NEURAL NETWORK WEIGHTS AS A NEW DATA MODALITY

WORKSHOP PROPOSAL - ICLR 2025

Workshop Summary

Scope. The ongoing deep learning revolution of the last decade has brought about hundreds of millions of neural networks (NNs) trained on diverse datasets. At the same time, the recent rise of foundation models has led to a rapid increase in the number of publicly available neural network models. On Hugging Face alone, there are over a million models, with thousands more added daily. As a result, the ample knowledge contained in the data, the abstraction learned via training, as well the trained models' behaviours themselves are stored in *the architectures and parameters of trained NNs*. Despite this massive growth, little research has been conducted into processing model weights, and they are rarely considered a data modality. This workshop aims to create a community around *Weight Space Learning* by bringing together the scattered sub-communities that already interface with model weights, with the ultimate goal of democratizing model weights as a proper data modality.

The proposed workshop aims to emphasise the potential of the *weight space modality* across several dimensions: (1) *Model analysis*: inferring model properties from its parameters (e.g., generalisation error) and mining information from trained networks by solely inspecting their weights. (2) *Model synthesis*: generating new models based on certain criteria (3) *Learning from model populations*, e.g. learning to select architectures/hyperparameters, learning to optimise, etc., or in general, replacing hand-designed components of ML pipelines with data-driven ones trained with data from NN populations (e.g., merging models to improve generalisation or robustness, pruning to improve efficiency, or even generating models to solve a new task without training). (4) *Applications to neural fields*: the advent of Implicit Neural Representations (INRs) in several fields (3D shape analysis, NeRF, dynamical systems modelling, etc.) has created the need for a new unifying toolbox for signal processing, where signals are represented as NNs - our workshop aims to contribute towards this goal.

Community Building. Although model weights hold great promise, current research is in its infancy. Recently, there has been a growing number of works that treat model weights as data points and attempt to learn weight representations (Eilertsen et al., 2020; Unterthiner et al., 2020; Schürholt et al., 2021, 2022; Navon et al., 2023a; Zhou et al., 2023a, 2024, 2023b; Lim et al., 2024; Kofinas et al., 2024; Shamsian et al., 2024; Kalogeropoulos et al., 2024; Schürholt et al.). Despite these advancements, research into model weights remains underexplored and is poised for significant growth in the coming years. Encouragingly, many research sub-communities interface with the weights of models and can gain from standardization of the field, e.g., model merging (Entezari et al., 2022; Wortsman et al., 2022; Ainsworth et al., 2023; Singh and Jaggi, 2020; Peña et al., 2023; Navon et al., 2023b), meta-learning & transfer learning (Metz et al., 2022; Knyazev et al., 2021, 2023), neural architecture search & hyperparameter optimization (Zhang et al., 2019; Thost and Chen, 2021; Mehta et al., 2024) and neural field processing/synthesis (Luigi et al., 2023; Serrano et al., 2024; Erkoç et al., 2023).

However, the field of weight space learning is in its early stages, with research scattered across multiple domains and researchers often unaware of each other's work. The goal of this workshop is to provide a bridge between these different communities and establish a common ground and language for researchers engaging in the field. By aligning terminology and methodologies, we aim to create better orientation and clarity within the field. By fostering dialogue between subfields, the workshop also aims to establish a clearer understanding of the current landscape and define key challenges moving forward.

Additionally, the workshop will serve as a **platform for exchange and future collaboration**, inviting the broader ICLR community to become aware of this emerging domain. By facilitating interdisciplinary exchange, the event encourages sustained progress in weight space learning, ensuring that ideas from different subfields converge into new research directions.

Research Goals and Key questions. This workshop aims to bring together researchers from the aforementioned diverse communities to explore fundamental questions about weight spaces:

- *What properties of weights, such as symmetries and invariances, present challenges or can be leveraged for optimization, learning and generalization?*
- *How can model weights be efficiently represented, manipulated, and used for downstream tasks?*
- *What model information can be decoded from model weights?*
- *Can model weights be generated for specific applications, to make training and model selection more efficient?*
- *Can weight space learning benefit research in processing and synthesising neural fields, for e.g. scientific applications and 3D vision?*
- *How can we democratize the usage of weight spaces, enabling more efficient research progress?*

The workshop will work towards establishing model weights as a new data modality through poster sessions, orals, and invited talks. In particular, it will feature weight space-relevant research in a wide variety of settings, including, but not limited to, theory of model weights, weight space representation learning, datasets and model zoos, efficient weight space learning, and novel applications of weight space learning.

Topics and call for papers. We invite contributions from researchers, practitioners, and industry experts that explore the promise of weight space learning. We invite submissions on a wide range of topics, including but not limited to:

- **Weight Space as a Modality:**
  - Characterization of weight space properties such as symmetries (e.g. permutations, scaling and beyond).
  - Weight space augmentations, scaling laws,
  - Model zoo datasets, etc.
- **Weight Space Learning Tasks/Learning paradigms:**
  - Supervised approaches, e.g., weight embeddings, meta-learning networks, (graph) hyper-networks.
  - Unsupervised approaches, e.g., autoencoders or hyper-representations.
  - Weight space learning backbones, e.g., plain MLPs, transformers, equivariant architectures, such as GNNs and neural functionals.
- **Theoretical Foundations:**
  - Expressivity of weight space processing modules.
  - Theoretical analysis of model weight properties.
  - Generalisation bounds of weight space learning methods.
- **Model/Weight Analysis:**
  - Inferring model properties and behaviours from their weights.
  - Investigating neural lineage and model trees through weights.
  - Learning dynamics in population-based training.
  - Interpretability of models via their weights.
- **Model/Weight Synthesis and Generation:**
  - Modeling weight distributions to facilitate weight sampling.
  - Generating weights in the context of e.g. transfer learning, learnable optimizers, implicit neural representation (INR) synthesis.
  - Model operations/editing (e.g. model merging, model soups, model pruning, task arithmetic).
  - Meta-learning and continual learning using model weights.
- **Applications of Weight Space Learning:**
  - Computer vision tasks (e.g. using NeRFs/INRs)
  - Applications to physics and dynamical system modelling.
  - Backdoor detection and adversarial robustness in weight space.

We invite submissions in the form of extended abstracts (4-6 pages) or full papers (8-12 pages). Accepted contributions will be presented in poster sessions and spotlight talks. Submissions should follow the ICLR formatting guidelines and be submitted via the official workshop submission portal.

We are interested in contributions that push the boundaries of weight space learning, including both theoretical advancements and novel applications across various domains.

Tentative Schedule and Organizational Details

Accessibility/Modality

Accepted papers and posters will be available on the workshop website https://weight-space-learning.github.io and the recorded talks will be made available on youtube. Accepted papers and posters will be made available for remote attendees. We will accommodate orals to be given remotely if the authors cannot attend in person. Additionally, we will set up a dedicated Discord channel to foster community building and interaction among participants, allowing remote attendees to engage with in-person participants and helping to cultivate a lasting community.

Anticipated Audience Size

Although weight-space learning is a relatively new field, its connections with various related communities make it highly relevant to a broad audience. We anticipate approximately 100 attendees, based on the increasing number of papers submitted to leading conferences like ICML 2024 that either directly explore weight-space learning or use model weights as their primary input for studying other topics. These scattered fields have yet to converge under a single interdisciplinary forum, and we hope this workshop will serve as the foundational nucleus for such a community.

Advertisement

We will promote the workshop website through social media platforms and email lists, as well as by sharing the workshop and call-for-papers within relevant research communities and institutional channels, such as those focused on model merging and model interpretability.

Conflicts of interest: No organizer will be involved in the assessment of a paper from the same institution or organization. The organizers will use their best efforts to ensure that the same will apply to reviewers, subject to feasibility during paper assignments. The above will be managed via openreview.

Tentative Timeline

- Submission deadline: February 3, 2025
- Author notification: March 5, 2025
- Camera-ready deadline: March 26, 2025.
- Import Workshop Program and Accepted Papers to iclr.cc: 27 March 2025
- Workshop day: April 27 or 28, 2025.

Tentative Schedule

The workshop aims to bring together the nascent community of weight-space learning, to establish a framework for the field, and future research directions and facilitate exchange and collaboration. To that end, the workshop is divided into three thematic blocks: **Graphs and Symmetries**, **Representation Learning**, and **Downstream Applications**. Each block begins with a contextualization from the steering committee, followed by an invited talk. The **Graphs and Symmetries** block will focus on leveraging symmetries (or lack thereof) and invariances in neural networks to better understand model behavior, while the **Representation Learning** session will explore the use of different backbones for this specific domain. Finally, the **Downstream Applications** block will cover how weight space learning can be applied to practical problems like transfer learning and model merging. Our invited talks will cover these domains and highlight interdisciplinary expertise across machine learning theory, neural network representation learning, and applied neural network analysis.

In addition to the invited talks, the workshop will feature spotlight talks and an interactive poster session to facilitate community engagement. These sessions allow participants to present their latest research and exchange ideas with peers. The poster session during the coffee break offers a casual environment for discussion, networking, and

collaboration. The workshop will conclude with closing remarks that summarize key takeaways and outline future opportunities for research and community-building in the growing field of weight space learning.

Table 1: Workshop Schedule

Time	Duration	Session	Session Content
9:00-9:30	30 min	**Introduction and opening remarks**	- Welcome / workshop origin - Overview of the weight space domain - Agenda & Logistics
9:30-10:30	60 min	**Opening Keynote**	Invited Talk
10:30-11:00	30 min	**Coffee Break**	
11:00-11:15	15 min	**Session 1 - Graphs and Symmetries**	Steering Committee / contextualization
11:15-11:45	30 min		Invited Talk (20 min + 10 min Q&A)
11:45-12:15	30 min		2x Spotlight Talks (each 15 min)
12:15-13:30	75 min	**Lunch Break**	
13:30-13:45	15 min	**Session 2 - Representation Learning**	Steering Committee / contextualization
13:45-14:15	30 min		Invited Talk (20 min + 10 min Q&A)
14:15-14:45	30 min		2x Spotlight Talks (each 15 min)
14:45-16:00	75 min	**Coffee Break + Poster Session**	Interactive poster session + coffee break
16:00-16:15	15 min	**Session 3 - Downstream Applications**	Steering Committee / contextualization
16:15-16:45	30 min		Invited Talk (20 min + 10 min Q&A)
16:45-17:00	15 min	**Closing**	- Recap - Next workshop / other opportunities - Next steps for the community
17:00	-	**End**	

Invited Speakers

The invited talks complement the steering committee contextualization as in-depth talks from experts of their respective fields. The speakers are chosen to cover a broad spectrum of weight-space learning topics, from exploring the weight space domain (Michael Mahoney, Naomi Saphra), over weight-space learning methods (Chelsea Finn, Stella Yu, Boris Knyazev) to applications (Michael Mahoney, Boris Knyazev, Chelsea Finn). The talks make these different perspectives accessible to the attendees and provide an introduction to the speaker's leading research on their respective domains. Together, they cover the field of weight space learning and help define cornerstones of the weight space learning framework.

- **Chelsea Finn (tentative)** Chelsea Finn is an Assistant Professor in Computer Science and Electrical Engineering at Stanford University and co-founder of Pi. Her lab, IRIS, studies intelligence through robotic interaction at scale, and is affiliated with SAIL and the ML Group. She is interested in the capability of robots and other agents to develop broadly intelligent behavior through learning and interaction. Previously, Chelsea completed her Ph.D. in computer science at UC Berkeley and my B.S. in electrical engineering and computer science at MIT. She also spent time at Google as part of the Google Brain team.
- **Michael Mahoney (confirmed)** Michael W. Mahoney is a professor at the University of California at Berkeley in the Department of Statistics and at the International Computer Science Institute (ICSI). He is also an Amazon Scholar as well as head of the Machine Learning and Analytics Group at the Lawrence Berkeley National Laboratory. He works on algorithmic and statistical aspects of modern large-scale data analysis. Much of his recent research has focused on large-scale machine learning, including randomized matrix algorithms and randomized numerical linear algebra, scientific machine learning, scalable stochastic optimization,

geometric network analysis tools for structure extraction in large informatics graphs, scalable implicit regularization methods, computational methods for neural network analysis, physics informed machine learning, and applications in genetics, astronomy, medical imaging, social network analysis, and internet data analysis. He received his PhD from Yale University with a dissertation in computational statistical mechanics, and he has worked and taught at Yale University in the mathematics department, at Yahoo Research, and at Stanford University in the mathematics department.

- **Stella X. Yu (confirmed)** Stella X. Yu is a professor of computer science at the University of Michigan, where she focuses on research in computer vision and machine learning. Prior to joining the University of Michigan, she was the Director of the Vision Group at the International Computer Science Institute (ICSI) in Berkeley and held various academic roles at UC Berkeley, including positions in computer science, vision science, and cognitive sciences. Stella earned her Ph.D. from Carnegie Mellon University, where she specialized in robotics and vision science. Her research explores visual perception through multiple lenses of representation learning aiming to develop models that can exceed human capabilities. She is particularly focused on actionable representation learning and structure-aware models, emphasizing that structure in visual data should naturally emerge or be reflected in the model structures.
- **Boris Knyazev (confirmed)** Boris Knyazev is a Research Scientist at Samsung - SAIT AI Lab, Montreal, Canada. He completed his PhD at the Machine Learning Research Group, University of Guelph and Vector Institute under supervision of Graham Taylor in 2022. His research interests lie at the intersection of graph neural networks (GNNs), computer vision and meta-learning. In the past, he interned at Facebook AI Research (FAIR) working with Adriana Romero and Michal Drozdal on parameter prediction for neural networks. He also interned at Mila working with Eugene Belilovsky and Aaron Courville on visual compositional generalization. He also interned at SRI International with Mohamed Amer, where he worked on training GNNs on image superpixels. Before starting his PhD, he worked on unsupervised learning and pretraining of neural networks, face, emotion and facial attributes recognition, and video recognition.
- **Naomi Saphra (confirmed)** is a research fellow at the Kempner Institute at Harvard University. She is interested in NLP training dynamics: how models learn to encode linguistic patterns or other structure and how we can encode useful inductive biases into the training process. Naomi has earned a PhD from the University of Edinburgh on Training Dynamics of Neural Language Models; worked at NYU, Google and Facebook; and attended Johns Hopkins and Carnegie Mellon University. Outside of research, she plays roller derby under the name Gaussian Retribution, perform standup comedy, and shepherd disabled programmers into the world of code dictation.

Committees

Organizing Committee

Our organizing committee benefits from diverse backgrounds and the research experience of its members spans a broad range of the weight learning spectrum: weight space symmetries and their interplay with optimization, generalization, weight space networks etc (Yoav Gelberg, Bo Zhao, Derek Lim, Allan Zhou, Giorgos Bouritsas, Stefanie Jegelka), representation learning and generative modelling of weights (Damian Borth, Konstantin Schürholt), model analysis and model interpretability (Eliahu Horwitz). Additionally, the team is composed of both academic and industry researchers at different levels of seniority, from PhD students over postdocs to full professors. Please refer to our Diversity Commitment section for additional information on diversity w.r.t. geographic location, gender, ethnicity etc.

Several of our committee members have served as program chairs, have organized workshops, and have the necessary experience to make the workshop a success. In addition, the organizing committee can rely on the input and guidance from the experienced members of the steering committee.

- **Konstantin Schürholt (University of St. Gallen)**
  - Email: `konstantin.schuerholt@unisg.ch`
  - Website: https://kschuerholt.github.io/
  - Google Scholar: https://scholar.google.com/citations?user=refZxl4AAAAJ&hl=de
  - Bio: Konstantin Schürholt is a postdoctoral researcher at University of St. Gallen, where he obtained his PhD in Computer Science earlier this year. His research is focused on representation learning on neural network weights to identify latent weight structure and use it for model analysis and generation, as well as on phase transitions in neural networks. During his PhD, he visited ICSI Berkeley as a visiting scholar and interned at Google Deepmind.

- **Giorgos Bouritsas (Archimedes AI & University of Athens)**.
  - Email: g.bouritsas@athenarc.gr
  - Google Scholar: https://scholar.google.com/citations?user=eNUJDXUAAAAJ&hl=en
  - Bio: Giorgos Bouritsas is a machine learning researcher and holds a postdoctoral fellow position at the Archimedes AI unit and the University of Athens, as well as an adjunct lecturer position at NCSR Demokritos. He obtained his PhD in computer science from Imperial College London, and his MEng in electrical and computer engineering from the National Technical University of Athens. Previously, he spent time at Google Deepmind, École Polytechnique Fédérale de Lausanne, KU Leuven, NCSR Demokritos and Universitat Politècnica de Catalunya. He has conducted basic and applied research on several machine learning topics, particularly within Geometric/Graph Deep Learning, where he designs and theoretically analyses data-driven methodologies for geometric and symmetric data, such as complex networks, physical systems and 3D objects. Currently, he aims to deepen the understanding of learning on weight spaces, with his latest publication involving a study through the lens of scaling symmetries. His work has been published in leading conferences (NeurIPS, ICML, CVPR, ICCV, ECCV) and journals (TPAMI), while he regularly performs educational activities (tutorials, postgraduate courses) and provides academic service as a reviewer in leading conferences and journals (e.g. ICLR, ICML, NeurIPS, TPAMI, TMLR), for which he has received outstanding reviewer awards.
- **Eliahu Horwitz (HUJI)**
  - Email: eliahu.horwitz@mail.huji.ac.il
  - Website: https://pages.cs.huji.ac.il/eliahu-horwitz/
  - Google Scholar: https://scholar.google.com/citations?user=NyLx5nIAAAAJ&hl=en
  - Bio: Eliahu Horwitz is a PhD candidate in Computer Science at the Hebrew University of Jerusalem. His research focuses on learning representations of neural network weights and exploring their applications in downstream tasks. He is particularly interested in how weight-space learning can benefit other areas of machine learning. Eliahu is a recipient of the Israeli Council for Higher Education Scholarship and has previously interned at Google Research.
- **Derek Lim (MIT & Liquid AI)**
  - Email: dereklim@mit.edu
  - Website: https://cptq.github.io/
  - Google Scholar: https://scholar.google.com/citations?user=y9YTBIsAAAAJ&hl=en
  - Bio: Derek Lim is a PhD student in computer science at MIT, and a research scientist at Liquid AI. His research spans several topics related to symmetries in deep learning. He has worked and published on equivariant weight-space networks, and on understanding the empirical impact of parameter symmetries in various neural networks. Previously, he has done research internships at Meta, NVIDIA, Cornell, and Johns Hopkins University. Derek is also the founder and organizer of the Boston Symmetry Day workshop https://bostonsymmetry.github.io/, the founder and 2022 organizer of the Learning on Graphs Conference https://logconference.org/, and the 2024 organizer of the Learning on Graphs New York City Meetup https://logmeetupnyc.github.io/.
- **Yoav Gelberg (University of Oxford)**
  - Email: yoav@robots.ox.ac.uk
  - Google Scholar: https://scholar.google.com/citations?user=FMbR_TkAAAAJ&hl=en
  - Bio: Yoav Gelberg is a computer science PhD student at the University of Oxford, where he is part of the AIMS CDT program. He completed his undergraduate studies at the Technion as a Rothschild scholar. Yoav's research interests include geometric and topological deep learning, and neural networks weight spaces. His work has explored the impacts of parameter symmetries on variational inference and Bayesian neural networks, as well as the expressive capabilities of topological deep learning architectures.
- **Bo Zhao (University of California San Diego)**
  - Email: bozhao@ucsd.edu
  - Website: https://b-zhao.github.io
  - Google Scholar: https://scholar.google.com/citations?hl=en&user=ZCCrFoIAAAAJ
  - Bio: Bo Zhao is a computer science PhD student at the University of California San Diego. Her research interests lie in mathematical structures in deep learning, with recent focus on theoretical frameworks of symmetry in neural network weight spaces. Her work has also explored applications of weight space

symmetries on optimization, generalization, and understanding the loss landscape. She has co-organized the Women in Machine Learning workshop and served as the workshop chair assistant for NeurIPS 2024.

- **Allan Zhou (Google Deepmind)**
  - Email: ayz@cs.stanford.edu
  - Website: https://bland.website/
  - Google Scholar: https://scholar.google.com/citations?user=6S9C8XoAAAAJ
  - Bio: Allan Zhou is a research engineer at Google DeepMind, where he works on AI and robotics. He recently completed his PhD at Stanford, advised by Chelsea Finn. There, his research focused on geometric deep learning, nested optimization, and imitation & reinforcement learning. Allan also worked on learned optimizers at Google DeepMind and on reinforcement learning at FAIR.
- **Damian Borth (HSG, UW & TU/e)**
  - Email: damian.borth@unisg.ch
  - Website: https://www.unisg.ch/en/university/about-us/organisation/detail/person-id/f72a2a1d-94ae-4b74-ae82-cc898ae53284/
  - Google Scholar: https://scholar.google.com/citations?user=J-8Z038AAAAJ&hl=en
  - Bio: Damian Borth holds a full professorship in Artificial Intelligence and Machine Learning (AIML) at the University of St.Gallen, Switzerland. Previously, Damian was the founding director of the Deep Learning Competence Center at the German Research Center for Artificial Intelligence (DFKI) in Kaiserslautern, where he was also PI of the NVIDIA AI Lab at the DFKI. His research focuses on representation learning of neural network's weight spaces. His work has been awarded with the ACM SIGMM Test of Time Award 2023, Google Research Scholar Award 2022, and the NVIDIA AI Lab at NVIDIA GTC Europe 2016. Damian did his postdoctoral research at ICSI and UC Berkeley, where he was involved in large-scale data projects at the Lawrence Livermore National Laboratory. Damian has served as Area Chair at NeurIPS, WACV and has organized several workshops in the past. He received his PhD from the University of Kaiserslautern and the German Research Center for Artificial Intelligence (DFKI). During that time, Damian stayed as a visiting researcher at the Digital Video and Multimedia Lab at Columbia University, New York City, USA. Currently, he is visiting professor at University of Washington, Seattle, USA and EAISI visiting professor at TU Eindhoven, Netherlands. Damian served as the organization chair for the MMCommons workshops in 2015 and 2016, held in conjunction with the release of the YFCC100m dataset. He was also the organization chair for the Deep Learning workshop at the International Supercomputing Conference (ISC) in 2017.
- **Stefanie Jegelka (MIT & TU Munich)**
  - Email: stefje@mit.edu
  - Website: https://people.csail.mit.edu/stefje/
  - Google Scholar: https://scholar.google.ch/citations?hl=en&user=gTWUZlsAAAAJ&view_op=list_works
  - Bio: Stefanie Jegelka is a Humboldt Professor at TU Munich and an Associate Professor in the Department of EECS at MIT (on leave). Before joining MIT, she was a postdoctoral researcher at UC Berkeley, and obtained her PhD from ETH Zurich and the Max Planck Institute for Intelligent Systems. Stefanie has received a Sloan Research Fellowship, an NSF CAREER Award, a DARPA Young Faculty Award, an Alexander von Humboldt Professorship, Google research awards, a Two Sigma faculty research award, the German Pattern Recognition Award and a Best Paper Award at ICML, and she was a sectional lecturer at the ICM 2022. She has served as (senior) Area Chair for NeurIPS and ICML and as Program Chair for ICML 2022, and she has organized multiple workshops. Her research interests span the theory and practice of learning with graphs, learning with symmetries, robustness to distribution shifts, and learning with limited supervision. Stefanie served as the program chair of ICML 2022, is a member of the advisory board for the Learning on Graphs conference, and is a founder and organizer of the Boston Symmetry Day workshop. She has also organized several workshops, including Heavy Tails in ML (NeurIPS 2023) and New Frontiers in Graph Learning (NeurIPS 2022).

Steering Committee

- **Michael Bronstein (Oxford)**
  - Email: michael.bronstein@cs.ox.ac.uk
  - Website: https://www.cs.ox.ac.uk/people/michael.bronstein/

- Google Scholar: https://scholar.google.com/citations?user=UU3N6-UAAAAJ&hl=en
- Bio: Michael Bronstein is the DeepMind Professor of AI at the University of Oxford. He was previously Head of Graph Learning Research at Twitter, a professor at Imperial College London and held visiting appointments at Stanford, MIT, and Harvard. He has been affiliated with three Institutes for Advanced Study (at TUM as a Rudolf Diesel Fellow (2017-2019), at Harvard as a Radcliffe fellow (2017-2018), and at Princeton as a short-time scholar (2020)). Michael received his PhD from the Technion in 2007. He is the recipient of the EPSRC Turing AI World Leading Research Fellowship, Royal Society Wolfson Research Merit Award, Royal Academy of Engineering Silver Medal, five ERC grants, two Google Faculty Research Awards, and two Amazon AWS ML Research Awards. He is a Member of the Academia Europaea, Fellow of IEEE, IAPR, BCS, and ELLIS, ACM Distinguished Speaker, and World Economic Forum Young Scientist. In addition to his academic career, Michael is a serial entrepreneur and founder of multiple startup companies, including Novafora, Invision (acquired by Intel in 2012), Videocites, and Fabula AI (acquired by Twitter in 2019).

- **Gal Chechik(NVIDIA/BIU)**
  - Email: gal.chechik@gmail.com
  - Website: https://research.nvidia.com/person/gal-chechik
  - Google Scholar: https://scholar.google.com/citations?user=Wk2gAZUAAAAJ&hl=en
  - Bio: Gal Chechik is a Sr. Director of AI research, leading NVIDIA research in Israel. Gal is also a Professor of computer science at Bar-Ilan University. Before joining NVIDIA, he was a Staff Research Scientist at Google, a postdoctoral research associate at Stanford University, and received his PhD from the Hebrew University of Jerusalem. Gal published 140 papers, including publications in Nature Biotechnology, Cell and PNAS, and holds 50 issued patents. His work won awards for outstanding papers at NeurIPS and ICML. Gal's group of NVIDIA research in Israel spans algorithms, theory and applications of deep learning, with a focus on computer vision and reinforcement learning. A particular interest is in perception, action, and reasoning (PAR) and their intersection for the purpose of smarter generalization.
- **Stella Yu (University of Michigan)**
  - Email: stellayu@umich.edu
  - Website: https://web.eecs.umich.edu/~stellayu/
  - Google Scholar: https://scholar.google.com/citations?user=uqWkLzMAAAAJ&hl=en
  - Bio: Stella X. Yu is a professor of computer science at the University of Michigan, where she focuses on research in computer vision and machine learning. Prior to joining the University of Michigan, she was the Director of the Vision Group at the International Computer Science Institute (ICSI) in Berkeley and held various academic roles at UC Berkeley, including positions in computer science, vision science, and cognitive sciences. Stella earned her Ph.D. from Carnegie Mellon University, where she specialized in robotics and vision science. Her research explores visual perception through multiple lenses of representation learning aiming to develop models that can exceed human capabilities. She is particularly focused on actionable representation learning and structure-aware models, emphasizing that structure in visual data should naturally emerge or be reflected in the model structures.
- **Haggai Maron (Technion,NVIDIA)**
  - Email: hmaron@nvidia.com
  - Website: https://haggaim.github.io/
  - Google Scholar: https://scholar.google.co.il/citations?user=4v8uJrIAAAAJ
  - Haggai Maron is an Assistant Professor at the Technion's Faculty of Electrical and Computer Engineering and a senior research scientist at NVIDIA Research. He earned his PhD in Computer Science and Applied Mathematics from the Weizmann Institute of Science. His research focuses on machine learning, particularly deep learning for structured data such as sets, graphs, point clouds, surfaces, and weight spaces. Haggai's work has been recognized with an outstanding paper award at ICML 2020. He co-organized the Israeli Geometric Deep Learning Workshops and co-organizes the Graph Learning Meets Theoretical Computer Science Workshop at the Simons Institute in 2025.
- **Yedid Hoshen (HUJI)**
  - Email: yedid.hoshen@mail.huji.ac.il
  - Website: https://www.cs.huji.ac.il/~yedid/
  - Google Scholar: https://scholar.google.co.il/citations?user=6y1-qS4AAAAJ&hl=en

- Bio: Yedid Hoshen is an Associate Professor at the Hebrew University of Jerusalem, Israel (2019-present). He is also a visiting faculty researcher at Google. Prior to that he was a Postdoc and researcher scientist in Facebook AI Research, New York, New York and Tel Aviv. He earned his PhD at the Hebrew University and MPhys at Oxford University. Yedid's research focuses on representation learning and its applications e.g., anomaly detection, disentanglement, cross-domain retrieval and generation. Yedid's lab is currently very active in developing new ways to represent models and its real-world use cases. Yedid has served as an area chair in ECCV'22, CVPR'23,24, NeurIPS'23,24 and will do so in CVPR'25.

Diversity Commitment

Our workshop brings together interdisciplinary perspectives from geometric deep learning, representation learning, and applications for weight spaces, ensuring a broad and inclusive dialogue. The invited speakers, organizers and steering committee members reflect our commitment to gender diversity, with 60%, 22% and 20% female representation among the speakers, organizing committee and steering committee, respectively.

The composition of our committees reflects a commitment to diversity in geographic and institutional affiliations as well as seniority. Geographically, the organizing team and speakers span Europe, the Middle East, Asia, and America. The team includes representatives from both academia and industry. We include participants from various career stages, ranging from PhD students to postdoctoral researchers, professors, and industry professionals.

With a combination of invited talks, contributed talks, and poster sessions, the program encourages engagement and discussion while addressing diversity and inclusion through a variety of voices from academia and industry, at different seniority, from different backgrounds and of different genders.

The schedule fosters collaboration and interaction, allowing both early-career researchers and under-represented groups to contribute meaningfully to the growing field.

Under-represented/under-resourced researchers. We aim for inclusivity within the broader ICLR community and encourage participation from underrepresented communities. To that end, we will seek funding from industry sponsors to cover scholarships for workshop fees and travel cost for attendees of under-represented communities. Additionally, as per our call for papers, we will invite extended abstract submissions (short papers track) to facilitate under-resourced researchers to submit preliminary or modest results (not yet ready to be published as full conference papers, but that will greatly benefit from feedback and visibility in our workshop).

Previous Related Workshops

Some past workshops have touched on subsets of the topics and themes of our workshop, and can be grouped as follows:

- Several workshops are related to *symmetry and geometry* in various aspects of machine learning, such as NeurReps, TAG, GRaM, HiLD, Re-Align. These workshops focus on properties of loss landscapes (HiLD), properties of representations (NeurReps, Re-Align), equivariance/symmetries in functions (NeurReps, TAG, GRaM), or other unrelated geometric topics (such as using tools from geometry or topology to make new neural network modules). Our workshop intersects with the above since symmetries play a crucial role in weight processing (equivariant) architectures (also known as neural functionals, metanetworks, etc.), optimisation and model merging. However, none of these workshops focused on weight spaces as a primary modality for analysis and processing, and to the best of our knowledge the progress achieved in this particular field via these workshops has been little.
- Other workshops have considered various *focused aspects of weight space analysis/processing*. For instance, the ATTRIB workshop covers attributing model behavior to factors including data, activations, and weights. Additionally, the NeurIPS 2024 Model Merging Challenge is about merging models in weight spaces. Our workshop aims to host diverse types of weight space analysis problems and aims to be a superset of the above.
- Additionally, starting from NeRF there has been strong interest in the computer vision and the broader machine learning community in Implicit Neural Representations/Neural Fields. Examples of workshops include but are not limited to Neural Fields across Fields, INRV, etc. As in the previous bullet point, our call for papers includes diverse applications similar to the aforementioned.
- Finally, the UniReps workshop focused on *similarities across various representations, including neural ones* hosting topics that intersect with our call for papers in certain aspects (e.g. model merging, how symmetries affect neural representations, etc.), but did not deal with the modality nature of weight spaces.

References

Samuel K. Ainsworth, Jonathan Hayase, and Siddhartha S. Srinivasa. Git re-basin: Merging models modulo permutation symmetries. In *The Eleventh International Conference on Learning Representations, ICLR 2023, Kigali, Rwanda, May 1-5, 2023*. OpenReview.net, 2023. 1

Gabriel Eilertsen, Daniel Jönsson, Timo Ropinski, Jonas Unger, and Anders Ynnerman. Classifying the classifier: Dissecting the weight space of neural networks. In *ECAI 2020 - 24th European Conference on Artificial Intelligence*, pages 1119–1126. IOS Press, 2020. 1

Rahim Entezari, Hanie Sedghi, Olga Saukh, and Behnam Neyshabur. The role of permutation invariance in linear mode connectivity of neural networks. In *The Tenth International Conference on Learning Representations, ICLR 2022, Virtual Event, April 25-29, 2022*. OpenReview.net, 2022. 1

Ziya Erkoç, Fangchang Ma, Qi Shan, Matthias Nießner, and Angela Dai. Hyperdiffusion: Generating implicit neural fields with weight-space diffusion. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 14300–14310, 2023. 1

Ioannis Kalogeropoulos, Giorgos Bouritsas, and Yannis Panagakis. Scale equivariant graph metanetworks. *Advances in Neural Information Processing Systems*, 2024. 1

Boris Knyazev, Michal Drozdzal, Graham W Taylor, and Adriana Romero Soriano. Parameter prediction for unseen deep architectures. *Advances in Neural Information Processing Systems*, 34:29433–29448, 2021. 1

Boris Knyazev, Doha Hwang, and Simon Lacoste-Julien. Can we scale transformers to predict parameters of diverse imagenet models? In *International Conference on Machine Learning*, pages 17243–17259. PMLR, 2023. 1

Miltiadis Kofinas, Boris Knyazev, Yan Zhang, Yunlu Chen, Gertjan J. Burghouts, Efstratios Gavves, Cees G. M. Snoek, and David W. Zhang. Graph neural networks for learning equivariant representations of neural networks. In *The Twelfth International Conference on Learning Representations*, 2024. 1

Derek Lim, Haggai Maron, Marc T. Law, Jonathan Lorraine, and James Lucas. Graph metanetworks for processing diverse neural architectures. In *The Twelfth International Conference on Learning Representations (ICLR)*, 2024. 1

Luca De Luigi, Adriano Cardace, Riccardo Spezialetti, Pierluigi Zama Ramirez, Samuele Salti, and Luigi Di Stefano. Deep learning on implicit neural representations of shapes. In *The Eleventh International Conference on Learning Representations, ICLR 2023, Kigali, Rwanda, May 1-5, 2023*. OpenReview.net, 2023. 1

Nikhil Mehta, Jonathan Lorraine, Steve Masson, Ramanathan Arunachalam, Zaid Pervaiz Bhat, James Lucas, and Arun George Zachariah. Improving hyperparameter optimization with checkpointed model weights. *The 18th European Conference on Computer Vision (ECCV) Efficient Foundation Models (EFM) Workshop*, 2024. 1

Luke Metz, James Harrison, C. Daniel Freeman, Amil Merchant, Lucas Beyer, James Bradbury, Naman Agrawal, Ben Poole, Igor Mordatch, Adam Roberts, and Jascha Sohl-Dickstein. Velo: Training versatile learned optimizers by scaling up. *CoRR*, abs/2211.09760, 2022. 1

Aviv Navon, Aviv Shamsian, Idan Achituve, Ethan Fetaya, Gal Chechik, and Haggai Maron. Equivariant architectures for learning in deep weight spaces. In *International Conference on Machine Learning*, pages 25790–25816. PMLR, 2023a. 1

Aviv Navon, Aviv Shamsian, Ethan Fetaya, Gal Chechik, Nadav Dym, and Haggai Maron. Equivariant deep weight space alignment. *arXiv preprint arXiv:2310.13397*, 2023b. 1

Fidel A Guerrero Peña, Heitor Rapela Medeiros, Thomas Dubail, Masih Aminbeidokhti, Eric Granger, and Marco Pedersoli. Re-basin via implicit sinkhorn differentiation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 20237–20246, 2023. 1

Konstantin Schürholt, Michael W Mahoney, and Damian Borth. Towards scalable and versatile weight space learning. In *International Conference on Machine Learning*. 1

Konstantin Schürholt, Dimche Kostadinov, and Damian Borth. Self-supervised representation learning on neural network weights for model characteristic prediction. *Advances in Neural Information Processing Systems*, 34:16481–16493, 2021. 1

Konstantin Schürholt, Boris Knyazev, Xavier Giró-i Nieto, and Damian Borth. Hyper-representations as generative models: Sampling unseen neural network weights. *Advances in Neural Information Processing Systems*, 35:27906–27920, 2022. 1

Louis Serrano, Lise Le Boudec, Armand Kassaï Koupaï, Thomas X Wang, Yuan Yin, Jean-Noël Vittaut, and Patrick Gallinari. Operator learning with neural fields: Tackling pdes on general geometries. *Advances in Neural Information Processing Systems*, 36, 2024. 1

Aviv Shamsian, Aviv Navon, David W. Zhang, Yan Zhang, Ethan Fetaya, Gal Chechik, and Haggai Maron. Improved generalization of weight space networks via augmentations. *CoRR*, abs/2402.04081, 2024. 1

Sidak Pal Singh and Martin Jaggi. Model fusion via optimal transport. *Advances in Neural Information Processing Systems*, 33: 22045–22055, 2020. 1

Veronika Thost and Jie Chen. Directed acyclic graph neural networks. In *International Conference on Learning Representations*, 2021. 1

Thomas Unterthiner, Daniel Keysers, Sylvain Gelly, Olivier Bousquet, and Ilya Tolstikhin. Predicting neural network accuracy from weights. *arXiv preprint arXiv:2002.11448*, 2020. 1

Mitchell Wortsman, Gabriel Ilharco, Samir Ya Gadre, Rebecca Roelofs, Raphael Gontijo-Lopes, Ari S Morcos, Hongseok Namkoong, Ali Farhadi, Yair Carmon, Simon Kornblith, et al. Model soups: averaging weights of multiple fine-tuned models improves accuracy without increasing inference time. In *International conference on machine learning*, pages 23965–23998. PMLR, 2022. 1

Muhan Zhang, Shali Jiang, Zhicheng Cui, Roman Garnett, and Yixin Chen. D-vae: A variational autoencoder for directed acyclic graphs. *Advances in neural information processing systems*, 32, 2019. 1

Allan Zhou, Kaien Yang, Kaylee Burns, Adriano Cardace, Yiding Jiang, Samuel Sokota, J Zico Kolter, and Chelsea Finn. Permutation equivariant neural functionals. *Advances in Neural Information Processing Systems*, 36, 2023a. 1

Allan Zhou, Kaien Yang, Yiding Jiang, Kaylee Burns, Winnie Xu, Samuel Sokota, J Zico Kolter, and Chelsea Finn. Neural functional transformers. In *Thirty-seventh Conference on Neural Information Processing Systems*, 2023b. 1

Allan Zhou, Chelsea Finn, and James Harrison. Universal neural functionals. *CoRR*, abs/2402.05232, 2024. 1