Tue 2024.04.30

Multi-Stream Cellular Test-Time Adaptation of Real-Time Models Evolving in Dynamic Environments

Benoît Gérin, Anaïs Halin, Anthony Cioppa, Maxim Henry, Bernard Ghanem, Benoît Macq, Christophe De Vleeschouwer, Marc Van Droogenbroeck

In the era of the Internet of Things (IoT), objects connect through a dynamic network, empowered by technologies like 5G, enabling real-time data sharing. However, smart objects, notably autonomous vehicles, face challenges in critical local computations due to limited resources. Lightweight AI models offer a solution but struggle with diverse data distributions. To address this limitation, we propose a novel Multi-Stream Cellular Test-Time Adaptation (MSC-TTA) setup where models adapt on the fly to a dynamic environment divided into cells. Then, we propose a real-time adaptive student-teacher method that leverages the multiple streams available in each cell to quickly adapt to changing data distributions. We validate our methodology in the context of autonomous vehicles navigating across cells defined based on location and weather conditions. To facilitate future benchmarking, we release a new multi-stream large-scale synthetic semantic segmentation dataset, called DADE, and show that our multi-stream approach outperforms a single-stream baseline. We believe that our work will open research opportunities in the IoT and 5G eras, offering solutions for real-time model adaptation.

link: http://arxiv.org/abs/2404.17930v1

Critical Review for One-class Classification: recent advances and the reality behind them

Toshitaka Hayashi, Dalibor Cimr, Hamido Fujita, Richard Cimler

This paper offers a comprehensive review of one-class classification (OCC), examining the technologies and methodologies employed in its implementation. It delves into various approaches utilized for OCC across diverse data types, such as feature data, image, video, time series, and others. Through a systematic review, this paper synthesizes promi-nent strategies used in OCC from its inception to its current advance-ments, with a particular emphasis on the promising application. Moreo-ver, the article criticizes the state-of-the-art (SOTA) image anomaly de-tection (AD) algorithms dominating one-class experiments. These algo-rithms include outlier exposure (binary classification) and pretrained model (multi-class classification), conflicting with the fundamental con-cept of learning from one class. Our investigation reveals that the top nine algorithms for one-class CIFAR10 benchmark are not OCC. We ar-gue that binary/multi-class classification algorithms should not be com-pared with OCC.

link: http://arxiv.org/abs/2404.17931v1

FDCE-Net: Underwater Image Enhancement with Embedding Frequency and Dual Color Encoder

Zheng Cheng, Guodong Fan, Jingchun Zhou, Min Gan, C. L. Philip Chen

Underwater images often suffer from various issues such as low brightness, color shift, blurred details, and noise due to light absorption and scattering caused by water and suspended particles. Previous underwater image enhancement (UIE) methods have primarily focused on spatial domain enhancement, neglecting the frequency domain information inherent in the images. However, the degradation factors of underwater images are closely intertwined in the spatial domain. Although certain methods focus on enhancing images in the frequency domain, they overlook the inherent relationship between the image degradation factors and the information present in the frequency domain. As a result, these methods frequently enhance certain attributes of the improved image while inadequately addressing or even exacerbating other attributes. Moreover, many existing methods heavily rely on prior knowledge to address color shift problems in underwater images, limiting their flexibility and robustness. In order to overcome these limitations, we propose the Embedding Frequency and Dual Color Encoder Network (FDCE-Net) in our paper. The FDCE-Net consists of two main structures: (1) Frequency Spatial Network (FS-Net) aims to achieve initial enhancement by utilizing our designed Frequency Spatial Residual Block (FSRB) to decouple image degradation factors in the frequency domain and enhance different attributes separately. (2) To tackle the color shift issue, we introduce the Dual-Color Encoder (DCE). The DCE establishes correlations between color and semantic representations through cross-attention and leverages multi-scale image features to guide the optimization of adaptive color query. The final enhanced images are generated by combining the outputs of FS-Net and DCE through a fusion network. These images exhibit rich details, clear textures, low noise and natural colors.

link: http://arxiv.org/abs/2404.17936v1

DTization: A New Method for Supervised Feature Scaling

Niful Islam

Artificial intelligence is currently a dominant force in shaping various aspects of the world. Machine learning is a sub-field in artificial intelligence. Feature scaling is one of the data pre-processing techniques that improves the performance of machine learning algorithms. The traditional feature scaling techniques are unsupervised where they do not have influence of the dependent variable in the scaling process. In this paper, we have presented a novel feature scaling technique named DTization that employs decision tree and robust scaler for supervised feature scaling. The proposed method utilizes decision tree to measure the feature importance and based on the importance, different features get scaled differently with the robust scaler algorithm. The proposed method has been extensively evaluated on ten classification and regression datasets on various evaluation matrices and the results show a noteworthy performance improvement compared to the traditional feature scaling methods.

link: http://arxiv.org/abs/2404.17937v1

CBMAP: Clustering-based manifold approximation and projection for dimensionality reduction

Berat Dogan

Dimensionality reduction methods are employed to decrease data dimensionality, either to enhance machine learning performance or to facilitate data visualization in two or three-dimensional spaces. These methods typically fall into two categories: feature selection and feature transformation. Feature selection retains significant features, while feature transformation projects data into a lower-dimensional space, with linear and nonlinear methods. While nonlinear methods excel in preserving local structures and capturing nonlinear relationships, they may struggle with interpreting global structures and can be computationally intensive. Recent algorithms, such as the t-SNE, UMAP, TriMap, and PaCMAP prioritize preserving local structures, often at the expense of accurately representing global structures, leading to clusters being spread out more in lower-dimensional spaces. Moreover, these methods heavily rely on hyperparameters, making their results sensitive to parameter settings. To address these limitations, this study introduces a clustering-based approach, namely CBMAP (Clustering-Based Manifold Approximation and Projection), for dimensionality reduction. CBMAP aims to preserve both global and local structures, ensuring that clusters in lower-dimensional spaces closely resemble those in high-dimensional spaces. Experimental evaluations on benchmark datasets demonstrate CBMAP's efficacy, offering speed, scalability, and minimal reliance on hyperparameters. Importantly, CBMAP enables low-dimensional projection of test data, addressing a critical need in machine learning applications. CBMAP is made freely available at https://github.com/doganlab/cbmap and can be installed from the Python Package Directory (PyPI) software repository with the command pip install cbmap.

link: http://arxiv.org/abs/2404.17940v1

Interaction Event Forecasting in Multi-Relational Recursive HyperGraphs: A Temporal Point Process Approach

Tony Gracious, Ambedkar Dukkipati

Modeling the dynamics of interacting entities using an evolving graph is an essential problem in fields such as financial networks and e-commerce. Traditional approaches focus primarily on pairwise interactions, limiting their ability to capture the complexity of real-world interactions involving multiple entities and their intricate relationship structures. This work addresses the problem of forecasting higher-order interaction events in multi-relational recursive hypergraphs. This is done using a dynamic graph representation learning framework that can capture complex relationships involving multiple entities. The proposed model, \textit{Relational Recursive Hyperedge Temporal Point Process} (RRHyperTPP) uses an encoder that learns a dynamic node representation based on the historical interaction patterns and then a hyperedge link prediction based decoder to model the event's occurrence. These learned representations are then used for downstream tasks involving forecasting the type and time of interactions. The main challenge in learning from hyperedge events is that the number of possible hyperedges grows exponentially with the number of nodes in the network. This will make the computation of negative log-likelihood of the temporal point process expensive, as the calculation of survival function requires a summation over all possible hyperedges. In our work, we use noise contrastive estimation to learn the parameters of our model, and we have experimentally shown that our models perform better than previous state-of-the-art methods for interaction forecasting.

link: http://arxiv.org/abs/2404.17943v1

Mobile Edge Computing

Sohaib Ahmed, Hassan Khalid, Muhammad Hamza, Danyal Farhat

Mobile Edge Computing (MEC) has emerged as a solution to the high latency and suboptimal Quality of Experience (QoE) associated with Mobile Cloud Computing (MCC). By processing data near the source, MEC reduces the need to send information to distant data centers, resulting in faster response times and lower latency. This paper explores the differences between MEC and traditional cloud computing, emphasizing architecture, data flow, and resource allocation. Key technologies like Network Function Virtualization (NFV) and Software-Defined Networking (SDN) are discussed for their role in achieving scalability and flexibility. Additionally, security and privacy challenges are addressed, underscoring the need for robust frameworks. We conclude with an examination of various edge computing applications and suggest future research directions to enhance the effectiveness and adoption of MEC in the evolving technological landscape.

link: http://arxiv.org/abs/2404.17944v1

Bounding the Expected Robustness of Graph Neural Networks Subject to Node Feature Attacks

Yassine Abbahaddou, Sofiane Ennadir, Johannes F. Lutzeyer, Michalis Vazirgiannis, Henrik Boström

Graph Neural Networks (GNNs) have demonstrated state-of-the-art performance in various graph representation learning tasks. Recently, studies revealed their vulnerability to adversarial attacks. In this work, we theoretically define the concept of expected robustness in the context of attributed graphs and relate it to the classical definition of adversarial robustness in the graph representation learning literature. Our definition allows us to derive an upper bound of the expected robustness of Graph Convolutional Networks (GCNs) and Graph Isomorphism Networks subject to node feature attacks. Building on these findings, we connect the expected robustness of GNNs to the orthonormality of their weight matrices and consequently propose an attack-independent, more robust variant of the GCN, called the Graph Convolutional Orthonormal Robust Networks (GCORNs). We further introduce a probabilistic method to estimate the expected robustness, which allows us to evaluate the effectiveness of GCORN on several real-world datasets. Experimental experiments showed that GCORN outperforms available defense methods. Our code is publicly available at: \href{https://github.com/Sennadir/GCORN}{https://github.com/Sennadir/GCORN}.

link: http://arxiv.org/abs/2404.17947v1

Transfer Learning Enhanced Single-choice Decision for Multi-choice Question Answering

Chenhao Cui, Yufan Jiang, Shuangzhi Wu, Zhoujun Li

Multi-choice Machine Reading Comprehension (MMRC) aims to select the correct answer from a set of options based on a given passage and question. The existing methods employ the pre-trained language model as the encoder, share and transfer knowledge through fine-tuning.These methods mainly focus on the design of exquisite mechanisms to effectively capture the relationships among the triplet of passage, question and answers. It is non-trivial but ignored to transfer knowledge from other MRC tasks such as SQuAD due to task specific of MMRC.In this paper, we reconstruct multi-choice to single-choice by training a binary classification to distinguish whether a certain answer is correct. Then select the option with the highest confidence score as the final answer. Our proposed method gets rid of the multi-choice framework and can leverage resources of other tasks. We construct our model based on the ALBERT-xxlarge model and evaluate it on the RACE and DREAM datasets. Experimental results show that our model performs better than multi-choice methods. In addition, by transferring knowledge from other kinds of MRC tasks, our model achieves state-of-the-art results in both single and ensemble settings.

link: http://arxiv.org/abs/2404.17949v1

Cauchy-Schwarz Divergence Information Bottleneck for Regression

Shujian Yu, Xi Yu, Sigurd Løkse, Robert Jenssen, Jose C. Principe

The information bottleneck (IB) approach is popular to improve the generalization, robustness and explainability of deep neural networks. Essentially, it aims to find a minimum sufficient representation $\mathbf{t}$ by striking a trade-off between a compression term $I(\mathbf{x};\mathbf{t})$ and a prediction term $I(y;\mathbf{t})$, where $I(\cdot;\cdot)$ refers to the mutual information (MI). MI is for the IB for the most part expressed in terms of the Kullback-Leibler (KL) divergence, which in the regression case corresponds to prediction based on mean squared error (MSE) loss with Gaussian assumption and compression approximated by variational inference. In this paper, we study the IB principle for the regression problem and develop a new way to parameterize the IB with deep neural networks by exploiting favorable properties of the Cauchy-Schwarz (CS) divergence. By doing so, we move away from MSE-based regression and ease estimation by avoiding variational approximations or distributional assumptions. We investigate the improved generalization ability of our proposed CS-IB and demonstrate strong adversarial robustness guarantees. We demonstrate its superior performance on six real-world regression tasks over other popular deep IB approaches. We additionally observe that the solutions discovered by CS-IB always achieve the best trade-off between prediction accuracy and compression ratio in the information plane. The code is available at \url{https://github.com/SJYuCNEL/Cauchy-Schwarz-Information-Bottleneck}.

link: http://arxiv.org/abs/2404.17951v1

PhishGuard: A Convolutional Neural Network Based Model for Detecting Phishing URLs with Explainability Analysis

Md Robiul Islam, Md Mahamodul Islam, Mst. Suraiya Afrin, Anika Antara, Nujhat Tabassum, Al Amin

Cybersecurity is one of the global issues because of the extensive dependence on cyber systems of individuals, industries, and organizations. Among the cyber attacks, phishing is increasing tremendously and affecting the global economy. Therefore, this phenomenon highlights the vital need for enhancing user awareness and robust support at both individual and organizational levels. Phishing URL identification is the best way to address the problem. Various machine learning and deep learning methods have been proposed to automate the detection of phishing URLs. However, these approaches often need more convincing accuracy and rely on datasets consisting of limited samples. Furthermore, these black box intelligent models decision to detect suspicious URLs needs proper explanation to understand the features affecting the output. To address the issues, we propose a 1D Convolutional Neural Network (CNN) and trained the model with extensive features and a substantial amount of data. The proposed model outperforms existing works by attaining an accuracy of 99.85%. Additionally, our explainability analysis highlights certain features that significantly contribute to identifying the phishing URL.

link: http://arxiv.org/abs/2404.17960v1

Random Walk on Pixel Manifolds for Anomaly Segmentation of Complex Driving Scenes

Zelong Zeng, Kaname Tomite

In anomaly segmentation for complex driving scenes, state-of-the-art approaches utilize anomaly scoring functions to calculate anomaly scores. For these functions, accurately predicting the logits of inlier classes for each pixel is crucial for precisely inferring the anomaly score. However, in real-world driving scenarios, the diversity of scenes often results in distorted manifolds of pixel embeddings in embedding space. This effect is not conducive to directly using the pixel embeddings for the logit prediction during inference, a concern overlooked by existing methods. To address this problem, we propose a novel method called Random Walk on Pixel Manifolds (RWPM). RWPM utilizes random walks to reveal the intrinsic relationships among pixels to refine the pixel embeddings. The refined pixel embeddings alleviate the distortion of manifolds, improving the accuracy of anomaly scores. Our extensive experiments show that RWPM consistently improve the performance of the existing anomaly segmentation methods and achieve the best results. Code: \url{https://github.com/ZelongZeng/RWPM}.

link: http://arxiv.org/abs/2404.17961v1

Deep Learning for Low-Latency, Quantum-Ready RF Sensing

Pranav Gokhale, Caitlin Carnahan, William Clark, Frederic T. Chong

Recent work has shown the promise of applying deep learning to enhance software processing of radio frequency (RF) signals. In parallel, hardware developments with quantum RF sensors based on Rydberg atoms are breaking longstanding barriers in frequency range, resolution, and sensitivity. In this paper, we describe our implementations of quantum-ready machine learning approaches for RF signal classification. Our primary objective is latency: while deep learning offers a more powerful computational paradigm, it also traditionally incurs latency overheads that hinder wider scale deployment. Our work spans three axes. (1) A novel continuous wavelet transform (CWT) based recurrent neural network (RNN) architecture that enables flexible online classification of RF signals on-the-fly with reduced sampling time. (2) Low-latency inference techniques for both GPU and CPU that span over 100x reductions in inference time, enabling real-time operation with sub-millisecond inference. (3) Quantum-readiness validated through application of our models to physics-based simulation of Rydberg atom QRF sensors. Altogether, our work bridges towards next-generation RF sensors that use quantum technology to surpass previous physical limits, paired with latency-optimized AI/ML software that is suitable for real-time deployment.

link: http://arxiv.org/abs/2404.17962v1

SCorP: Statistics-Informed Dense Correspondence Prediction Directly from Unsegmented Medical Images

Krithika Iyer, Jadie Adams, Shireen Y. Elhabian

Statistical shape modeling (SSM) is a powerful computational framework for quantifying and analyzing the geometric variability of anatomical structures, facilitating advancements in medical research, diagnostics, and treatment planning. Traditional methods for shape modeling from imaging data demand significant manual and computational resources. Additionally, these methods necessitate repeating the entire modeling pipeline to derive shape descriptors (e.g., surface-based point correspondences) for new data. While deep learning approaches have shown promise in streamlining the construction of SSMs on new data, they still rely on traditional techniques to supervise the training of the deep networks. Moreover, the predominant linearity assumption of traditional approaches restricts their efficacy, a limitation also inherited by deep learning models trained using optimized/established correspondences. Consequently, representing complex anatomies becomes challenging. To address these limitations, we introduce SCorP, a novel framework capable of predicting surface-based correspondences directly from unsegmented images. By leveraging the shape prior learned directly from surface meshes in an unsupervised manner, the proposed model eliminates the need for an optimized shape model for training supervision. The strong shape prior acts as a teacher and regularizes the feature learning of the student network to guide it in learning image-based features that are predictive of surface correspondences. The proposed model streamlines the training and inference phases by removing the supervision for the correspondence prediction task while alleviating the linearity assumption.

link: http://arxiv.org/abs/2404.17967v1

Usefulness of Emotional Prosody in Neural Machine Translation

Charles Brazier, Jean-Luc Rouas

Neural Machine Translation (NMT) is the task of translating a text from one language to another with the use of a trained neural network. Several existing works aim at incorporating external information into NMT models to improve or control predicted translations (e.g. sentiment, politeness, gender). In this work, we propose to improve translation quality by adding another external source of information: the automatically recognized emotion in the voice. This work is motivated by the assumption that each emotion is associated with a specific lexicon that can overlap between emotions. Our proposed method follows a two-stage procedure. At first, we select a state-of-the-art Speech Emotion Recognition (SER) model to predict dimensional emotion values from all input audio in the dataset. Then, we use these predicted emotions as source tokens added at the beginning of input texts to train our NMT model. We show that integrating emotion information, especially arousal, into NMT systems leads to better translations.

link: http://arxiv.org/abs/2404.17968v1