Mon 2024.04.08

MarsSeg: Mars Surface Semantic Segmentation with Multi-level Extractor and Connector

Junbo Li, Keyan Chen, Gengju Tian, Lu Li, Zhenwei Shi

The segmentation and interpretation of the Martian surface play a pivotal role in Mars exploration, providing essential data for the trajectory planning and obstacle avoidance of rovers. However, the complex topography, similar surface features, and the lack of extensive annotated data pose significant challenges to the high-precision semantic segmentation of the Martian surface. To address these challenges, we propose a novel encoder-decoder based Mars segmentation network, termed MarsSeg. Specifically, we employ an encoder-decoder structure with a minimized number of down-sampling layers to preserve local details. To facilitate a high-level semantic understanding across the shadow multi-level feature maps, we introduce a feature enhancement connection layer situated between the encoder and decoder. This layer incorporates Mini Atrous Spatial Pyramid Pooling (Mini-ASPP), Polarized Self-Attention (PSA), and Strip Pyramid Pooling Module (SPPM). The Mini-ASPP and PSA are specifically designed for shadow feature enhancement, thereby enabling the expression of local details and small objects. Conversely, the SPPM is employed for deep feature enhancement, facilitating the extraction of high-level semantic category-related information. Experimental results derived from the Mars-Seg and AI4Mars datasets substantiate that the proposed MarsSeg outperforms other state-of-the-art methods in segmentation performance, validating the efficacy of each proposed component.

link: http://arxiv.org/abs/2404.04155v1

Noisy Label Processing for Classification: A Survey

Mengting Li, Chuang Zhu

In recent years, deep neural networks (DNNs) have gained remarkable achievement in computer vision tasks, and the success of DNNs often depends greatly on the richness of data. However, the acquisition process of data and high-quality ground truth requires a lot of manpower and money. In the long, tedious process of data annotation, annotators are prone to make mistakes, resulting in incorrect labels of images, i.e., noisy labels. The emergence of noisy labels is inevitable. Moreover, since research shows that DNNs can easily fit noisy labels, the existence of noisy labels will cause significant damage to the model training process. Therefore, it is crucial to combat noisy labels for computer vision tasks, especially for classification tasks. In this survey, we first comprehensively review the evolution of different deep learning approaches for noisy label combating in the image classification task. In addition, we also review different noise patterns that have been proposed to design robust algorithms. Furthermore, we explore the inner pattern of real-world label noise and propose an algorithm to generate a synthetic label noise pattern guided by real-world data. We test the algorithm on the well-known real-world dataset CIFAR-10N to form a new real-world data-guided synthetic benchmark and evaluate some typical noise-robust methods on the benchmark.

link: http://arxiv.org/abs/2404.04159v1

Wireless Resource Optimization in Hybrid Semantic/Bit Communication Networks

Le Xia, Yao Sun, Dusit Niyato, Lan Zhang, Muhammad Ali Imran

Recently, semantic communication (SemCom) has shown great potential in significant resource savings and efficient information exchanges, thus naturally introducing a novel and practical cellular network paradigm where two modes of SemCom and conventional bit communication (BitCom) coexist. Nevertheless, the involved wireless resource management becomes rather complicated and challenging, given the unique background knowledge matching and time-consuming semantic coding requirements in SemCom. To this end, this paper jointly investigates user association (UA), mode selection (MS), and bandwidth allocation (BA) problems in a hybrid semantic/bit communication network (HSB-Net). Concretely, we first identify a unified performance metric of

message throughput for both SemCom and BitCom links. Next, we specially develop a knowledge matching-aware two-stage tandem packet queuing model and theoretically derive the average packet loss ratio and queuing latency. Combined with practical constraints, we then formulate a joint optimization problem for UA, MS, and BA to maximize the overall message throughput of HSB-Net. Afterward, we propose an optimal resource management strategy by utilizing a Lagrange primal-dual transformation method and a preference list-based heuristic algorithm with polynomial-time complexity. Numerical results not only demonstrate the accuracy of our analytical queuing model, but also validate the performance superiority of our proposed strategy compared with different benchmarks.

link: http://arxiv.org/abs/2404.04162v1

Dwell in the Beginning: How Language Models Embed Long Documents for Dense Retrieval

João Coelho, Bruno Martins, João Magalhães, Jamie Callan, Chenyan Xiong

This study investigates the existence of positional biases in Transformer-based models for text representation learning, particularly in the context of web document retrieval. We build on previous research that demonstrated loss of information in the middle of input sequences for causal language models, extending it to the domain of representation learning. We examine positional biases at various stages of training for an encoder-decoder model, including language model pre-training, contrastive pre-training, and contrastive fine-tuning. Experiments with the MS-MARCO document collection reveal that after contrastive pre-training the model already generates embeddings that better capture early contents of the input, with fine-tuning further aggravating this effect.

link: http://arxiv.org/abs/2404.04163v1

Chinese Tiny LLM: Pretraining a Chinese-Centric Large Language Model

Xinrun Du, Zhouliang Yu, Songyang Gao, Ding Pan, Yuyang Cheng, Ziyang Ma, Ruibin Yuan, Xingwei Qu, Jiaheng Liu, Tianyu Zheng, Xinchen Luo, Guorui Zhou, Binhang Yuan, Wenhu Chen, Jie Fu, Ge Zhang

In this study, we introduce CT-LLM, a 2B large language model (LLM) that illustrates a pivotal shift towards prioritizing the Chinese language in developing LLMs. Uniquely initiated from scratch, CT-LLM diverges from the conventional methodology by primarily incorporating Chinese textual data, utilizing an extensive corpus of 1,200 billion tokens, including 800 billion Chinese tokens, 300 billion English tokens, and 100 billion code tokens. This strategic composition facilitates the model's exceptional proficiency in understanding and processing Chinese, a capability further enhanced through alignment techniques. Demonstrating remarkable performance on the CHC-Bench, CT-LLM excels in Chinese language tasks, and showcases its adeptness in English through SFT. This research challenges the prevailing paradigm of training LLMs predominantly on English corpora and then adapting them to other languages, broadening the horizons for LLM training methodologies. By open-sourcing the full process of training a Chinese LLM, including a detailed data processing procedure with the obtained Massive Appropriate Pretraining Chinese Corpus (MAP-CC), a well-chosen multidisciplinary Chinese Hard Case Benchmark (CHC-Bench), and the 2B-size Chinese Tiny LLM (CT-LLM), we aim to foster further exploration and innovation in both academia and industry, paving the way for more inclusive and versatile language models.

link: http://arxiv.org/abs/2404.04167v1

Do Sentence Transformers Learn Quasi-Geospatial Concepts from General Text?

Ilya Ilyankou, Aldo Lipani, Stefano Cavazzi, Xiaowei Gao, James Haworth

Sentence transformers are language models designed to perform semantic search. This study investigates the capacity of sentence transformers, fine-tuned on general question-answering datasets for asymmetric semantic search, to associate descriptions of human-generated routes across Great Britain with queries often used to describe hiking experiences. We find that sentence transformers have some zero-shot capabilities to understand quasi-geospatial concepts, such as

route types and difficulty, suggesting their potential utility for routing recommendation systems.

link: http://arxiv.org/abs/2404.04169v1

H3DFact: Heterogeneous 3D Integrated CIM for Factorization with Holographic Perceptual Representations

Zishen Wan, Che-Kai Liu, Mohamed Ibrahim, Hanchen Yang, Samuel Spetalnick, Tushar Krishna, Arijit Raychowdhury

Disentangling attributes of various sensory signals is central to human-like perception and reasoning and a critical task for higher-order cognitive and neuro-symbolic AI systems. An elegant approach to represent this intricate factorization is via high-dimensional holographic vectors drawing on brain-inspired vector symbolic architectures. However, holographic factorization involves iterative computation with high-dimensional matrix-vector multiplications and suffers from non-convergence problems. In this paper, we present H3DFact, a heterogeneous 3D integrated in-memory compute engine capable of efficiently factorizing high-dimensional holographic representations. H3DFact exploits the computation-in-superposition capability of holographic vectors and the intrinsic stochasticity associated with memristive-based 3D compute-in-memory. Evaluated on large-scale factorization and perceptual problems, H3DFact demonstrates superior capability in factorization accuracy and operational capacity by up to five orders of magnitude, with 5.5x compute density, 1.2x energy efficiency improvements, and 5.9x less silicon footprint compared to iso-capacity 2D designs.

link: http://arxiv.org/abs/2404.04173v1

SCAResNet: A ResNet Variant Optimized for Tiny Object Detection in Transmission and Distribution Towers

Weile Li, Muqing Shi, Zhonghua Hong

Traditional deep learning-based object detection networks often resize images during the data preprocessing stage to achieve a uniform size and scale in the feature map. Resizing is done to facilitate model propagation and fully connected classification. However, resizing inevitably leads to object deformation and loss of valuable information in the images. This drawback becomes particularly pronounced for tiny objects like distribution towers with linear shapes and few pixels. To address this issue, we propose abandoning the resizing operation. Instead, we introduce Positional-Encoding Multi-head Criss-Cross Attention. This allows the model to capture contextual information and learn from multiple representation subspaces, effectively enriching the semantics of distribution towers. Additionally, we enhance Spatial Pyramid Pooling by reshaping three pooled feature maps into a new unified one while also reducing the computational burden. This approach allows images of different sizes and scales to generate feature maps with uniform dimensions and can be employed in feature map propagation. Our SCAResNet incorporates these aforementioned improvements into the backbone network ResNet. We evaluated our SCAResNet using the Electric Transmission and Distribution Infrastructure Imagery dataset from Duke University. Without any additional tricks, we employed various object detection models with Gaussian Receptive Field based Label Assignment as the baseline. When incorporating the SCAResNet into the baseline model, we achieved a 2.1% improvement in mAPs. This demonstrates the advantages of our SCAResNet in detecting transmission and distribution towers and its value in tiny object detection. The source code is available at https://github.com/LisavilaLee/SCAResNet_mmdet.

link: http://dx.doi.org/10.1109/LGRS.2023.3315376

RACS and SADL: Towards Robust SMR in the Wide-Area Network

Pasindu Tennage, Antoine Desjardins, Lefteris Kokoris-Kogias

Consensus algorithms deployed in the crash fault tolerant setting chose a leader-based architecture in order to achieve the lowest latency possible. However, when deployed in the wide area they face two key robustness challenges. First, they lose liveness when the network is unreliable because they rely on timeouts to find a leader. Second, they cannot have a high replication factor because of the high load imposed on the leader-replica making it a bottleneck. This effectively limits the

replication factor allowed, for a given level of throughput, thus lowering the fault tolerance threshold. In this paper, we propose RACS and SADL, a modular state machine replication algorithm that addresses these two robustness challenges. To achieve robustness under adversarial network conditions, we propose RACS, a novel crash fault-tolerant consensus algorithm. RACS consists of two modes of operations: synchronous and asynchronous, that always ensure liveness. RACS leverages the synchronous network to minimize the communication cost to O(n) and matches the lower bound of O(n2) at adversarial-case executions. To avoid the leader bottleneck and to allow higher replication factor, without sacrificing the throughput, we then propose SADL, a novel consensus-agnostic asynchronous dissemination layer. SADL separates client command dissemination from the critical path of consensus and distributes the overhead evenly among all the replicas. The combination of RACS and SADL (SADL-RACS) provides a robust and high-performing state machine replication system. We implement and evaluate RACS and SADL-RACS in a wide-area deployment running on Amazon EC2.

link: http://arxiv.org/abs/2404.04183v1

Reliable Feature Selection for Adversarially Robust Cyber-Attack Detection

João Vitorino, Miguel Silva, Eva Maia, Isabel Praça

The growing cybersecurity threats make it essential to use high-quality data to train Machine Learning (ML) models for network traffic analysis, without noisy or missing data. By selecting the most relevant features for cyber-attack detection, it is possible to improve both the robustness and computational efficiency of the models used in a cybersecurity system. This work presents a feature selection and consensus process that combines multiple methods and applies them to several network datasets. Two different feature sets were selected and were used to train multiple ML models with regular and adversarial training. Finally, an adversarial evasion robustness benchmark was performed to analyze the reliability of the different feature sets and their impact on the susceptibility of the models to adversarial examples. By using an improved dataset with more data diversity, selecting the best time-related features and a more specific feature set, and performing adversarial training, the ML models were able to achieve a better adversarially robust generalization. The robustness of the models was significantly improved without their generalization to regular traffic flows being affected, without increases of false alarms, and without requiring too many computational resources, which enables a reliable detection of suspicious activity and perturbed traffic flows in enterprise computer networks.

link: http://arxiv.org/abs/2404.04188v1

Are We Up to the Challenge? An analysis of the FCC Broadband Data Collection Fixed Internet Availability Challenges

Jonatas Marques, Alexis Schrubbe, Nicole P. Marwell, Nick Feamster

In 2021, the Broadband Equity, Access, and Deployment (BEAD) program allocated $42.45 billion to enhance high-speed internet access across the United States. As part of this funding initiative, The Federal Communications Commission (FCC) developed a national coverage map to guide the allocation of BEAD funds. This map was the key determinant to direct BEAD investments to areas in need of broadband infrastructure improvements. The FCC encouraged public participation in refining this coverage map through the submission of "challenges" to either locations on the map or the status of broadband at any location on the map. These challenges allowed citizens and organizations to report discrepancies between the map's data and actual broadband availability, ensuring a more equitable distribution of funds. In this paper, we present a study analyzing the nature and distribution of these challenges across different access technologies and geographic areas. Among several other insights, we observe, for example, that the majority of challenges (about 58%) were submitted against terrestrial fixed wireless technologies as well as that the state of Nebraska had the strongest engagement in the challenge process with more than 75% of its broadband-serviceable locations having submitted at least one challenge.

link: http://arxiv.org/abs/2404.04189v1

Exploring Probabilistic Models for Semi-supervised Learning

Jianfeng Wang

This thesis studies advanced probabilistic models, including both their theoretical foundations and practical applications, for different semi-supervised learning (SSL) tasks. The proposed probabilistic methods are able to improve the safety of AI systems in real applications by providing reliable uncertainty estimates quickly, and at the same time, achieve competitive performance compared to their deterministic counterparts. The experimental results indicate that the methods proposed in the thesis have great value in safety-critical areas, such as the autonomous driving or medical imaging analysis domain, and pave the way for the future discovery of highly effective and efficient probabilistic approaches in the SSL sector.

link: http://arxiv.org/abs/2404.04199v1

V-Star: Learning Visibly Pushdown Grammars from Program Inputs

Xiaodong Jia, Gang Tan

Accurate description of program inputs remains a critical challenge in the field of programming languages. Active learning, as a well-established field, achieves exact learning for regular languages. We offer an innovative grammar inference tool, V-Star, based on the active learning of visibly pushdown automata. V-Star deduces nesting structures of program input languages from sample inputs, employing a novel inference mechanism based on nested patterns. This mechanism identifies token boundaries and converts languages such as XML documents into VPLs. We then adapted Angluin's L-Star, an exact learning algorithm, for VPA learning, which improves the precision of our tool. Our evaluation demonstrates that V-Star effectively and efficiently learns a variety of practical grammars, including S-Expressions, JSON, and XML, and outperforms other state-of-the-art tools.

link: http://dx.doi.org/10.1145/3656458

Deep-learning Segmentation of Small Volumes in CT images for Radiotherapy Treatment Planning

Jianxin Zhou, Kadishe Fejza, Massimiliano Salvatori, Daniele Della Latta, Gregory M. Hermann, Angela Di Fulvio

Our understanding of organs at risk is progressing to include physical small tissues such as coronary arteries and the radiosensitivities of many small organs and tissues are high. Therefore, the accurate segmentation of small volumes in external radiotherapy is crucial to protect them from over-irradiation. Moreover, with the development of the particle therapy and on-board imaging, the treatment becomes more accurate and precise. The purpose of this work is to optimize organ segmentation algorithms for small organs. We used 50 three-dimensional (3-D) computed tomography (CT) head and neck images from StructSeg2019 challenge to develop a general-purpose V-Net model to segment 20 organs in the head and neck region. We applied specific strategies to improve the segmentation accuracy of the small volumes in this anatomical region, i.e., the lens of the eye. Then, we used 17 additional head images from OSF healthcare to validate the robustness of the V Net model optimized for small-volume segmentation. With the study of the StructSeg2019 images, we found that the optimization of the image normalization range and classification threshold yielded a segmentation improvement of the lens of the eye of approximately 50%, compared to the use of the V-Net not optimized for small volumes. We used the optimized model to segment 17 images acquired using heterogeneous protocols. We obtained comparable Dice coefficient values for the clinical and StructSeg2019 images (0.61 plus/minus 0.07 and 0.58 plus/minus 0.10 for the left and right lens of the eye, respectively)

link: http://arxiv.org/abs/2404.04202v1

Social Skill Training with Large Language Models

Diyi Yang, Caleb Ziems, William Held, Omar Shaikh, Michael S. Bernstein, John Mitchell

People rely on social skills like conflict resolution to communicate effectively and to thrive in both work and personal life. However, practice environments for social skills are typically out of reach for most people. How can we make social skill training more available, accessible, and inviting? Drawing upon interdisciplinary research from communication and psychology, this perspective paper identifies social skill barriers to enter specialized fields. Then we present a solution that leverages large language models for social skill training via a generic framework. Our AI Partner, AI Mentor framework merges experiential learning with realistic practice and tailored feedback. This work ultimately calls for cross-disciplinary innovation to address the broader implications for workforce development and social equality.

link: http://arxiv.org/abs/2404.04204v1