEG-ER articles published between 2018 and 2023. Especially, we focus on the evaluation protocol. Analyzing 231 papers, we show that there are inconsistencies between evaluation protocols, which are due to ground truth definition, evaluation metric selection, data splitting types (e.g., subject-dependent or subject-independent) and the use of different datasets. We uncover that the field lacks a unified protocol, which is essential to fairly compare new approaches and to track the field's progress. Capitalizing on this state-of-the-art research, we propose a unified evaluation protocol, EEGAIN, which enables an easy and efficient evaluation of new methods and datasets. EEGAIN is a novel open source software framework, offering the capability to compare, and thus define, state-of-the-art results. EEGAIN includes standardized methods for data pre-processing, data splitting, evaluation metrics, and the ability to load the 6 most relevant datasets (i.e., Mahnob-HCI, Deap, Dreamer, Amigos, SEED, SEED IV) in EEG-ER with only a single line of code. In addition, we have assessed and validated EEGAIN using the 6 datasets on the 4 most common publicly available methods (TSception, EEGNet, Deepconvnet, Shallowconvnet). This is a significant step to make research on EEG-ER more reproducible and comparable, thereby accelerating the overall progress of the field.

Presentation 2

Emotion Recognition from Human Body Movements

Recent research has reported substantial progress in machine-learning based emotion recognition from outward behavior. These approaches, however, usually rely on facial expressions or speech and are thus not applicable to material with low resolution, without audio, or filmed from large distance or challenging angles (e.g., CCTV footage). Therefore, we aim to develop an approach to emotion recognition from full body movements. To this end, we use the MPIIEmo dataset which contains videos of dyadic interactions with emotional content. The dataset consists of 224 video sequences filmed from eight viewpoints each. For feature extraction, we used ByteTrack which successfully created bounding boxes and IDs for the depicted persons. Based on these bounding boxes, we were able to create crops for feature extraction. To increase the stability of the crops, we smoothed the coordinates using a moving-window median correction. Furthermore, we annotated the videos to have a ground truth for supervised learning algorithms. Based on this preliminary work, we plan to extract features using DinoV2 and use those features within several machine-learning approaches such as contrastive learning.

Session Abstracts

Presentations of the young GAIN researchers

Presentation 3

Parameter-Efficient Transfer Learning for CNNs

Parameter-efficient fine-tuning methods, including adapters, have gained significant attention in the field of transformer models. However, their origins trace back to Convolutional Neural Networks (CNNs). Since Rebuffi et al. introduced this approach, numerous variations have emerged for specialized downstream tasks. Despite this proliferation, a systematic comparison of these variations in a more general setting has been lacking.

Concurrently, various tensor decomposition techniques have surfaced as parameter reduction strategies for CNN architectures. These methods share fundamental concepts with low-rank weight approximations, an approach that has also gained traction in transformer models. This commonality naturally leads to the generalization of LoRA-like adaptation schemas. However, the diversity of decomposition methods necessitates individual study and performance evaluation across a wide range of downstream tasks.

In this presentation, we offer a systematic exploration and side-by-side comparison of parameter-efficient fine-tuning methods for CNNs.

Presentation 4

A Comparison of Different Tokenization Methods for the Georgian Language

While the impact of tokenization on language modeling is well-researched in common languages, fewer studies on this topic exist for challenging low-resource languages. In this work, we present the first systematic evaluation of tokenization methods for Georgian, a low-resource language with high morphological complexity. We compare standard subword tokenizers, such as WordPiece, Byte Pair Encoding, SentencePiece with Unigram, and a recently proposed token-free approach.

We also investigate the multilingual BERT tokenizer (mBERT), which includes the Georgian language. In addition to these different classes of tokenization algorithms we also evaluate the impact of different vocabulary sizes, a key parameter for subword tokenizers. We evaluate the performance of all tokenizers on masked language modeling and on four downstream tasks: part-of-speech tagging, named entity recognition, toxicity detection, and sentiment analysis. We observe that larger vocabulary sizes for subword tokenizers generally lead to better performance across most tasks, with a notable exception in the toxicity detection task, where finer granularity is more effective. For the remaining tasks, pre-training tokenizers on Georgian text consistently yield better results compared to mBERT. Additionally, the token-free method is consistently outperformed by all other tokenizers.

Taken together, our comprehensive evaluation of tokenizers will be highly valuable in making informed tokenization choices in future language model development for the Georgian language.

Session Abstracts

Presentations of the young GAIN researchers

Presentation 5

EEG Dataset for Event Related Potentials (ERP)

The auditory evoked potential refers to the change in potential in the brain in response to an auditory stimulus. This change in potential can be detected using an electroencephalogram (EEG). Responses to auditory stimuli differ among individuals, but it is possible to identify common components of evoked waves. In this presentation, we discuss the components of auditory evoked stimuli and their relationship to various external and internal factors. Additionally, we review the oddball paradigm — frequently used method to investigate auditory evoked potentials, and present the data recorded at MICM with the commercial EEG device Unicorn Hybrid Black, discussing the obtained results.

Scientific Committee

1. Jan Alexandersson (DFKI)
2. Francois Bremond (INRIA)
3. Kartlos Kachiashvili (MICM)
4. Philipp Müller (DFKI)
5. Benedikt Emanuel Wirth (DFKI)

Local Organizing Committee

1. Valeri Berikashvili
2. Eka Chkonia
3. Giorgi Ghlonti
4. Giorgi Giorgobiani
5. Tamar Giorgobiani
6. Tsotne Javakhishvili
7. Rapael Kalandadze
8. David Karapetian
9. Ioseb Katchiashvili
10. Natia Kukhilava
11. Vakhtang Kvaratskhelia
12. Marina Menteshashvili
13. Luka Nadiradze
14. Teimuraz Saghinadze
15. Zaza Tabagari
16. Luka Tabagari