Wed 2024.03.20

Exploring Distance Query Processing in Edge Computing Environments

Xiubo Zhang, Yujie He, Ye Li, Yan Li, Zijie Zhou, Dongyao Wei, Ryan

In the context of changing travel behaviors and the expanding user base of Geographic Information System (GIS) services, conventional centralized architectures responsible for handling shortest distance queries are facing increasing challenges, such as heightened load pressure and longer response times. To mitigate these concerns, this study is the first to develop an edge computing framework specially tailored for processing distance queries. In conjunction with this innovative system, we have developed a straightforward, yet effective, labeling technique termed Border Labeling. Furthermore, we have devised and implemented a range of query strategies intended to capitalize on the capabilities of the edge computing infrastructure. Our experiments demonstrate that our solution surpasses other methods in terms of both indexing time and query speed across various road network datasets. The empirical evidence from our experiments supports the claim that our edge computing architecture significantly reduces the latency encountered by end-users, thus markedly decreasing their waiting times.

link: http://arxiv.org/abs/2403.11246v1

Compact 3D Gaussian Splatting For Dense Visual SLAM

Tianchen Deng, Yaohui Chen, Leyan Zhang, Jianfei Yang, Shenghai Yuan, Danwei Wang, Weidong Chen

Recent work has shown that 3D Gaussian-based SLAM enables high-quality reconstruction, accurate pose estimation, and real-time rendering of scenes. However, these approaches are built on a tremendous number of redundant 3D Gaussian ellipsoids, leading to high memory and storage costs, and slow training speed. To address the limitation, we propose a compact 3D Gaussian Splatting SLAM system that reduces the number and the parameter size of Gaussian ellipsoids. A sliding window-based masking strategy is first proposed to reduce the redundant ellipsoids. Then we observe that the covariance matrix (geometry) of most 3D Gaussian ellipsoids are extremely similar, which motivates a novel geometry codebook to compress 3D Gaussian geometric attributes, i.e., the parameters. Robust and accurate pose estimation is achieved by a global bundle adjustment method with reprojection loss. Extensive experiments demonstrate that our method achieves faster training and rendering speed while maintaining the state-of-the-art (SOTA) quality of the scene representation.

link: http://arxiv.org/abs/2403.11247v1

YOLOv9 for Fracture Detection in Pediatric Wrist Trauma X-ray Images

Chun-Tse Chien, Rui-Yang Ju, Kuang-Yi Chou, Jen-Shiun Chiang

The introduction of YOLOv9, the latest version of the You Only Look Once (YOLO) series, has led to its widespread adoption across various scenarios. This paper is the first to apply the YOLOv9 algorithm model to the fracture detection task as computer-assisted diagnosis (CAD) to help radiologists and surgeons to interpret X-ray images. Specifically, this paper trained the model on the GRAZPEDWRI-DX dataset and extended the training set using data augmentation techniques to improve the model performance. Experimental results demonstrate that compared to the mAP 50-95 of the current state-of-the-art (SOTA) model, the YOLOv9 model increased the value from 42.16% to 43.73%, with an improvement of 3.7%. The implementation code is publicly available at https://github.com/RuiyangJu/YOLOv9-Fracture-Detection.

link: http://arxiv.org/abs/2403.11249v1

NeoNeXt: Novel neural network operator and architecture based on the patch-wise matrix multiplications

Vladimir Korviakov, Denis Koposov

Most of the computer vision architectures nowadays are built upon the well-known foundation operations: fully-connected layers, convolutions and multi-head self-attention blocks. In this paper we propose a novel foundation operation - NeoCell - which learns matrix patterns and performs patchwise matrix multiplications with the input data. The main advantages of the proposed operator are (1) simple implementation without need in operations like im2col, (2) low computational complexity (especially for large matrices) and (3) simple and flexible implementation of up-/down-sampling. We validate NeoNeXt family of models based on this operation on ImageNet-1K classification task and show that they achieve competitive quality.

link: http://arxiv.org/abs/2403.11251v1

Uncertainty-Aware Pseudo-Label Filtering for Source-Free Unsupervised Domain Adaptation

Xi Chen, Haosen Yang, Huicong Zhang, Hongxun Yao, Xiatian Zhu

Source-free unsupervised domain adaptation (SFUDA) aims to enable the utilization of a pre-trained source model in an unlabeled target domain without access to source data. Self-training is a way to solve SFUDA, where confident target samples are iteratively selected as pseudo-labeled samples to guide target model learning. However, prior heuristic noisy pseudo-label filtering methods all involve introducing extra models, which are sensitive to model assumptions and may introduce additional errors or mislabeling. In this work, we propose a method called Uncertainty-aware Pseudo-label-filtering Adaptation (UPA) to efficiently address this issue in a coarse-to-fine manner. Specially, we first introduce a sample selection module named Adaptive Pseudo-label Selection (APS), which is responsible for filtering noisy pseudo labels. The APS utilizes a simple sample uncertainty estimation method by aggregating knowledge from neighboring samples and confident samples are selected as clean pseudo-labeled. Additionally, we incorporate Class-Aware Contrastive Learning (CACL) to mitigate the memorization of pseudo-label noise by learning robust pair-wise representation supervised by pseudo labels. Through extensive experiments conducted on three widely used benchmarks, we demonstrate that our proposed method achieves competitive performance on par with state-of-the-art SFUDA methods. Code is available at https://github.com/chenxi52/UPA.

link: http://dx.doi.org/10.1016/j.neucom.2023.127190

A learning-based solution approach to the application placement problem in mobile edge computing under uncertainty

Taha-Hossein Hejazi, Zahra Ghadimkhani, Arezoo Borji

Placing applications in mobile edge computing servers presents a complex challenge involving many servers, users, and their requests. Existing algorithms take a long time to solve high-dimensional problems with significant uncertainty scenarios. Therefore, an efficient approach is required to maximize the quality of service while considering all technical constraints. One of these approaches is machine learning, which emulates optimal solutions for application placement in edge servers. Machine learning models are expected to learn how to allocate user requests to servers based on the spatial positions of users and servers. In this study, the problem is formulated as a two-stage stochastic programming. A sufficient amount of training records is generated by varying parameters such as user locations, their request rates, and solving the optimization model. Then, based on the distance features of each user from the available servers and their request rates, machine learning models generate decision variables for the first stage of the stochastic optimization model, which is the user-to-server request allocation, and are employed as independent decision agents that reliably mimic the optimization model. Support Vector Machines (SVM) and Multi-layer Perceptron (MLP) are used in this research to achieve practical decisions from the stochastic optimization models. The performance of each model has shown an execution effectiveness of over 80%. This research aims to provide a more efficient approach for tackling high-dimensional problems and scenarios with uncertainties in mobile edge computing by leveraging machine learning models for optimal decision-making in request allocation to edge servers. These results suggest that machine-learning models can significantly improve solution times compared to conventional approaches.

link: http://arxiv.org/abs/2403.11259v1

A Lie Group Approach to Riemannian Batch Normalization

Ziheng Chen, Yue Song, Yunmei Liu, Nicu Sebe

Manifold-valued measurements exist in numerous applications within computer vision and machine learning. Recent studies have extended Deep Neural Networks (DNNs) to manifolds, and concomitantly, normalization techniques have also been adapted to several manifolds, referred to as Riemannian normalization. Nonetheless, most of the existing Riemannian normalization methods have been derived in an ad hoc manner and only apply to specific manifolds. This paper establishes a unified framework for Riemannian Batch Normalization (RBN) techniques on Lie groups. Our framework offers the theoretical guarantee of controlling both the Riemannian mean and variance. Empirically, we focus on Symmetric Positive Definite (SPD) manifolds, which possess three distinct types of Lie group structures. Using the deformation concept, we generalize the existing Lie groups on SPD manifolds into three families of parameterized Lie groups. Specific normalization layers induced by these Lie groups are then proposed for SPD neural networks. We demonstrate the effectiveness of our approach through three sets of experiments: radar recognition, human action recognition, and electroencephalography (EEG) classification. The code is available at https://github.com/GitZH-Chen/LieBN.git.

link: http://arxiv.org/abs/2403.11261v1

Stylized Face Sketch Extraction via Generative Prior with Limited Data

Kwan Yun, Kwanggyoon Seo, Chang Wook Seo, Soyeon Yoon, Seongcheol Kim, Soohyun Ji, Amirsaman Ashtari, Junyong Noh

Facial sketches are both a concise way of showing the identity of a person and a means to express artistic intention. While a few techniques have recently emerged that allow sketches to be extracted in different styles, they typically rely on a large amount of data that is difficult to obtain. Here, we propose StyleSketch, a method for extracting high-resolution stylized sketches from a face image. Using the rich semantics of the deep features from a pretrained StyleGAN, we are able to train a sketch generator with 16 pairs of face and the corresponding sketch images. The sketch generator utilizes part-based losses with two-stage learning for fast convergence during training for high-quality sketch extraction. Through a set of comparisons, we show that StyleSketch outperforms existing state-of-the-art sketch extraction methods and few-shot image adaptation methods for the task of extracting high-resolution abstract face sketches. We further demonstrate the versatility of StyleSketch by extending its use to other domains and explore the possibility of semantic editing. The project page can be found in https://kwanyun.github.io/stylesketch_project.

link: http://dx.doi.org/10.1111/cgf.15045

Forging the Forger: An Attempt to Improve Authorship Verification via Data Augmentation

Silvia Corbara, Alejandro Moreo

Authorship Verification (AV) is a text classification task concerned with inferring whether a candidate text has been written by one specific author or by someone else. It has been shown that many AV systems are vulnerable to adversarial attacks, where a malicious author actively tries to fool the classifier by either concealing their writing style, or by imitating the style of another author. In this paper, we investigate the potential benefits of augmenting the classifier training set with (negative) synthetic examples. These synthetic examples are generated to imitate the style of the author of interest. We analyze the improvements in classifier prediction that this augmentation brings to bear in the task of AV in an adversarial setting. In particular, we experiment with three different generator architectures (one based on Recurrent Neural Networks, another based on small-scale transformers, and another based on the popular GPT model) and with two training strategies (one inspired by standard Language Models, and another inspired by Wasserstein Generative Adversarial Networks). We evaluate our hypothesis on five datasets (three of which have been specifically collected to represent an adversarial setting) and using two learning

algorithms for the AV classifier (Support Vector Machines and Convolutional Neural Networks). This experimentation has yielded negative results, revealing that, although our methodology proves effective in many adversarial settings, its benefits are too sporadic for a pragmatical application.

link: http://arxiv.org/abs/2403.11265v1

Bilateral Propagation Network for Depth Completion

Jie Tang, Fei-Peng Tian, Boshi An, Jian Li, Ping Tan

Depth completion aims to derive a dense depth map from sparse depth measurements with a synchronized color image. Current state-of-the-art (SOTA) methods are predominantly propagation-based, which work as an iterative refinement on the initial estimated dense depth. However, the initial depth estimations mostly result from direct applications of convolutional layers on the sparse depth map. In this paper, we present a Bilateral Propagation Network (BP-Net), that propagates depth at the earliest stage to avoid directly convolving on sparse data. Specifically, our approach propagates the target depth from nearby depth measurements via a non-linear model, whose coefficients are generated through a multi-layer perceptron conditioned on both \emph{radiometric difference} and \emph{spatial distance}. By integrating bilateral propagation with multi-modal fusion and depth refinement in a multi-scale framework, our BP-Net demonstrates outstanding performance on both indoor and outdoor scenes. It achieves SOTA on the NYUv2 dataset and ranks 1st on the KITTI depth completion benchmark at the time of submission. Experimental results not only show the effectiveness of bilateral propagation but also emphasize the significance of early-stage propagation in contrast to the refinement stage. Our code and trained models will be available on the project page.

link: http://arxiv.org/abs/2403.11270v1

Studying and improving the performance of ETSI ITS contention-based forwarding (CBF) in urban and highway scenarios: S-FoT+

Oscar Amador, Ignacio Soto, Maria Calderon, Manuel Urueña

This paper evaluates the performance of ETSI ITS Contention-Based Forwarding (CBF) and ETSI Simple GeoBroadcast forwarding while disseminating warning messages over a Geographical Area in highway and urban scenarios. Our experimental evaluation considers the complete ETSI ITS architecture including the Decentralized Congestion Control (DCC) mechanism. We propose an enhanced CBF mechanism, named S-FOT+, which combines several improvements to the ETSI CBF algorithm. S-FoT+ has a similar or better performance than the ETSI forwarding algorithms regarding both reliability and end-to-end delay while requiring much fewer transmissions. The improvements are equally effective and efficient in both urban and highway scenarios with large Destination Areas. Finally, we evaluate the trade-offs that stem from using multi-hop broadcast mechanisms in urban settings with smaller Destination Areas when compared to single-hop broadcast. Results show that multi-hop mechanisms significantly improve coverage at the cost of an increased number of transmissions.

link: http://dx.doi.org/10.1016/j.comnet.2023.109899

BrightDreamer: Generic 3D Gaussian Generative Framework for Fast Text-to-3D Synthesis

Lutao Jiang, Lin Wang

Text-to-3D synthesis has recently seen intriguing advances by combining the text-to-image models with 3D representation methods, e.g., Gaussian Splatting (GS), via Score Distillation Sampling (SDS). However, a hurdle of existing methods is the low efficiency, per-prompt optimization for a single 3D object. Therefore, it is imperative for a paradigm shift from per-prompt optimization to one-stage generation for any unseen text prompts, which yet remains challenging. A hurdle is how to directly generate a set of millions of 3D Gaussians to represent a 3D object. This paper presents BrightDreamer, an end-to-end single-stage approach that can achieve generalizable and fast (77 ms) text-to-3D generation. Our key idea is to formulate the generation process as estimating the 3D deformation from an anchor shape with predefined positions. For this, we first propose a

Text-guided Shape Deformation (TSD) network to predict the deformed shape and its new positions, used as the centers (one attribute) of 3D Gaussians. To estimate the other four attributes (i.e., scaling, rotation, opacity, and SH coefficient), we then design a novel Text-guided Triplane Generator (TTG) to generate a triplane representation for a 3D object. The center of each Gaussian enables us to transform the triplane feature into the four attributes. The generated 3D Gaussians can be finally rendered at 705 frames per second. Extensive experiments demonstrate the superiority of our method over existing methods. Also, BrightDreamer possesses a strong semantic understanding capability even for complex text prompts. The project code is available at https://vlislab22.github.io/BrightDreamer.

link: http://arxiv.org/abs/2403.11273v1

Pattern-Based Peephole Optimizations with Java JIT Tests

Zhiqiang Zang, Aditya Thimmaiah, Milos Gligoric

We present JOG, a framework that facilitates developing Java JIT peephole optimizations alongside JIT tests. JOG enables developers to write a pattern, in Java itself, that specifies desired code transformations by writing code before and after the optimization, as well as any necessary preconditions. Such patterns can be written in the same way that tests of the optimization are already written in OpenJDK. JOG translates each pattern into C/C++ code that can be integrated as a JIT optimization pass. JOG also generates Java tests for optimizations from patterns. Furthermore, JOG can automatically detect possible shadow relation between a pair of optimizations where the effect of the shadowed optimization is overridden by another. Our evaluation shows that JOG makes it easier to write readable JIT optimizations alongside tests without decreasing the effectiveness of JIT optimizations. We wrote 162 patterns, including 68 existing optimizations in OpenJDK, 92 new optimizations adapted from LLVM, and two new optimizations that we proposed. We opened eight pull requests (PRs) for OpenJDK, including six for new optimizations, one on removing shadowed optimizations, and one for newly generated JIT tests; seven PRs have already been integrated into the master branch of OpenJDK.

link: http://dx.doi.org/10.1145/3597926.3598038

Fast Personalized Text-to-Image Syntheses With Attention Injection

Yuxuan Zhang, Yiren Song, Jinpeng Yu, Han Pan, Zhongliang Jing

Currently, personalized image generation methods mostly require considerable time to finetune and often overfit the concept resulting in generated images that are similar to custom concepts but difficult to edit by prompts. We propose an effective and fast approach that could balance the text-image consistency and identity consistency of the generated image and reference image. Our method can generate personalized images without any fine-tuning while maintaining the inherent text-to-image generation ability of diffusion models. Given a prompt and a reference image, we merge the custom concept into generated images by manipulating cross-attention and self-attention layers of the original diffusion model to generate personalized images that match the text description. Comprehensive experiments highlight the superiority of our method.

link: http://arxiv.org/abs/2403.11284v1

Advanced Knowledge Extraction of Physical Design Drawings, Translation and conversion to CAD formats using Deep Learning

Jesher Joshua M, Ragav V, Syed Ibrahim S P

The maintenance, archiving and usage of the design drawings is cumbersome in physical form in different industries for longer period. It is hard to extract information by simple scanning of drawing sheets. Converting them to their digital formats such as Computer-Aided Design (CAD), with needed knowledge extraction can solve this problem. The conversion of these machine drawings to its digital form is a crucial challenge which requires advanced techniques. This research proposes an innovative methodology utilizing Deep Learning methods. The approach employs object detection model, such as Yolov7, Faster R-CNN, to detect physical drawing objects present in the images followed by, edge detection algorithms such as canny filter to extract and refine the

identified lines from the drawing region and curve detection techniques to detect circle. Also ornaments (complex shapes) within the drawings are extracted. To ensure comprehensive conversion, an Optical Character Recognition (OCR) tool is integrated to identify and extract the text elements from the drawings. The extracted data which includes the lines, shapes and text is consolidated and stored in a structured comma separated values(.csv) file format. The accuracy and the efficiency of conversion is evaluated. Through this, conversion can be automated to help organizations enhance their productivity, facilitate seamless collaborations and preserve valuable design information in a digital format easily accessible. Overall, this study contributes to the advancement of CAD conversions, providing accurate results from the translating process. Future research can focus on handling diverse drawing types, enhanced accuracy in shape and line detection and extraction.

link: http://arxiv.org/abs/2403.11291v1