

Apprentissage profond pour l'analyse de graphes appliqué à la reconnaissance d'activités humaines

Mots-clés : Deep learning, Graph embedding, HAR, Spatio-temporal image encoding, CNN

Contexte

La recherche sur la reconnaissance de l'activité humaine ou *Human Activity Recognition* en anglais (HAR) a pris de l'ampleur ces dernières années en raison de son utilisation dans plusieurs domaines tels que la sécurité et l'assistance à la personne [5]. Cette HAR peut être réalisée en utilisant des séquences de données squelettes 3D issus d'une Kinect combinées aux techniques d'apprentissage automatique.

Actuellement, des travaux sur la HAR sont menés au CERV et se basent sur la détection d'actions à partir d'un encodage spatio-temporel de séquences de squelette 3D sous forme d'image et de modèles d'apprentissage profond (CNN) pour la reconnaissance [4]. Cependant, il existe d'autres représentations possibles de ces séquences, à l'exemple de l'utilisation de graphes, où les noeuds représentent les articulations, et les arêtes représentent les liens entre ces articulations [6, 2] (Figure 1). Certains travaux, à l'exemple de Chen & al. [1] proposent d'effectuer un embedding sur ces graphes, avant de passer à la phase d'apprentissage.

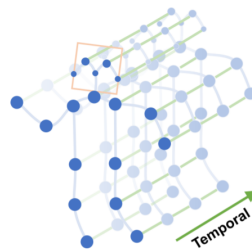


Figure 1: Spatial temporal graph of a skeleton sequence [6].

Stage

Dans le cadre de ce stage, nous nous intéresserons à l'utilisation de la technique dite de "l'embedding sur graphe" (à l'exemple du node2vec [3]) dans l'objectif d'exploiter ces derniers en utilisant un modèle de deep learning.

L'approche proposée sera ensuite comparée aux différentes techniques utilisées, dans l'état de l'art, pour la reconnaissance d'activités humaines, notamment la représentation sous la forme d'images.

Étapes principales du stage

- état de l'art sur l'embedding sur graphes.
- proposition d'un ou plusieurs encodages de graphes à partir des données squelettes 3D.
- tests sur différentes bases de données et sélection du meilleur encodage (si plusieurs possible).
- comparaison avec l'état de l'art en HAR

Environnement de travail

- Langage : Python
- Keras ou Pytorch pour l'apprentissage automatique
- Matériel : Kinect v2, Ordinateur portable

Contact

- **Encadrants** : Nassim Mokhtari et Alexis Nédélec
- **Contacts** : mokhtari@enib.fr nedelec@enib.fr
- **Lieu** : Centre Européen de Réalité Virtuelle (CERV)
- **Financement** : ENIB

References

- [1] S. Chen, K. Xu, X. Jiang, and T. Sun. Pyramid spatial-temporal graph transformer for skeleton-based action recognition. *Applied Sciences*, 12(18), 2022. ISSN 2076-3417. doi: 10.3390/app12189229. URL <https://www.mdpi.com/2076-3417/12/18/9229>.
- [2] M. Delamare, C. Laville, A. Cabani, and H. Chafouk. Graph convolutional networks skeleton-based action recognition for continuous data stream: A sliding window approach. 02 2021. doi: 10.5220/0010234904270435.
- [3] A. Grover and J. Leskovec. node2vec: Scalable feature learning for networks, 2016. URL <https://arxiv.org/abs/1607.00653>.
- [4] N. Mokhtari, A. Nédélec, and P. Loor. Human activity recognition: A spatio-temporal image encoding of 3d skeleton data for online action detection. pages 448–455, 01 2022. doi: 10.5220/0010835800003124.
- [5] J. Wang, Y. Chen, S. Hao, X. Peng, and L. Hu. Deep learning for sensor-based activity recognition: A survey. *Pattern Recognition Letters*, 119:3–11, mar 2019. doi: 10.1016/j.patrec.2018.02.010. URL <https://doi.org/10.1016%2Fj.patrec.2018.02.010>.
- [6] S. Yan, Y. Xiong, and D. Lin. Spatial temporal graph convolutional networks for skeleton-based action recognition. 01 2018.