PhD subject:

Generative model for spatio-temporal data

Hosting institute

ICube Laboratory (The Engineering science, computer science and imaging laboratory) at the University of Strasbourg is a leading research center in Computer Science, with more than 300 permanent researchers, with the recently opened AI graduate school supported by the French government.

Work place

The thesis work will take place in the MLMS (Machine Learning, Modeling & Simulation) research team of the ICube laboratory, in collaboration with EE Dept. at KAIST, South Korea. The workplace is located on the hospital site of the laboratory, a 10-minute walk from the heart of downtown Strasbourg, listed as a UNESCO World Heritage Site.

Supervisors

– Supervisor: Hyewon Seo (Directrice de Recherche, CNRS—Univ. Strasbourg)

– Co-supervisor: Chan-Hyun Youn (Professor, KAIST)

Staring date

October 2024, over 3 years.

Context and work description

Spatiotemporal data provides a comprehensive view of the evolution of phenomena or organic structures over time and space, with a wide range of applications in various domains, such as environmental monitoring and climate analysis [BXZ+23], transportation and traffic management [BHL+19], and longitudinal analysis of disease progression [QCFAK23]. The Transformer architecture, initially conceived and demonstrated as a powerful model for natural language processing tasks, has been adapted to become a formidable tool for spatiotemporal analysis, demonstrating superior parallelization and scaling capabilities to RNNs, in part due to its attention mechanism. Graph neural networks (GNNs), used in some studies to extract spatial information [WMW+20], further enhance spatiotemporal analyses by exploiting their adaptability to irregularly distributed data and their flexibility in integrating various sources of information.

In this study, we aim to develop a novel model specifically tailored for generating spatiotemporal data. Our strategy involves leveraging diffusion probabilistic models for denoising (DDPM) while exploring other innovative theories with the aim of comprehensively incorporating spatiotemporal conditions, possibly from different modalities. DDPM has demonstrated remarkable generative capability in the text-image generation task [DN21], and possesses inherent advantages for dealing with the many-to-many distribution fitting problem, making it a superior candidate for spatiotemporal data generation. However, diffusion models typically rely on U-Net, which is not explicitly crafted for sequential data but excels in feature extraction. Although some researchers have attempted to incorporate Transformers into diffusion models [ZFY*23], there remains a notable absence of tailor-made adaptations to meet the demands of spatiotemporal data.

Recognizing the limitations of existing model architectures, we intend to transition to a new model capable of exploiting spatial information while exploring temporal relationships in depth. At the same time, we aim to establish a reliable latent space for the representation of spatiotemporal data, with potential benefits for

several downstream tasks such as anomaly detection or longitudinal analysis. Our specific aims are as follows:

1. Generative model for the generation of spatiotemporal data with spatiotemporal conditioning. The advantages of probabilistic diffusion models for denoising (DDPM) will be deployed, with advanced spatiotemporal encoding capabilities [LFL*23]. The underlying spatiotemporal constraints will be modeled, such as time-varying tasks or stimuli given to the subject at the time of biometric data acquisition.
2. Extension to multimodality. Taking inspiration from the multi-model diffusion model [RMY*23], we will design a joint denoising process so as to efficiently generate multimodal data consistent across modalities. Going beyond existing works that deal with at most two modalities [WYM*24], we will tackle the challenging cases involving more than three modalities, where relying on paired modeling can be suboptimal.

The project will encompass several applications, including but not limited to, boat surveillance and research on neurodegenerative diseases.

Candidate profile

− Master student in Computer Science, Electronic & Electrical Engineering, or in Applied Mathematics
− Solid programming skills: Python/Pytorch/C++
− Proficiency in Deep Learning techniques
− Good communication skills

Application

Send your CV and your academic transcripts (Bachelor and Master) to seo@unistra.fr.

Bibliography

[BHL+19] Bai, S., He, Z., Lei, Y., et al., Traffic anomaly detection via perspective map based on spatial-temporal information matrix. In CVPR Workshops, 2019.

[BXZ+23] Bi, K., Xie, L., Zhang, H. et al. Accurate medium-range global weather forecasting with 3D neural networks. Nature 619, 2023.

[DN21] Dhariwal, P., and Nichol A., Diffusion models beat GANs on image synthesis, Advances in neural information processing systems 34, 2021.

[LFL*23] Liu C., Fan W., Liu Y., et al.: Generative Diffusion Models on Graphs: Methods and Applications, Int'l Joint Conference on Artificial Intelligence, 2023.

[QCFAK23] Qin C, Cao J, Fu H, Anwer RM, Khan FS. A Spatial-Temporal Deformable Attention Based Framework for Breast Lesion Detection in Videos. In Int'l Conf. Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2023.

[RMY*23] Ruan L., Ma Y., Yang H., et al, MM-Diffusion: Learning Multi-Modal Diffusion Models for Joint Audio and Video Generation, IEEE Computer Vision and Pattern Recognition (CVPR) 2023.

[WMW+20] Wang, X., Ma, Y., Wang, Y., et al, Traffic flow prediction via spatial temporal graph neural network. In Proc. web conference, 2020.

[WYM*24] Wang D., Yuan K., Muller C., Blanc F., Padoy N., Seo H., Enhancing Gait Video Analysis in Neurodegenerative Diseases by Knowledge Augmentation in Vision Language Model, https://arxiv.org/abs/2403.13756.

[ZFY*23] Zou K., Faisan S., Yu B., Valette S., Seo H., 4D Facial Expression Diffusion Model, https://arxiv.org/abs/2303.16611, accepted for publication, ACM Trans. Multimedia Computing, Communications and Applications, 2023.