Second Workshop on Multimodal AI

Tuesday, 25th June 2024

Social Media: #MultimodalAI24

Venue: The Edge, The Endcliffe Village, 34 Endcliffe Crescent, Sheffield S10 3ED (direction)

Programme

Time	Event
09:00 - 09:30	Registration, morning refreshments, and poster session 1
09:30 - 09:35	Welcome: Guy Brown, Deputy Director of Centre for Machine Intelligence, University of Sheffield
09:35 - 10:00	Introduction: exploring multimodal AI beyond vision and language, Haiping Lu
10:00 - 10:40	Keynote 1: Daniel Zügner, Microsoft Research AI4Science **MatterGen: a generative model for inorganic materials design**
10:40 - 11:20	Community talks
11:20 - 11:50	Break and poster session 2
11:50 - 12:30	Keynote 2: Maria Liakata, Queen Mary University of London **Longitudinal language processing for dementia**
12:30 - 12:40	Group photos
12:40 - 14:00	Lunch and poster session 3
14:00 - 14:40	Keynote 3: Nataliya Tkachenko, Lloyds Banking Group **Ethical challenges for multimodal conversational banking & parametric insurance**
14:40 - 15:20	Community talks
15:20 - 15:30	Break
15:30 - 16:10	Keynote 4: Adam Steventon, Our Future Health **An incredibly detailed picture of human health: the exciting potential of Our Future Health to prevent, detect and treat diseases**
16:10 - 16:40	Panel discussion
16:40 - 17:00	Best talk/poster prize winner announcement and closing
17:00 - 17:30	Tea/coffee and networking

Morning Talk Session (10:40 - 11:20)

Name	Title
Yao Zhang	AI in Maritime Engineering Control System
Douglas Amoke	Geo-located Multimodal Data for Maritime Downstream Tasks
Yan Ge	Multimodal Multi-task Asset Pricing with Numeral Learning
Ruyi Wang	Multimodal Affective Computing for Mental Health Support
Luigi Moretti	Integrating Affective Computing and Smart Sensing into Treatment Pathways for Anxiety Disorders
Jiawei Zheng	Process-aware Human Activity Recognition
Lucas Farndale	Super Vision without Supervision: Self-supervised Learning from Multimodal Data for Enhanced Biomedical Imaging
Ruizhe Li	It's Never Too Late: Fusing Acoustic Information into Large Language Models for Automatic Speech Recognition

Afternoon Talk Session (14:40 - 15:20)

Name	Title
Valentin Danchev	Data Governance, Ethics, and Safety of Multimodal vs Unimodal AI Models: A Review of Evidence and Challenges
Salah (Sam) Hammouche	Beyond Regulatory Compliance: The RCR College Report
Nee Ling Wong	How Would AI Work with Us in Healthcare
Jiayan Zhang	Interdisciplinary Multimodal AI Research
Martin Callaghan	Multimodal AI for Enhanced Information Extraction from Complex HPC Documentation
Madhurananda Pahar	CognoSpeak: An Automatic, Remote Assessment of Early Cognitive Decline in Real-world Conversational Speech
Hubin Zhao	Wearable Intelligent Multimodal Neuroimaging for Health
Peter Charlton	Understanding Determinants of Health: Leveraging Routinely Collected Data

Keynotes

Keynote 1: Daniel Zügner, Senior Researcher, Microsoft Research AI4Science

Title: MatterGen: a generative model for inorganic materials design

Time: 10:00 - 10:40

Abstract: The design of functional materials with desired properties is essential in driving technological advances in areas like energy storage, catalysis, and carbon capture. Traditionally, materials design is achieved by screening a large database of known materials and filtering down candidates based on the application. Generative models provide a new paradigm for materials design by directly generating entirely novel materials given desired property constraints. In this talk, we present MatterGen, a generative model that generates stable, diverse inorganic materials across the periodic table and can further be fine-tuned to steer the generation towards a broad range of property constraints. To enable this, we introduce a new diffusion-based generative process that produces crystalline structures by gradually refining atom types, coordinates, and the periodic lattice. We further introduce adapter modules to enable fine-tuning towards any given property constraints with a labeled dataset. Compared to prior generative models, structures produced by MatterGen are more than twice as likely to be novel and stable, and more than 15 times closer to the local energy minimum. After fine-tuning, MatterGen successfully generates stable, novel materials with desired chemistry, symmetry, as well as mechanical, electronic and magnetic properties. Finally, we demonstrate multi-property materials design capabilities by proposing structures that have both high magnetic density and a chemical composition with low supply-chain risk. We believe that the quality of generated materials and the breadth of MatterGen's capabilities represent a major advancement towards creating a universal generative model for materials design.

Keynote 2: Maria Liakata, Professor of Natural Language Processing, Queen Mary University of London and Alan Turing Institute AI Fellow

Title: Longitudinal language processing for dementia

Time: 11:50 - 12:30

Abstract: While the advent of Large Language Modes (LLMs) has brought great promise to the field of AI there are many unresolved challenges especially around appropriate generation, temporal robustness, temporal and other reasoning and privacy concerns especially when working with sensitive content such as mental health data. The programme of work I have been leading consists in three core research directions: (1) data representation and generation (2) methods for personalised longitudinal models and temporal understanding (3) evaluation in real-world settings, with a focus on mental health. I will give an overview of work within my group on these topics and focus on work on longitudinal monitoring for dementia.

Keynote 3: Nataliya Tkachenko, Generative AI Ethics and Assurance Lead, Lloyds Banking Group

Title: Ethical challenges for multimodal conversational banking & parametric insurance

Time: 14:00 - 14:40

Abstract: Ever since mass-propagation of generative AI models, multimodal data has been getting increased attention from the customer-focused industries. Multimodal chatbots, which can process and respond to customer queries using enriched context, such as text, voice, and even visual data, offer significant advantages in customer banking and parametric insurance by enhancing user interaction, speed and overall service efficiency. Customers now have an option to choose their preferred mode of communication, whether through typing, speaking, or even using gestures. By analysing customer data from various sources, chatbots can offer personalised financial advice, investment recommendations, and alert about unusual activities. They even can help with the immediate payouts, by promptly verifying predefined parameters, such as weather data for crop insurance for example. However, with enriched context also come multi-dimensional ethical considerations, such as bias, fairness, transparency and confabulations. In this presentation I will cover how these risks emerge and mutually diffuse in highly automated interfaces.

Keynote 4: Adam Steventon, Director of Data Platforms, Our Future Health

Title: An incredibly detailed picture of human health: the exciting potential of Our Future Health to prevent, detect and treat diseases

Time: 15:30 - 16:10

Abstract: In this presentation, I will detail the groundbreaking efforts of Our Future Health to construct a multimodal dataset encompassing 5 million individuals, representative of the UK's diverse population. I will explore the transformative potential of this dataset to enhance our capabilities in predicting, detecting, and treating major diseases. Additionally, I will discuss the roles of artificial intelligence in this context, focusing on the opportunities and challenges it presents. This exploration will underscore the potential of AI and large-scale data in shaping the future of healthcare.

Posters

Name	Title
Douglas Amoke	Geo-located Multimodal Data for Maritime Downstream Tasks
Sedat Dogan	Enhanced Multimodal Learning for Meme Virality Prediction
Wenrui Fan	MeDSLIP: Medical Dual-Stream Language-Image Pre-training for Fine-grained Alignment
Lucas Farndale	Super Vision without Supervision: Self-supervised Learning from Multimodal Data for Enhanced Biomedical Imaging
Yan Ge	Multimodal Multi-task Asset Pricing with Numeral Learning
Ruizhe Li	Large Language Models are Efficient Learners of Noise-robust Speech Recognition
Xianyuan Liu	Exploring Multimodal AI beyond Vision and Language
Sabrina McCallum	Learning Generalisable Representations for Embodied Tasks with Multimodal Feedback
Luigi Moretti	Integrating Affective Computing and Smart Sensing into Treatment Pathways for Anxiety Disorders
Madhuranand Pahar	CognoSpeak: An Automatic, Remote Assessment of Early Cognitive Decline in Real-world Conversational Speech
Mohammod Suvon	Multimodal Variational Autoencoder for Low-cost Cardiac Hemodynamics Instability Detection
Ruyi Wang	Multimodal Affects Spreading and Development
Jiawei Zheng	Process-aware Human Activity Recognition

Panel Discussion (16:10 - 16:40)

Panel Members	Questions
• Halimat Afolabi • Maria Liakata • Venet Osmani • Adam Steventon • Nataliya Tkachenko • Fiona Young • Yurong Yu	1. What are the major barriers to deploying multimodal AI systems in real-world applications? 2. How can we best identify and utilise diverse data sources to advance multimodal AI research and applications?

Who to ask for help on site

If you have any questions or need assistance during the workshop, look for individuals wearing **YELLOW** lanyards. They are available to provide support and help make your experience enjoyable.

Parking at the Venue

We have a limited number of free parking permits available for parking at the Edge venue for those who require them. Please email multimodal-ai-enquiry-group@shef.ac.uk if you would like to receive a parking permit and have not already informed us of this need. Please note parking is on a first come first serve basis, and unfortunately, we are unable to reserve parking spaces.

Accessing the Internet

Complimentary WiFi is available throughout the venue. Connect to the network WiFiGuest by creating an account, with no password required (see instructions for more details). The Eduroam network is also available (connect to Eduroam).

Catering

Complimentary refreshments and lunch will be provided during the workshop. If you have specific dietary needs or catering questions, please speak to our team at the registration desk for assistance.

Filming and Photography

Please be advised that there will be media coverage, including filming and photography, during the workshop. The images and video taken may be used for promotional purposes on the Centre for Machine Intelligence and Department of Computer Science websites and social media channels. If you do not wish to appear in any video or photography, please inform the organisers via the contact multimodal-ai-enquiry-group@shef.ac.uk, or at the registration desk or speak to the photographer/ videographer.

Accessibility and Quiet Room

Please contact the organisers at multimodal-ai-enquiry-group@shef.ac.uk if you have any accessibility requirements that you would like to discuss, and we will endeavour to meet your requirements. There will be a quiet, private room available at the venue for delegates to use for various purposes. If you require access to the room, please contact the organisers via multimodal-ai-enquiry-group@shef.ac.uk or at the registration desk.

Partners

This workshop is jointly organised by the Alan Turing Institute's Interest Group on Meta-learning for Multimodal Data (welcome to sign-up and join) and the Multimodal AI Community (welcome to subscribe to our Google Group) supported by the University of Sheffield's Centre for Machine Intelligence.

The Alan Turing Institute

University of Sheffield | Centre for Machine Intelligence

Sponsors

We gratefully acknowledge the support of our sponsors, whose generosity has ensured the success of this workshop. Their contributions have been valuable in enabling students and researchers to attend and present their work.