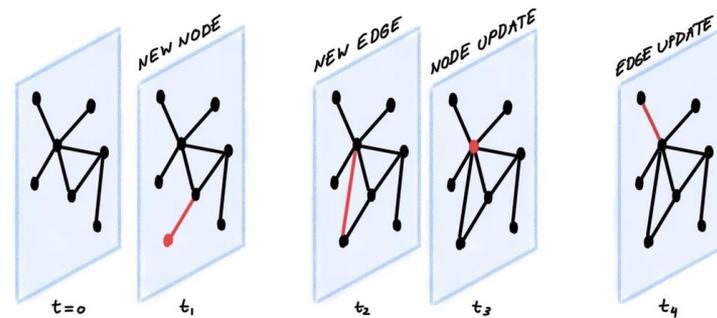
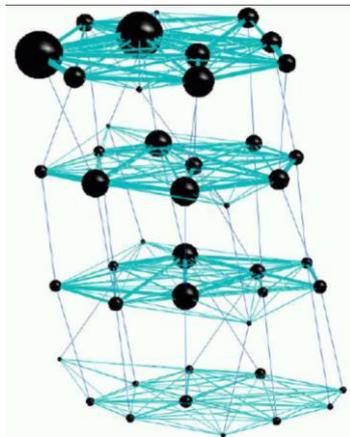


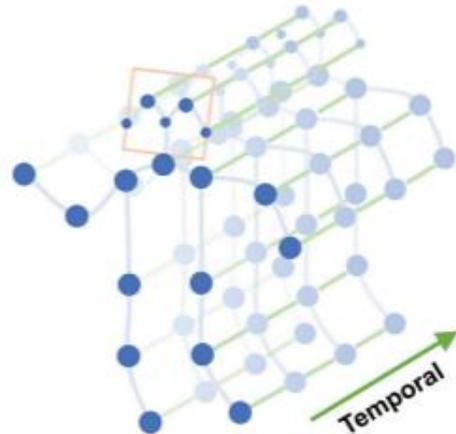
Les GNN pour les graphes spatio-temporels pour la reconnaissance d'actions



- 1. Spécificité du problème**
- 2. Représentations sous différentes formes de graphes**
- 3. Problématiques**
- 4. Rôle du partitionnement**
- 5. Convolution sur graphes (Spatiale/temporelle)**
- 6. Pooling**
- 7. Progrès dans la reconnaissance d'actions humaines**
- 8. Opportunité**
- 9. Conclusion**
- 10. Questions**

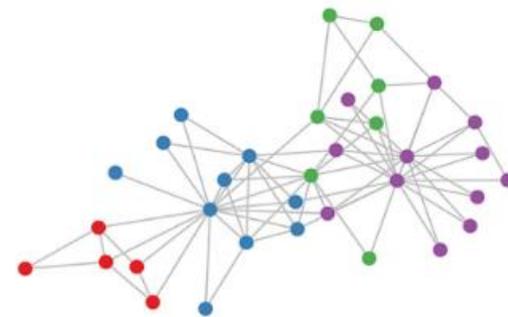
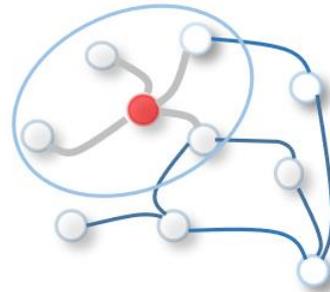
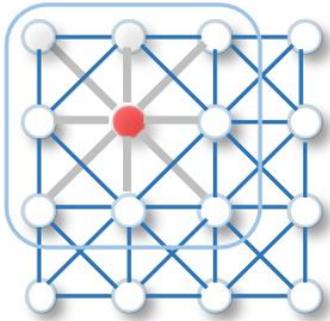
Spécificité du problème :

Graphes spatio-temporels : les séquences de graphes.



- Séquence de graphes
- Possibilité à tout temps t d'être plongé dans un espace euclidien.
- La structure reste inchangée seule les caractéristiques évoluent.

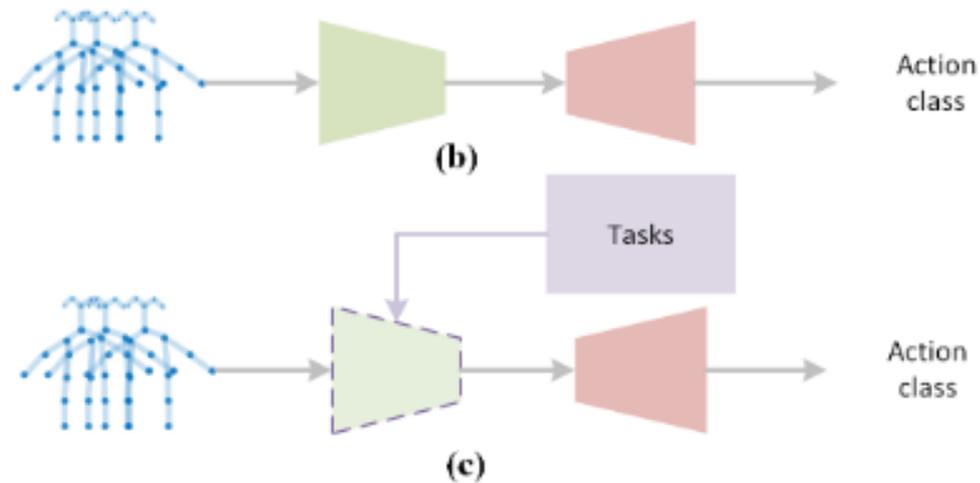
- Données euclidiennes Vs Non Euclidiennes



- Evolution en fonction d'événements
- Apparition/disparition de nœuds ou d'arêtes

Spécificité du problème :

Différents types de réseaux.



2 paradigmes :

- un réseau pour la classification uniquement supervisé.
- Les réseaux génératifs : un modèle non supervisé accompagne la tâche de classification

Spécificité du problème :

Séquences de graphes

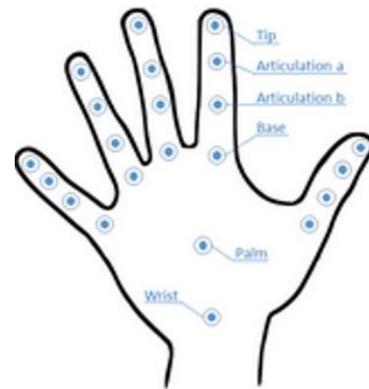


Figure 1

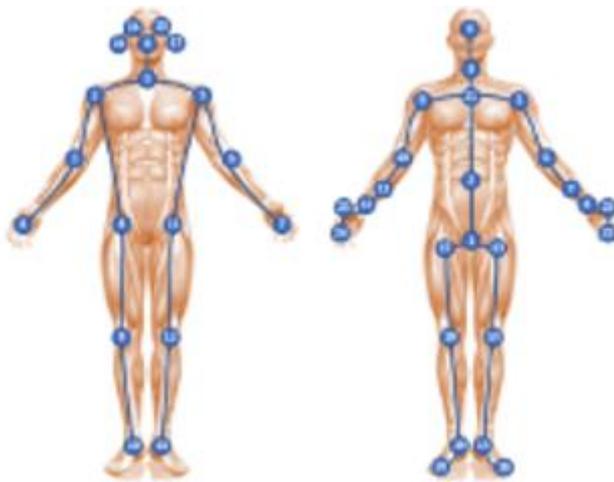


Figure 2



Figure 3

- Plusieurs type de données possédant des caractéristiques communes.

Fig 2.:SHI, Lei, ZHANG, Yifan, CHENG, Jian, et al. Skeleton-based action recognition with multi-stream adaptive graph convolutional networks. *IEEE Transactions on Image Processing*, 2020, vol. 29, p. 9532-9545.

Fig 3:PAPADOPOULOS, Konstantinos, KACEM, Anis, SHABAYEK, Abdelrahman, et al. Face-gcn: A graph convolutional network for 3d dynamic face identification/recognition. *arXiv preprint arXiv:2104.09145*, 2021

Spécificité du problème :

Différents types d'arêtes

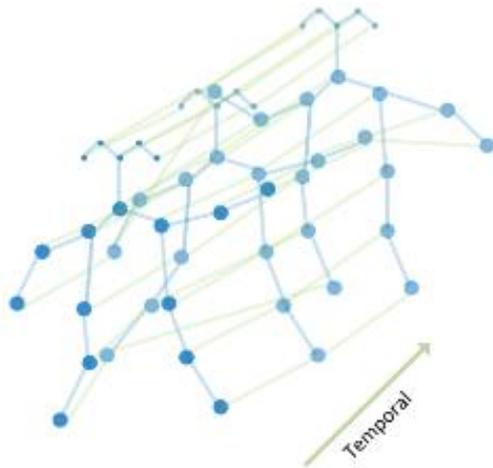


Figure 1

- Arêtes spatiales
- Arêtes temporelles



Figure 2

- Graphes d'interactions
 - Arêtes intra-sujet
 - Arêtes inter-sujets

Figure 1,2 FENG, Miao et MEUNIER, Jean. Skeleton graph-neural-network-based human action recognition: A survey. Sensors, 2022, vol. 22, no 6, p. 2091.

Représentations sous différentes formes de graphes :

3 considérations d'un graphe spatio-temporel

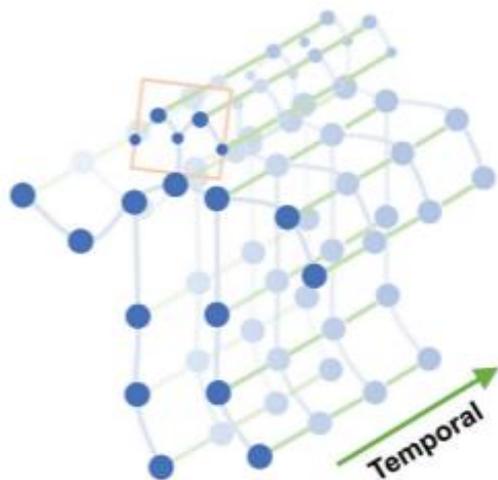


Figure 1

1- Succession de *frames*

3- Cas dérivé un Graphe pour toutes la séquence

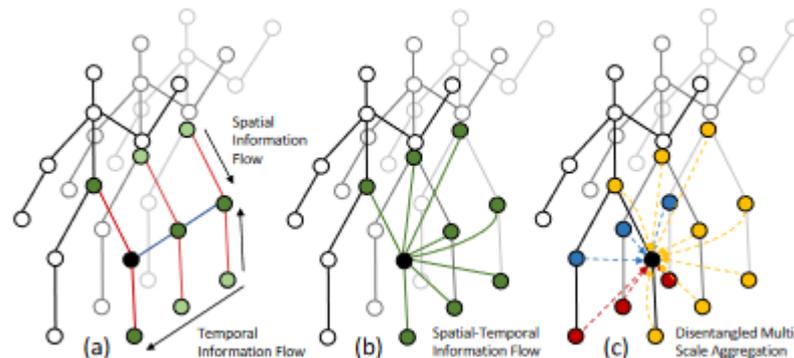
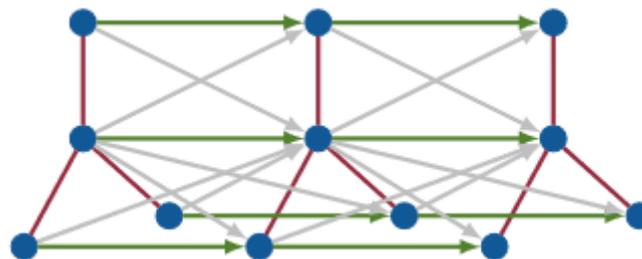


Figure 2

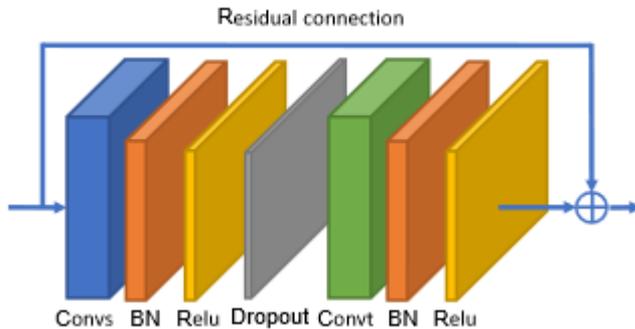
2- Regroupement de plusieurs frames = Mini-graphes spatio-temporelles



YAN, Sijie, XIONG, Yuanjun, et LIN, Dahua. *Spatial temporal graph convolutional networks for skeleton-based action recognition*. In : *Proceedings of the AAAI conference on artificial intelligence*. 2018.

LIU, Ziyu, ZHANG, Hongwen, CHEN, Zhenghao, et al. *Disentangling and unifying graph convolutions for skeleton-based action recognition*. In : *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2020. p. 143-152.

Architecture de réseaux



- Structure d'un réseau capable de traiter une succession de frames

Figure 1

- On trouve bien une partie pour l'aspect spatial et une pour l'aspect temporel

Figure 2

Fig 1, 2 : SHI, Lei, ZHANG, Yifan, CHENG, Jian, et al. Skeleton-based action recognition with multi-stream adaptive graph convolutional networks. *IEEE Transactions on Image Processing*, 2020, vol. 29, p. 9532-9545.

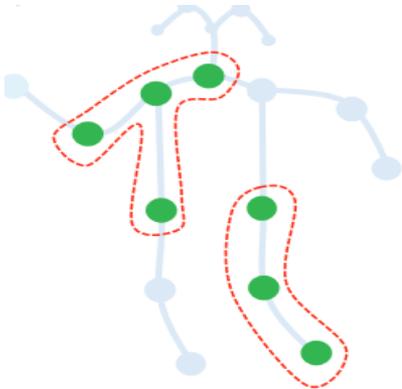
Problématiques

Cas de graphes Spatio-temporels:
Comment traiter les dimensions spatiales et temporelles ?

Convolution sur graphes :

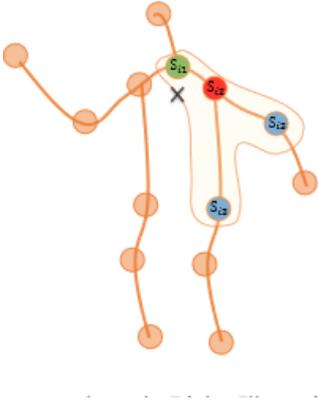
Sans partitionnement

$$f_{out} = Af_{in}W$$



Avec partitionnement

$$f_{out} = \sum_k^K A_k f_{in} W_k$$



Rôle du partitionnement ?

Masques dans la convolution

- Rôle de pondération des arêtes
 - Obtenu par somme de matrices portant un sens sémantique indépendant
 - Appris par un mécanisme d'attention
- Intégration dans la matrice d'adjacence par
 - Soit une multiplication élément par élément
 - Soit par addition de matrices

Exemple usage de masque

Usage de différents Masques typiques:

- A_k structure naturelle du graphe
- B_k masque appris par rétropropagation
- C_k masque guidé par mécanisme de self-attention

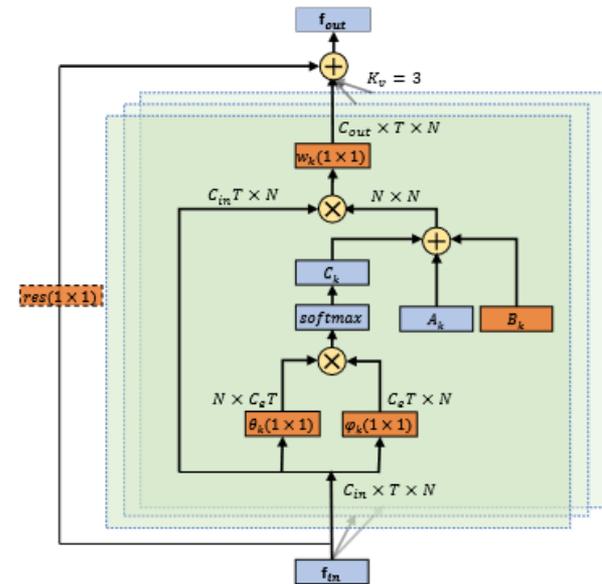
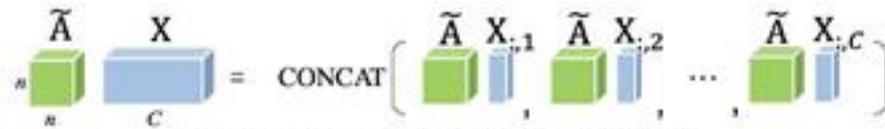


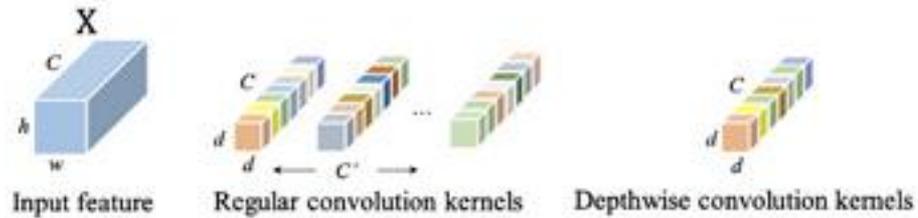
Figure 1

Convolution sur graphe :

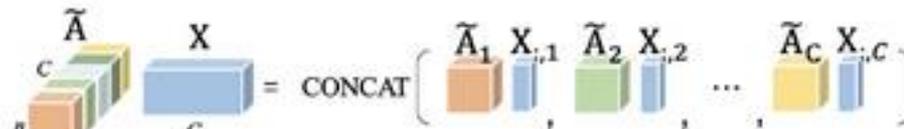
Usage de la matrice d'Adjacence



(a) Coupling aggregation in GCNs



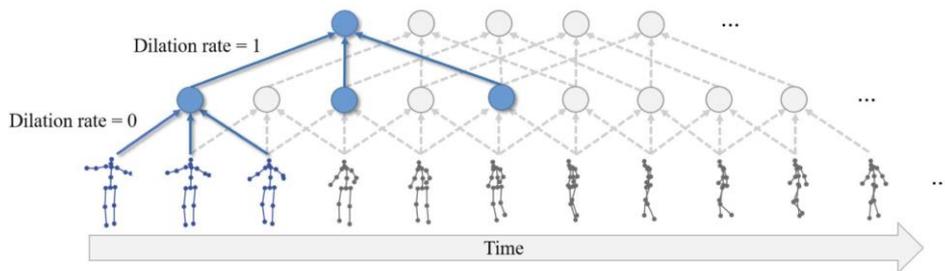
(b) Decoupling aggregation in CNNs



(c) Decoupling aggregation in GCNs

Aspect temporel

- Comment dépasser la convolution temporelle limitée aux frames contigüe et avec la même articulation ?



The temporal dilated graph convolution

Figure 1

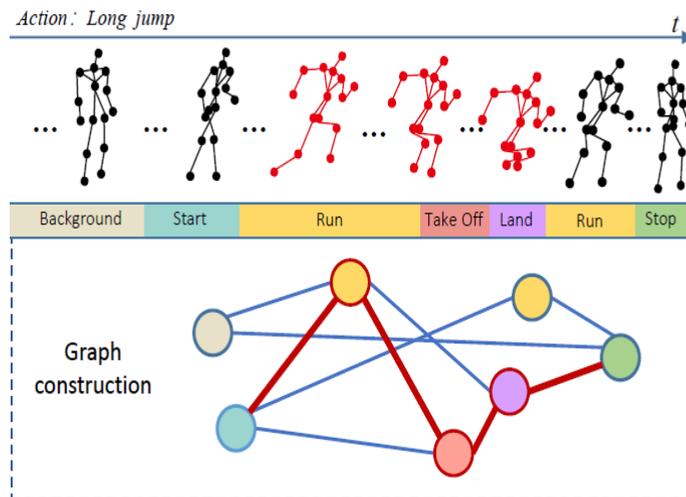


Figure 2

Fig 1 :ZHANG, Jiayu, YE, Gaoxiang, TU, Zhigang, et al. A spatial attentive and temporal dilated (SATD) GCN for skeleton-based action recognition. CAAI Transactions on Intelligence Technology, 2022, vol. 7, no 1, p. 46-55.

Fig 2: LI, Jianan, XIE, Xuemei, ZHAO, Zhifu, et al. Temporal graph modeling for skeleton-based action recognition. arXiv preprint arXiv:2012.08804, 2020

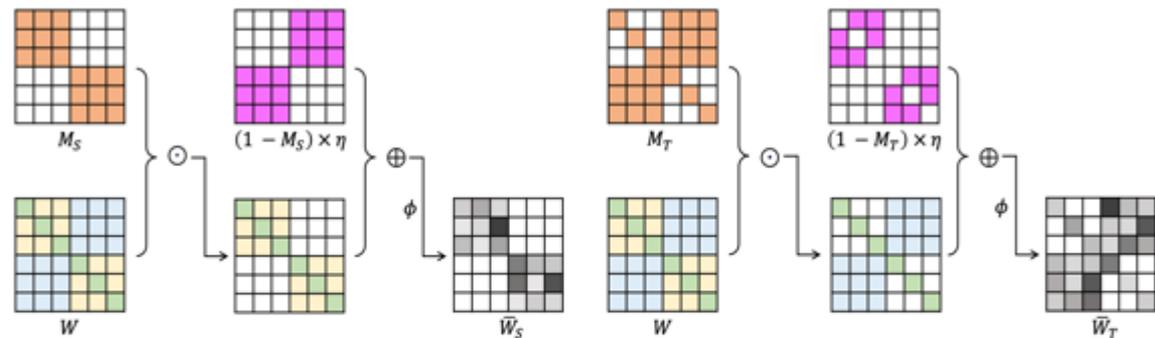
Aspect temporelle :
géré par des réseaux de type RNN

- Pourquoi des RNN ?
 - Capables de modéliser des systèmes non Markoviens
- Limite
 - Pas conçu initialement pour le spatial
- Evolution :
 - Prise en compte du voisinage dans les RNN.
- Désavantage
 - Beaucoup de paramètres par rapport à la convolution

Aspect temporelle :

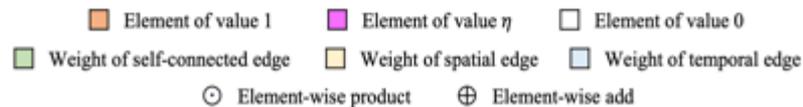
dans le cas de mini-graphes ou graphes uniques

$$\tilde{\mathbf{A}}_{(\tau)} = \begin{bmatrix} \tilde{\mathbf{A}} & \dots & \tilde{\mathbf{A}} \\ \vdots & \ddots & \vdots \\ \tilde{\mathbf{A}} & \dots & \tilde{\mathbf{A}} \end{bmatrix} \in \mathbb{R}^{\tau N \times \tau N}.$$



(a) Spatial Mask Operation

(b) Temporal Mask Operation



L'attention dans les graphes :

Mécanisme de self-attention incorporé au graphe : cas spatial et temporel

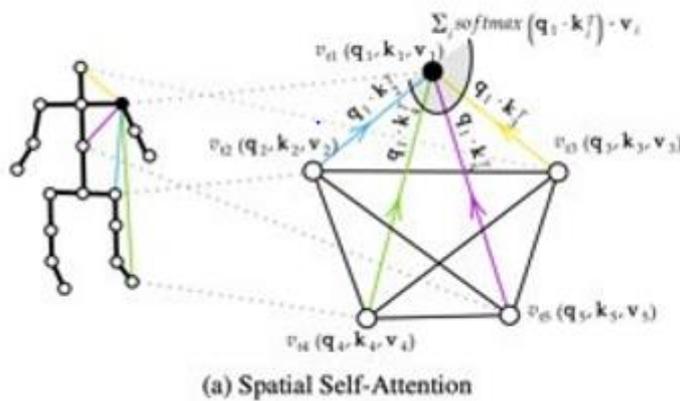


Figure 1

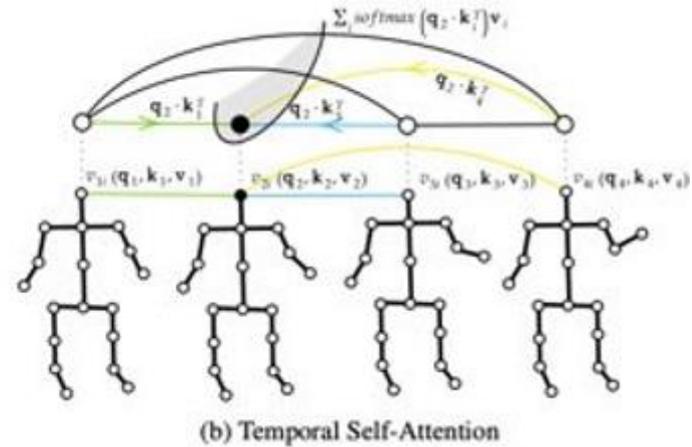


Figure 2

- Connexion en fonction du contexte

Fig 1, 2 : PLIZZARI, Chiara, CANNICI, Marco, et MATTEUCCI, Matteo. Skeleton-based action recognition via spatial and temporal transformer networks. *Computer Vision and Image Understanding*, 2021, vol. 208, p. 103219.

Le pooling :

Généralités

- Flat pooling
- Hierarchical pooling
 - Node drop graph
 - Node clustering pooling

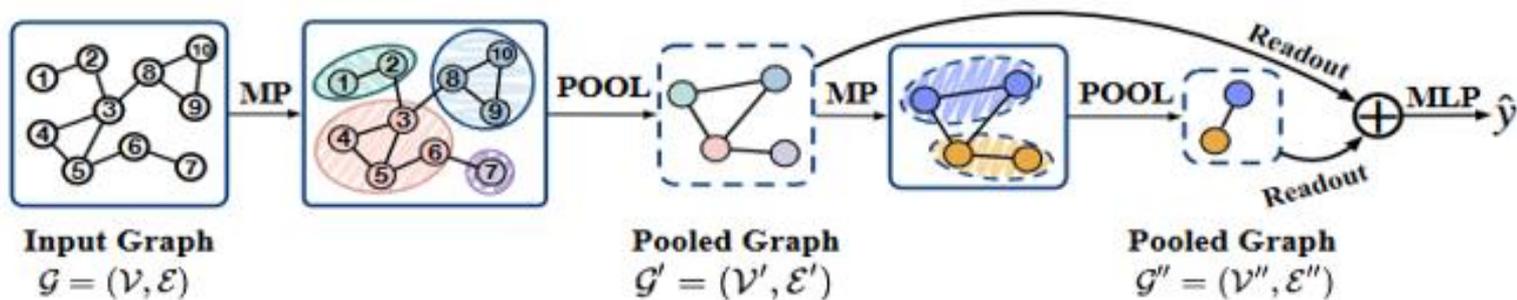


Figure 1

Fig 1 :LIU, Chuang, ZHAN, Yibing, WU, Jia, et al. Graph pooling for graph neural networks: Progress, challenges, and opportunities. arXiv preprint arXiv:2204.07321, 2022

Pooling :

Illustration de Pooling Hiérarchique

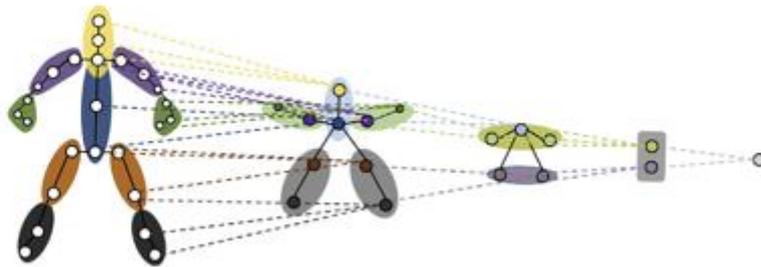


Figure 1

- Basé sur la structure

- Basé sur les parties du corps mis en jeu durant le mouvement

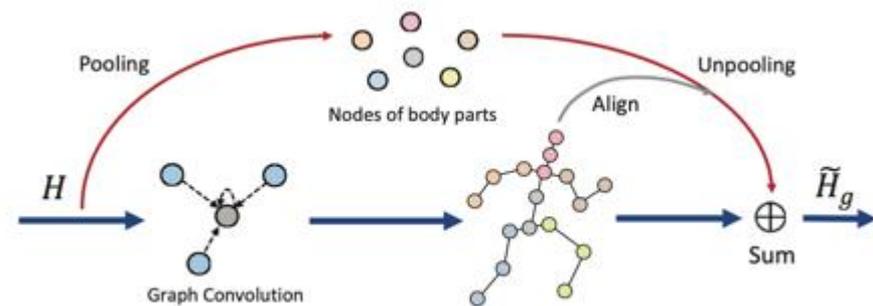


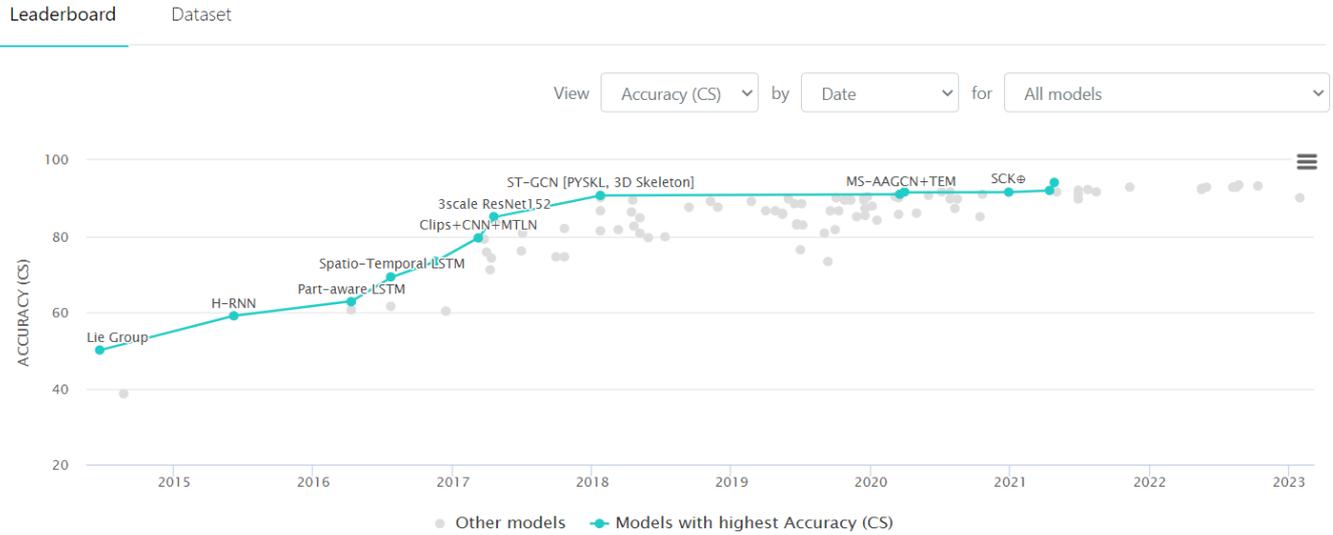
Figure 2

Fig 1: CHEN, Yuxin, MA, Gaoqun, YUAN, Chunfeng, et al. Graph convolutional network with structure pooling and joint-wise channel attention for action recognition. *Pattern Recognition*, 2020, vol. 103, p. 107321.

Des améliorations :

liées à une complexifications des réseaux

Skeleton Based Action Recognition on NTU RGB+D



Conclusion:

Depuis ST-GCN de nombreuses travaux :

- Différents types de données
 - Squelette
 - main
 - Landmark
- Différentes méthodes
 - Convolution
 - Pooling
 - Attention
- Sujet non totalement exploré , en particulier
 - Pooling
 - modèle génératif
- Perspectives
 - Améliorer la généralité.
 - Alléger les réseaux

Questions