Wed 2024.05.15

Point Resampling and Ray Transformation Aid to Editable NeRF Models

Zhenyang Li, Zilong Chen, Feifan Qu, Mingqing Wang, Yizhou Zhao, Kai Zhang, Yifan Peng

In NeRF-aided editing tasks, object movement presents difficulties in supervision generation due to the introduction of variability in object positions. Moreover, the removal operations of certain scene objects often lead to empty regions, presenting challenges for NeRF models in inpainting them effectively. We propose an implicit ray transformation strategy, allowing for direct manipulation of the 3D object's pose by operating on the neural-point in NeRF rays. To address the challenge of inpainting potential empty regions, we present a plug-and-play inpainting module, dubbed differentiable neural-point resampling (DNR), which interpolates those regions in 3D space at the original ray locations within the implicit space, thereby facilitating object removal & scene inpainting tasks. Importantly, employing DNR effectively narrows the gap between ground truth and predicted implicit features, potentially increasing the mutual information (MI) of the features across rays. Then, we leverage DNR and ray transformation to construct a point-based editable NeRF pipeline PR^2T-NeRF. Results primarily evaluated on 3D object removal & inpainting tasks indicate that our pipeline achieves state-of-the-art performance. In addition, our pipeline supports high-quality rendering visualization for diverse editing operations without necessitating extra supervision.

link: http://arxiv.org/abs/2405.07306v1

DiffGen: Robot Demonstration Generation via Differentiable Physics Simulation, Differentiable Rendering, and Vision-Language Model

Yang Jin, Jun Lv, Shuqiang Jiang, Cewu Lu

Generating robot demonstrations through simulation is widely recognized as an effective way to scale up robot data. Previous work often trained reinforcement learning agents to generate expert policies, but this approach lacks sample efficiency. Recently, a line of work has attempted to generate robot demonstrations via differentiable simulation, which is promising but heavily relies on reward design, a labor-intensive process. In this paper, we propose DiffGen, a novel framework that integrates differentiable physics simulation, differentiable rendering, and a vision-language model to enable automatic and efficient generation of robot demonstrations. Given a simulated robot manipulation scenario and a natural language instruction, DiffGen can generate realistic robot demonstrations by minimizing the distance between the embedding of the language instruction and the embedding of the simulated observation after manipulation. The embeddings are obtained from the vision-language model, and the optimization is achieved by calculating and descending gradients through the differentiable simulation, differentiable rendering, and vision-language model components, thereby accomplishing the specified task. Experiments demonstrate that with DiffGen, we could efficiently and effectively generate robot data with minimal human effort or training time.

link: http://arxiv.org/abs/2405.07309v1

Nonparametric Control-Koopman Operator Learning: Flexible and Scalable Models for Prediction and Control

Petar Bevanda, Bas Driessen, Lucian Cristian Iacob, Roland Toth, Stefan Sosnowski, Sandra Hirche

Linearity of Koopman operators and simplicity of their estimators coupled with model-reduction capabilities has lead to their great popularity in applications for learning dynamical systems. While nonparametric Koopman operator learning in infinite-dimensional reproducing kernel Hilbert spaces is well understood for autonomous systems, its control system analogues are largely unexplored. Addressing systems with control inputs in a principled manner is crucial for fully data-driven learning of controllers, especially since existing approaches commonly resort to representational heuristics or parametric models of limited expressiveness and scalability. We address the aforementioned challenge by proposing a universal framework via control-affine reproducing kernels that enables direct estimation of a single operator even for control systems. The proposed approach, called control-Koopman operator regression (cKOR), is thus completely analogous to Koopman operator regression of the autonomous case. First in the literature, we present a nonparametric framework for learning Koopman operator representations of nonlinear control-affine systems that does not suffer from the curse of control input dimensionality. This allows for reformulating the infinite-dimensional learning problem in a finite-dimensional space based solely on data without apriori loss of precision due to a restriction to a finite span of functions or inputs as in other approaches. For enabling applications to large-scale control systems, we also enhance the scalability of control-Koopman operator estimators by leveraging random projections (sketching). The efficacy of our novel cKOR approach is demonstrated on both forecasting and control tasks.

link: http://arxiv.org/abs/2405.07312v1

VALID: a Validated Algorithm for Learning in Decentralized Networks with Possible Adversarial Presence

Mayank Bakshi, Sara Ghasvarianjahromi, Yauhen Yakimenka, Allison Beemer, Oliver Kosut, Joerg Kliewer

We introduce the paradigm of validated decentralized learning for undirected networks with heterogeneous data and possible adversarial infiltration. We require (a) convergence to a global empirical loss minimizer when adversaries are absent, and (b) either detection of adversarial presence of convergence to an admissible consensus irrespective of the adversarial configuration. To this end, we propose the VALID protocol which, to the best of our knowledge, is the first to achieve a validated learning guarantee. Moreover, VALID offers an O(1/T) convergence rate (under pertinent regularity assumptions), and computational and communication complexities comparable to non-adversarial distributed stochastic gradient descent. Remarkably, VALID retains optimal performance metrics in adversary-free environments, sidestepping the robustness penalties observed in prior byzantine-robust methods. A distinctive aspect of our study is a heterogeneity metric based on the norms of individual agents' gradients computed at the global empirical loss minimizer. This not only provides a natural statistic for detecting significant byzantine disruptions but also allows us to prove the optimality of VALID in wide generality. Lastly, our numerical results reveal that, in the absence of adversaries, VALID converges faster than state-of-the-art byzantine robust algorithms, while when adversaries are present, VALID terminates with each honest either converging to an admissible consensus of declaring adversarial presence in the network.

link: http://arxiv.org/abs/2405.07316v1

Machine Unlearning in Contrastive Learning

Zixin Wang, Kongyang Chen

Machine unlearning is a complex process that necessitates the model to diminish the influence of the training data while keeping the loss of accuracy to a minimum. Despite the numerous studies on machine unlearning in recent years, the majority of them have primarily focused on supervised learning models, leaving research on contrastive learning models relatively underexplored. With the conviction that self-supervised learning harbors a promising potential, surpassing or rivaling that of supervised learning, we set out to investigate methods for machine unlearning centered around contrastive learning models. In this study, we introduce a novel gradient constraint-based approach for training the model to effectively achieve machine unlearning. Our method only necessitates a minimal number of training epochs and the identification of the data slated for unlearning. Remarkably, our approach demonstrates proficient performance not only on contrastive learning models but also on supervised learning models, showcasing its versatility and adaptability in various learning paradigms.

link: http://arxiv.org/abs/2405.07317v1

LayGA: Layered Gaussian Avatars for Animatable Clothing Transfer

Siyou Lin, Zhe Li, Zhaoqi Su, Zerong Zheng, Hongwen Zhang, Yebin Liu

Animatable clothing transfer, aiming at dressing and animating garments across characters, is a challenging problem. Most human avatar works entangle the representations of the human body and clothing together, which leads to difficulties for virtual try-on across identities. What's worse, the entangled representations usually fail to exactly track the sliding motion of garments. To overcome these limitations, we present Layered Gaussian Avatars (LayGA), a new representation that formulates body and clothing as two separate layers for photorealistic animatable clothing transfer from multi-view videos. Our representation is built upon the Gaussian map-based avatar for its excellent representation power of garment details. However, the Gaussian map produces unstructured 3D Gaussians distributed around the actual surface. The absence of a smooth explicit surface raises challenges in accurate garment tracking and collision handling between body and garments. Therefore, we propose two-stage training involving single-layer reconstruction and multi-layer fitting. In the single-layer reconstruction stage, we propose a series of geometric constraints to reconstruct smooth surfaces and simultaneously obtain the segmentation between body and clothing. Next, in the multi-layer fitting stage, we train two separate models to represent body and clothing and utilize the reconstructed clothing geometries as 3D supervision for more accurate garment tracking. Furthermore, we propose geometry and rendering layers for both high-quality geometric reconstruction and high-fidelity rendering. Overall, the proposed LayGA realizes photorealistic animations and virtual try-on, and outperforms other baseline methods. Our project page is https://jsnln.github.io/layga/index.html.

link: http://arxiv.org/abs/2405.07319v1

L(u)PIN: LLM-based Political Ideology Nowcasting

Ken Kato, Annabelle Purnomo, Christopher Cochrane, Raeid Saqur

The quantitative analysis of political ideological positions is a difficult task. In the past, various literature focused on parliamentary voting data of politicians, party manifestos and parliamentary speech to estimate political disagreement and polarization in various political systems. However previous methods of quantitative political analysis suffered from a common challenge which was the amount of data available for analysis. Also previous methods frequently focused on a more general analysis of politics such as overall polarization of the parliament or party-wide political ideological positions. In this paper, we present a method to analyze ideological positions of individual parliamentary representatives by leveraging the latent knowledge of LLMs. The method allows us to evaluate the stance of politicians on an axis of our choice allowing us to flexibly measure the stance of politicians in regards to a topic/controversy of our choice. We achieve this by using a fine-tuned BERT classifier to extract the opinion-based sentences from the speeches of representatives and projecting the average BERT embeddings for each representative on a pair of reference seeds. These reference seeds are either manually chosen representatives known to have opposing views on a particular topic or they are generated sentences which where created using the GPT-4 model of OpenAI. We created the sentences by prompting the GPT-4 model to generate a speech that would come from a politician defending a particular position.

link: http://arxiv.org/abs/2405.07320v1

QACM: QoS-Aware xApp Conflict Mitigation in Open RAN

Abdul Wadud, Fatemeh Golpayegani, Nima Afraz

The advent of Open Radio Access Network (RAN) has revolutionized the field of RAN by introducing elements of native support of intelligence and openness into the next generation of mobile network infrastructure. Open RAN paves the way for standardized interfaces and enables the integration of network applications from diverse vendors, thereby enhancing network management flexibility. However, control decision conflicts occur when components from different vendors are deployed together. This article provides an overview of various types of conflicts that may occur in Open RAN, with a particular focus on intra-component conflict mitigation among Extended Applications (xApps) in the Near Real Time RAN Intelligent Controller (Near-RT-RIC). A QoS-Aware Conflict Mitigation (QACM) method is proposed that finds the optimal configuration of conflicting parameters while maximizing the number of xApps that have their Quality of Service (QoS) requirements met. We compare the performance of the proposed QACM method with two benchmark methods for priority and non-priority cases. The results indicate that our proposed method is the most effective in maintaining QoS requirements for conflicting xApps.

link: http://arxiv.org/abs/2405.07324v1

Power Evaluation of IOT Application Layer Protocols

Amirhossein Shahrokhi, Mahmood Ahmadi

The Internet of Things has affected all aspects of daily life, and the number of IoT devices is increasing day by day. According to forecasts, the number of Internet of Things devices will reach one trillion devices by 2035. The increase in the number of devices connected to the Internet will cause various concerns. One of the most important concerns is the energy and power consumption of these devices. Although Internet of Things modules are low in energy consumption, their widespread and large-scale use has made the issue of power consumption become the most important challenge in this field. For this reason, it is necessary to use communication protocols that, in addition to establishing efficient communication, impose minimal power consumption on the network. In this paper, application layer protocols such as MQTT, MQTT-SN, CoAP, and HTTP are simulated using the tools available in the Contiki operating system, including COOJA and Powertrace, and they { are evaluated} and compared with each other in terms of power consumption. According to the simulations performed by the mentioned tools, the MQTT-SN protocol was the least consuming protocol in terms of power consumption. After that, the CoAP protocol is placed, and with a slight difference, the MQTT protocol, which consumes more than MQTT-SN. Finally, the HTTP protocol consumes the most power, which makes it unsuitable for communication in the Internet of Things

link: http://arxiv.org/abs/2405.07326v1

Liquid Ensemble Selection for Continual Learning

Carter Blair, Ben Armstrong, Kate Larson

Continual learning aims to enable machine learning models to continually learn from a shifting data distribution without forgetting what has already been learned. Such shifting distributions can be broken into disjoint subsets of related examples; by training each member of an ensemble on a different subset it is possible for the ensemble as a whole to achieve much higher accuracy with less forgetting than a naive model. We address the problem of selecting which models within an ensemble should learn on any given data, and which should predict. By drawing on work from delegative voting we develop an algorithm for using delegation to dynamically select which models in an ensemble are active. We explore a variety of delegation methods and performance metrics, ultimately finding that delegation is able to provide a significant performance boost over naive learning in the face of distribution shifts.

link: http://arxiv.org/abs/2405.07327v1

Stochastic Bandits with ReLU Neural Networks

Kan Xu, Hamsa Bastani, Surbhi Goel, Osbert Bastani

We study the stochastic bandit problem with ReLU neural network structure. We show that a $\tilde{O}(\sqrt{T})$ regret guarantee is achievable by considering bandits with one-layer ReLU neural networks; to the best of our knowledge, our work is the first to achieve such a guarantee. In this specific setting, we propose an OFU-ReLU algorithm that can achieve this upper bound. The algorithm first explores randomly until it reaches a linear regime, and then implements a UCB-type linear bandit algorithm to balance exploration and exploitation. Our key insight is that we can exploit the piecewise linear structure of ReLU activations and convert the problem into a linear bandit in a transformed feature space, once we learn the parameters of ReLU relatively accurately during the exploration stage. To remove dependence on model parameters, we design an OFU-ReLU+ algorithm based on a batching strategy, which can provide the same theoretical guarantee.

link: http://arxiv.org/abs/2405.07331v1

PotatoGANs: Utilizing Generative Adversarial Networks, Instance Segmentation, and Explainable AI for Enhanced Potato Disease Identification and Classification

Mohammad Shafiul Alam, Fatema Tuj Johora Faria, Mukaffi Bin Moin, Ahmed Al Wase, Md. Rabius Sani, Khan Md Hasib

Numerous applications have resulted from the automation of agricultural disease segmentation using deep learning techniques. However, when applied to new conditions, these applications frequently face the difficulty of overfitting, resulting in lower segmentation performance. In the context of potato farming, where diseases have a large influence on yields, it is critical for the agricultural economy to quickly and properly identify these diseases. Traditional data augmentation approaches, such as rotation, flip, and translation, have limitations and frequently fail to provide strong generalization results. To address these issues, our research employs a novel approach termed as PotatoGANs. In this novel data augmentation approach, two types of Generative Adversarial Networks (GANs) are utilized to generate synthetic potato disease images from healthy potato images. This approach not only expands the dataset but also adds variety, which helps to enhance model generalization. Using the Inception score as a measure, our experiments show the better quality and realisticness of the images created by PotatoGANs, emphasizing their capacity to resemble real disease images closely. The CycleGAN model outperforms the Pix2Pix GAN model in terms of image quality, as evidenced by its higher IS scores CycleGAN achieves higher Inception scores (IS) of 1.2001 and 1.0900 for black scurf and common scab, respectively. This synthetic data can significantly improve the training of large neural networks. It also reduces data collection costs while enhancing data diversity and generalization capabilities. Our work improves interpretability by combining three gradient-based Explainable AI algorithms (GradCAM, GradCAM++, and ScoreCAM) with three distinct CNN architectures (DenseNet169, Resnet152 V2, InceptionResNet V2) for potato disease classification.

link: http://arxiv.org/abs/2405.07332v1

Quantum Mini-Apps: A Framework for Developing and Benchmarking Quantum-HPC Applications

Nishant Saurabh, Pradeep Mantha, Florian J. Kiwit, Shantenu Jha, Andre Luckow

With the increasing maturity and scale of quantum hardware and its integration into HPC systems, there is a need to develop robust techniques for developing, characterizing, and benchmarking quantum-HPC applications and middleware systems. This requires a better understanding of interaction, coupling, and common execution patterns between quantum and classical workload tasks and components. This paper identifies six quantum-HPC execution motifs - recurring execution patterns characterized by distinct coupling and interaction modes. These motifs provide the basis for a suite of quantum mini-apps - simplified application prototypes that encapsulate essential characteristics of production systems. To support these developments, we introduce a mini-app framework that offers the necessary abstractions for creating and executing mini-apps across heterogeneous quantum-HPC infrastructure, making it a valuable tool for performance characterizations and middleware development.

link: http://arxiv.org/abs/2405.07333v1

Data Trading Combination Auction Mechanism based on the Exponential Mechanism

Kongyang Chen, Zeming Xu, Bing Mi

With the widespread application of machine learning technology in recent years, the demand for training data has increased significantly, leading to the emergence of research areas such as data trading. The work in this field is still in the developmental stage. Different buyers have varying degrees of demand for various types of data, and auctions play a role in such scenarios due to their authenticity and fairness. Recent related work has proposed combination auction mechanisms for different domains. However, such mechanisms have not addressed the privacy concerns of buyers. In this paper, we design a \textit{Data Trading Combination Auction Mechanism based on the exponential mechanism} (DCAE) to protect buyers' bidding privacy from being leaked. We apply the exponential mechanism to select the final settlement price for the auction and generate a probability distribution based on the relationship between the price and the revenue. In the experimental aspect, we consider the selection of different mechanisms under two scenarios, and the experimental results show that this method can ensure high auction revenue and protect buyers' privacy from being violated.

link: http://arxiv.org/abs/2405.07336v1

Explainable Convolutional Neural Networks for Retinal Fundus Classification and Cutting-Edge Segmentation Models for Retinal Blood Vessels from Fundus Images

Fatema Tuj Johora Faria, Mukaffi Bin Moin, Pronay Debnath, Asif Iftekher Fahim, Faisal Muhammad Shah

Our research focuses on the critical field of early diagnosis of disease by examining retinal blood vessels in fundus images. While automatic segmentation of retinal blood vessels holds promise for early detection, accurate analysis remains challenging due to the limitations of existing methods, which often lack discrimination power and are susceptible to influences from pathological regions. Our research in fundus image analysis advances deep learning-based classification using eight pre-trained CNN models. To enhance interpretability, we utilize Explainable AI techniques such as Grad-CAM, Grad-CAM++, Score-CAM, Faster Score-CAM, and Layer CAM. These techniques illuminate the decision-making processes of the models, fostering transparency and trust in their predictions. Expanding our exploration, we investigate ten models, including TransUNet with ResNet backbones, Attention U-Net with DenseNet and ResNet backbones, and Swin-UNET. Incorporating diverse architectures such as ResNet50V2, ResNet101V2, ResNet152V2, and DenseNet121 among others, this comprehensive study deepens our insights into attention mechanisms for enhanced fundus image analysis. Among the evaluated models for fundus image classification, ResNet101 emerged with the highest accuracy, achieving an impressive 94.17%. On the other end of the spectrum, EfficientNetB0 exhibited the lowest accuracy among the models, achieving a score of 88.33%. Furthermore, in the domain of fundus image segmentation, Swin-Unet demonstrated a Mean Pixel Accuracy of 86.19%, showcasing its effectiveness in accurately delineating regions of interest within fundus images. Conversely, Attention U-Net with DenseNet201 backbone exhibited the lowest Mean Pixel Accuracy among the evaluated models, achieving a score of 75.87%.

link: http://arxiv.org/abs/2405.07338v1