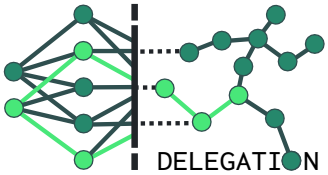
IRIMAS UNIVERSITÉ HAUTE-ALSACE
Institut de Recherche en Informatique, Mathématiques, Automatique et Signal

Master Internship Position

Deep Learning architectures for generating rehabilitation human motion



1 Information about the internship

- **Supervisors:**
  - Dr. Maxime Devanne (https://maxime-devanne.com)
  - Ali El Hadi Ismail Fawaz (https://hadifawaz1999.github.io)
  - Prof. Jonathan Weber (https://www.jonathan-weber.eu)
  - Prof. Germain Forestier (https://germain-forestier.info)
- **Location:** UHA/IRIMAS EA 7499, Mulhouse, France
- **Duration:** 6 months (starting from February or March 2022)
- **Salary:** 600.60€ monthly
- **Keywords:** Deep Learning, generative models, human motion, time series, rehabilitation

2 Context

Human motion analysis is crucial for studying people and understanding how they behave, communicate and interact with real world environments. Due to the complex nature of body movements as well as the high cost of motion capture systems, acquisition of human motion is not straightforward and thus constraints data production. Hopefully, recent approaches estimating human poses from videos offer new opportunities to analyze skeleton-based human motion [1, 2]. While skeleton-based human motion analysis has been extensively studied for behavior understanding like action recognition, some efforts are yet

to be done for the task of human motion generation. Particularly, the automatic generation of motion sequences is beneficial for rapidly increasing the amount of data and improving Deep Learning-based analysis algorithms. In particular, this is crucial in a medical context like in physical rehabilitation where acquiring data is challenging. Rehabilitation human motions are corresponding to rehabilitation exercises proposed by physiotherapists.

Since several years, human motion generation paradigms have been possible thanks to the appearance of Generative Adversarial Networks (GAN) [3], Variational AutoEncoder (VAE) [4] or Diffusion models [5]. While most of these works have considered motion capture (mocap) data, we consider noisy skeleton data estimated from videos as it is easily applicable in real-world scenarios for the general public.

3 Goals

The goal of this internship is to investigate deep generative models for skeleton-based human motion sequences with a particular focus on rehabilitation data. Inspiring from recent effective Deep Learning-based approaches [6, 7, 8, 4], the aim is to generate full skeleton-based rehabilitation motion sequences. It is therefore crucial to investigate how deep generative models can handle such noisy and possibly incomplete data in order to generate novel rehabilitation motion sequences as natural and variable as possible.

In particular, the candidate will work on the following tasks:

- **Deep generative models adapted to rehabilitation data:** based on studies from existing works, the goal is to build generative models for rehabilitation sequences. Therefore, the candidate will investigate different generative models, like GAN, VAE and Diffusion models, in order to propose and develop a complete Deep Learning model for generating skeleton-based human motions. These models will be trained using publicly available datasets such as the Kimore dataset [9].

- **Evaluation of deep generative models:** in order to validate the proposed model, experimental evaluation is crucial. In comparison to motion recognition where classification accuracy is a natural way to assess an approach, evaluating the task of motion generation is not as straightforward. Dedicated metrics evaluating both naturalness and diversity of generated sequences as well as the impact of new generated sequences in a classification task will be considered.

- **Text to rehabilitation motion:** The generated models will be then adapted to take as input text sequences corresponding to rehabilitation exercises' descriptions. This will be particularly useful to create new rehabilitation exercises.

4 Profile of applicant

The candidate must fit the following requirements:

- Registered in Master 2 or last year of Engineering School (or equivalent) in **Computer Science**
- Advanced skills in **Python programming** are mandatory
- Good skills in **Machine Learning & Deep Learning** using related libraries (scikit-learn, Tensorflow, Pytorch, etc.) are required
- Knowledge and/or a first experience in **human motion analysis** will be appreciated

5 Research environment

The proposed internship will be carried out within the MSD (Modeling and Data Science) team from the IRIMAS Institute. It will be part of the ANR DELEGATION project [1].

6 Application

For further information or for applying, candidates should send a **CV, academic records, personal projects (e.g. github repo) and a motivation letter** to maxime.devanne@uha.fr.

[1] https://maxime-devanne.com/delegation/

References

[1] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, "Realtime multi-person 2d pose estimation using part affinity fields," in *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7291–7299, 2017.

[2] V. Bazarevsky, I. Grishchenko, K. Raveendran, T. Zhu, F. Zhang, and M. Grundmann, "Blazepose: On-device real-time body pose tracking," *arXiv preprint arXiv:2006.10204*, 2020.

[3] C.-J. Chou, J.-T. Chien, and H.-T. Chen, "Self adversarial training for human pose estimation," in *Asia-Pacific Signal and Information Processing Association Annual Summit and Conference*, pp. 17–30, IEEE, 2018.

[4] M. Petrovich, M. J. Black, and G. Varol, "Action-conditioned 3d human motion synthesis with transformer vae," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 10985–10995, 2021.

[5] M. Zhang, Z. Cai, L. Pan, F. Hong, X. Guo, L. Yang, and Z. Liu, "Motiondiffuse: Text-driven human motion generation with diffusion model," *arXiv preprint arXiv:2208.15001*, 2022.

[6] J. Martinez, M. J. Black, and J. Romero, "On human motion prediction using recurrent neural networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2891–2900, 2017.

[7] C. Li, Z. Zhang, W. S. Lee, and G. H. Lee, "Convolutional sequence to sequence model for human dynamics," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5226–5234, 2018.

[8] Y. Cai, L. Huang, Y. Wang, T.-J. Cham, J. Cai, J. Yuan, J. Liu, X. Yang, Y. Zhu, X. Shen, *et al.*, "Learning progressive joint propagation for human motion prediction," in *European Conference on Computer Vision*, pp. 226–242, Springer, 2020.

[9] M. Capecci, M. G. Ceravolo, F. Ferracuti, S. Iarlori, A. Monteriu, L. Romeo, and F. Verdini, "The kimore dataset: Kinematic assessment of movement and clinical scores for remote monitoring of physical rehabilitation," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 7, pp. 1436–1448, 2019.