# Mon 2024.03.25

## Hyperbolic Metric Learning for Visual Outlier Detection

*Alvaro Gonzalez-Jimenez, Simone Lionetti, Dena Bazazian, Philippe Gottfrois, Fabian Gröger, Marc Pouly, Alexander Navarini*

Out-Of-Distribution (OOD) detection is critical to deploy deep learning models in safety-critical applications. However, the inherent hierarchical concept structure of visual data, which is instrumental to OOD detection, is often poorly captured by conventional methods based on Euclidean geometry. This work proposes a metric framework that leverages the strengths of Hyperbolic geometry for OOD detection. Inspired by previous works that refine the decision boundary for OOD data with synthetic outliers, we extend this method to Hyperbolic space. Interestingly, we find that synthetic outliers do not benefit OOD detection in Hyperbolic space as they do in Euclidean space. Furthermore we explore the relationship between OOD detection performance and Hyperbolic embedding dimension, addressing practical concerns in resource-constrained environments. Extensive experiments show that our framework improves the FPR95 for OOD detection from 22\% to 15\% and from 49% to 28% on CIFAR-10 and CIFAR-100 respectively compared to Euclidean methods.

link: http://arxiv.org/abs/2403.15260v1

## Federated Bayesian Deep Learning: The Application of Statistical Aggregation Methods to Bayesian Models

*John Fischer, Marko Orescanin, Justin Loomis, Patrick McClure*

Federated learning (FL) is an approach to training machine learning models that takes advantage of multiple distributed datasets while maintaining data privacy and reducing communication costs associated with sharing local datasets. Aggregation strategies have been developed to pool or fuse the weights and biases of distributed deterministic models; however, modern deterministic deep learning (DL) models are often poorly calibrated and lack the ability to communicate a measure of epistemic uncertainty in prediction, which is desirable for remote sensing platforms and safety-critical applications. Conversely, Bayesian DL models are often well calibrated and capable of quantifying and communicating a measure of epistemic uncertainty along with a competitive prediction accuracy. Unfortunately, because the weights and biases in Bayesian DL models are defined by a probability distribution, simple application of the aggregation methods associated with FL schemes for deterministic models is either impossible or results in sub-optimal performance. In this work, we use independent and identically distributed (IID) and non-IID partitions of the CIFAR-10 dataset and a fully variational ResNet-20 architecture to analyze six different aggregation strategies for Bayesian DL models. Additionally, we analyze the traditional federated averaging approach applied to an approximate Bayesian Monte Carlo dropout model as a lightweight alternative to more complex variational inference methods in FL. We show that aggregation strategy is a key hyperparameter in the design of a Bayesian FL system with downstream effects on accuracy, calibration, uncertainty quantification, training stability, and client compute requirements.

link: http://arxiv.org/abs/2403.15263v1

## Parametric PDE Control with Deep Reinforcement Learning and Differentiable L0-Sparse Polynomial Policies

*Nicolò Botteghi, Urban Fasel*

Optimal control of parametric partial differential equations (PDEs) is crucial in many applications in engineering and science. In recent years, the progress in scientific machine learning has opened up new frontiers for the control of parametric PDEs. In particular, deep reinforcement learning (DRL) has the potential to solve high-dimensional and complex control problems in a large variety of applications. Most DRL methods rely on deep neural network (DNN) control policies. However, for many dynamical systems, DNN-based control policies tend to be over-parametrized, which means they need large amounts of training data, show limited robustness, and lack interpretability. In this work, we leverage dictionary learning and differentiable L$_0$ regularization to learn sparse, robust, and interpretable control policies for parametric PDEs. Our sparse policy architecture is agnostic to the DRL method and can be used in different policy-gradient and actor-critic DRL algorithms without changing their policy-optimization procedure. We test our approach on the challenging tasks of controlling parametric Kuramoto-Sivashinsky and convection-diffusion-reaction PDEs. We show that our method (1) outperforms baseline DNN-based DRL policies, (2) allows for the derivation of interpretable equations of the learned optimal control laws, and (3) generalizes to unseen parameters of the PDE without retraining the policies.

link: http://arxiv.org/abs/2403.15267v1

### Imagination Augmented Generation: Learning to Imagine Richer Context for Question Answering over Large Language Models

*Huanxuan Liao, Shizhu He, Yao Xu, Yuanzhe Zhang, Kang Liu, Shengping Liu, Jun Zhao*

Retrieval-Augmented-Generation and Gener-ation-Augmented-Generation have been proposed to enhance the knowledge required for question answering over Large Language Models (LLMs). However, the former depends on external resources, and both require incorporating the explicit documents into the context, which results in longer contexts that lead to more resource consumption. Recent works indicate that LLMs have modeled rich knowledge, albeit not effectively triggered or activated. Inspired by this, we propose a novel knowledge-augmented framework, Imagination-Augmented-Generation (IAG), which simulates the human capacity to compensate for knowledge deficits while answering questions solely through imagination, without relying on external resources. Guided by IAG, we propose an imagine richer context method for question answering (IMcQA), which obtains richer context through the following two modules: explicit imagination by generating a short dummy document with long context compress and implicit imagination with HyperNetwork for generating adapter weights. Experimental results on three datasets demonstrate that IMcQA exhibits significant advantages in both open-domain and closed-book settings, as well as in both in-distribution performance and out-of-distribution generalizations. Our code will be available at https://github.com/Xnhyacinth/IAG.

link: http://arxiv.org/abs/2403.15268v1

### WSCLoc: Weakly-Supervised Sparse-View Camera Relocalization

*Jialu Wang, Kaichen Zhou, Andrew Markham, Niki Trigoni*

Despite the advancements in deep learning for camera relocalization tasks, obtaining ground truth pose labels required for the training process remains a costly endeavor. While current weakly supervised methods excel in lightweight label generation, their performance notably declines in scenarios with sparse views. In response to this challenge, we introduce WSCLoc, a system capable of being customized to various deep learning-based relocalization models to enhance their performance under weakly-supervised and sparse view conditions. This is realized with two stages. In the initial stage, WSCLoc employs a multilayer perceptron-based structure called WFT-NeRF to co-optimize image reconstruction quality and initial pose information. To ensure a stable learning process, we incorporate temporal information as input. Furthermore, instead of optimizing SE(3), we opt for $\mathfrak{sim}(3)$ optimization to explicitly enforce a scale constraint. In the second stage, we co-optimize the pre-trained WFT-NeRF and WFT-Pose. This optimization is enhanced by Time-Encoding based Random View Synthesis and supervised by inter-frame geometric constraints that consider pose, depth, and RGB information. We validate our approaches on two publicly available datasets, one outdoor and one indoor. Our experimental results demonstrate that our weakly-supervised relocalization solutions achieve superior pose estimation accuracy in sparse-view scenarios, comparable to state-of-the-art camera relocalization methods. We will make our code publicly available.

link: http://arxiv.org/abs/2403.15272v1

### Event Temporal Relation Extraction based on Retrieval-Augmented on LLMs

*Xiaobin Zhang, Liangjun Zang, Qianwen Liu, Shuchong Wei, Songlin Hu*

Event temporal relation (TempRel) is a primary subject of the event relation extraction task. However, the inherent ambiguity of TempRel increases the difficulty of the task. With the rise of prompt engineering, it is important to design effective prompt templates and verbalizers to extract relevant knowledge. The traditional manually designed templates struggle to extract precise temporal knowledge. This paper introduces a novel retrieval-augmented TempRel extraction approach, leveraging knowledge retrieved from large language models (LLMs) to enhance prompt templates and verbalizers. Our method capitalizes on the diverse capabilities of various LLMs to generate a wide array of ideas for template and verbalizer design. Our proposed method fully exploits the potential of LLMs for generation tasks and contributes more knowledge to our design. Empirical evaluations across three widely recognized datasets demonstrate the efficacy of our method in improving the performance of event temporal relation extraction tasks.

link: http://arxiv.org/abs/2403.15273v1

### Specifying Genericity through Inclusiveness and Abstractness Continuous Scales

*Claudia Collacciani, Andrea Amelio Ravelli, Marianna Marcella Bolognesi*

This paper introduces a novel annotation framework for the fine-grained modeling of Noun Phrases' (NPs) genericity in natural language. The framework is designed to be simple and intuitive, making it accessible to non-expert annotators and suitable for crowd-sourced tasks. Drawing from theoretical and cognitive literature on genericity, this framework is grounded in established linguistic theory. Through a pilot study, we created a small but crucial annotated dataset of 324 sentences, serving as a foundation for future research. To validate our approach, we conducted an evaluation comparing our continuous annotations with existing binary annotations on the same dataset, demonstrating the framework's effectiveness in capturing nuanced aspects of genericity. Our work offers a practical resource for linguists, providing a first annotated dataset and an annotation scheme designed to build real-language datasets that can be used in studies on the semantics of genericity, and NLP practitioners, contributing to the development of commonsense knowledge repositories valuable in enhancing various NLP applications.

link: http://arxiv.org/abs/2403.15278v1

### Fundus: A Simple-to-Use News Scraper Optimized for High Quality Extractions

*Max Dallabetta, Conrad Dobberstein, Adrian Breiding, Alan Akbik*

This paper introduces Fundus, a user-friendly news scraper that enables users to obtain millions of high-quality news articles with just a few lines of code. Unlike existing news scrapers, we use manually crafted, bespoke content extractors that are specifically tailored to the formatting guidelines of each supported online newspaper. This allows us to optimize our scraping for quality such that retrieved news articles are textually complete and without HTML artifacts. Further, our framework combines both crawling (retrieving HTML from the web or large web archives) and content extraction into a single pipeline. By providing a unified interface for a predefined collection of newspapers, we aim to make Fundus broadly usable even for non-technical users. This paper gives an overview of the framework, discusses our design choices, and presents a comparative evaluation against other popular news scrapers. Our evaluation shows that Fundus yields significantly higher quality extractions (complete and artifact-free news articles) than prior work. The framework is available on GitHub under https://github.com/flairNLP/fundus and can be simply installed using pip.

link: http://arxiv.org/abs/2403.15279v1

### Blockchain-based Pseudonym Management for Vehicle Twin Migrations in Vehicular Edge Metaverse

*Jiawen Kang, Xiaofeng Luo, Jiangtian Nie, Tianhao Wu, Haibo Zhou, Yonghua Wang, Dusit Niyato, Shiwen Mao, Shengli Xie*

Driven by the great advances in metaverse and edge computing technologies, vehicular edge metaverses are expected to disrupt the current paradigm of intelligent transportation systems. As highly computerized avatars of Vehicular Metaverse Users (VMUs), the Vehicle Twins (VTs) deployed in edge servers can provide valuable metaverse services to improve driving safety and on-board satisfaction for their VMUs throughout journeys. To maintain uninterrupted metaverse experiences, VTs must be migrated among edge servers following the movements of vehicles. This can raise concerns about privacy breaches during the dynamic communications among vehicular edge metaverses. To address these concerns and safeguard location privacy, pseudonyms as temporary identifiers can be leveraged by both VMUs and VTs to realize anonymous communications in the physical space and virtual spaces. However, existing pseudonym management methods fall short in meeting the extensive pseudonym demands in vehicular edge metaverses, thus dramatically diminishing the performance of privacy preservation. To this end, we present a cross-metaverse empowered dual pseudonym management framework. We utilize cross-chain technology to enhance management efficiency and data security for pseudonyms. Furthermore, we propose a metric to assess the privacy level and employ a Multi-Agent Deep Reinforcement Learning (MADRL) approach to obtain an optimal pseudonym generating strategy. Numerical results demonstrate that our proposed schemes are high-efficiency and cost-effective, showcasing their promising applications in vehicular edge metaverses.

link: http://arxiv.org/abs/2403.15285v1

## Sphere Neural-Networks for Rational Reasoning

*Tiansi Dong, Mateja Jamnik, Pietro Liò*

The success of Large Language Models (LLMs), e.g., ChatGPT, is witnessed by their planetary popularity, their capability of human-like question-answering, and also by their steadily improved reasoning performance. However, it remains unclear whether LLMs reason. It is an open problem how traditional neural networks can be qualitatively extended to go beyond the statistic paradigm and achieve high-level cognition. Here, we present a minimalist qualitative extension by generalising computational building blocks from vectors to spheres. We propose Sphere Neural Networks (SphNNs) for human-like reasoning through model construction and inspection, and develop SphNN for syllogistic reasoning, a microcosm of human rationality. Instead of training data, SphNN uses a neuro-symbolic transition map of neighbourhood spatial relations to guide transformations from the current sphere configuration towards the target. SphNN is the first neural model that can determine the validity of long-chained syllogistic reasoning in one epoch by constructing sphere configurations as Euler diagrams, with the worst computational complexity of O(N^2). SphNN can evolve into various types of reasoning, such as spatio-temporal reasoning, logical reasoning with negation and disjunction, event reasoning, neuro-symbolic reasoning, and humour understanding (the highest level of cognition). All these suggest a new kind of Herbert A. Simon's scissors with two neural blades. SphNNs will tremendously enhance interdisciplinary collaborations to develop the two neural blades and realise deterministic neural reasoning and human-bounded rationality and elevate LLMs to reliable psychological AI. This work suggests that the non-zero radii of spheres are the missing components that prevent traditional deep-learning systems from reaching the realm of rational reasoning and cause LLMs to be trapped in the swamp of hallucination.

link: http://arxiv.org/abs/2403.15297v1

## Planning with a Learned Policy Basis to Optimally Solve Complex Tasks

*Guillermo Infante, David Kuric, Anders Jonsson, Vicenç Gómez, Herke van Hoof*

Conventional reinforcement learning (RL) methods can successfully solve a wide range of sequential decision problems. However, learning policies that can generalize predictably across multiple tasks in a setting with non-Markovian reward specifications is a challenging problem. We propose to use successor features to learn a policy basis so that each (sub)policy in it solves a well-defined subproblem. In a task described by a finite state automaton (FSA) that involves the same set of subproblems, the combination of these (sub)policies can then be used to generate an optimal solution without additional learning. In contrast to other methods that combine (sub)policies via planning, our method asymptotically attains global optimality, even in stochastic environments.

link: http://arxiv.org/abs/2403.15301v1

### Network Calculus Characterization of Congestion Control for Time-Varying Traffic

*Harvinder Lehal, Natchanon Luangsomboon, Jörg Liebeherr*

Models for the dynamics of congestion control generally involve systems of coupled differential equations. Universally, these models assume that traffic sources saturate the maximum transmissions allowed by the congestion control method. This is not suitable for studying congestion control of intermittent but bursty traffic sources. In this paper, we present a characterization of congestion control for arbitrary time-varying traffic that applies to rate-based as well as window-based congestion control. We leverage the capability of network calculus to precisely describe the input-output relationship at network elements for arbitrary source traffic. We show that our characterization can closely track the dynamics of even complex congestion control algorithms.

link: http://arxiv.org/abs/2403.15303v1

### KTbench: A Novel Data Leakage-Free Framework for Knowledge Tracing

*Yahya Badran, Christine Preisach*

Knowledge Tracing (KT) is concerned with predicting students' future performance on learning items in intelligent tutoring systems. Learning items are tagged with skill labels called knowledge concepts (KCs). Many KT models expand the sequence of item-student interactions into KC-student interactions by replacing learning items with their constituting KCs. This often results in a longer sequence length. This approach addresses the issue of sparse item-student interactions and minimises model parameters. However, two problems have been identified with such models. The first problem is the model's ability to learn correlations between KCs belonging to the same item, which can result in the leakage of ground truth labels and hinder performance. This problem can lead to a significant decrease in performance on datasets with a higher number of KCs per item. The second problem is that the available benchmark implementations ignore accounting for changes in sequence length when expanding KCs, leading to different models being tested with varying sequence lengths but still compared against the same benchmark. To address these problems, we introduce a general masking framework that mitigates the first problem and enhances the performance of such KT models while preserving the original model architecture without significant alterations. Additionally, we introduce KTbench, an open-source benchmark library designed to ensure the reproducibility of this work while mitigating the second problem.

link: http://arxiv.org/abs/2403.15304v1

### Controlled Training Data Generation with Diffusion Models

*Teresa Yeo, Andrei Atanov, Harold Benoit, Aleksandr Alekseev, Ruchira Ray, Pooya Esmaeil Akhoondi, Amir Zamir*

In this work, we present a method to control a text-to-image generative model to produce training data specifically "useful" for supervised learning. Unlike previous works that employ an open-loop approach and pre-define prompts to generate new data using either a language model or human expertise, we develop an automated closed-loop system which involves two feedback mechanisms. The first mechanism uses feedback from a given supervised model and finds adversarial prompts that result in image generations that maximize the model loss. While these adversarial prompts result in diverse data informed by the model, they are not informed of the target distribution, which can be inefficient. Therefore, we introduce the second feedback mechanism that guides the generation process towards a certain target distribution. We call the method combining these two mechanisms Guided Adversarial Prompts. We perform our evaluations on different tasks, datasets and architectures, with different types of distribution shifts (spuriously correlated data, unseen domains) and demonstrate the efficiency of the proposed feedback mechanisms compared to open-loop approaches.

link: http://arxiv.org/abs/2403.15309v1

### A Wasserstein perspective of Vanilla GANs

*Lea Kunkel, Mathias Trabs*

The empirical success of Generative Adversarial Networks (GANs) caused an increasing interest in theoretical research. The statistical literature is mainly focused on Wasserstein GANs and generalizations thereof, which especially allow for good dimension reduction properties. Statistical results for Vanilla GANs, the original optimization problem, are still rather limited and require assumptions such as smooth activation functions and equal dimensions of the latent space and the ambient space. To bridge this gap, we draw a connection from Vanilla GANs to the Wasserstein distance. By doing so, existing results for Wasserstein GANs can be extended to Vanilla GANs. In particular, we obtain an oracle inequality for Vanilla GANs in Wasserstein distance. The assumptions of this oracle inequality are designed to be satisfied by network architectures commonly used in practice, such as feedforward ReLU networks. By providing a quantitative result for the approximation of a Lipschitz function by a feedforward ReLU network with bounded H\"older norm, we conclude a rate of convergence for Vanilla GANs as well as Wasserstein GANs as estimators of the unknown probability distribution.

link: http://arxiv.org/abs/2403.15312v1