

# **Title: Lightweight Graph Neural Networks for Image and Action Recognition**

**CNRS LIP6, UPMC Sorbonne University**

Deep convolutional networks are currently one of the most successful models in image processing and computer vision. Their principle consists in learning convolutional filters, together with attention and fully connected layers, that maximize classification performances. These models are mainly suitable for data sitting on top of regular domains (such as images) [1], but their adaptation to irregular data (namely graphs) requires extending convolutions to arbitrary domains [2, 3]; these extensions are known as graph convolutional networks (GCNs).

Two categories of GCNs exist in the literature: spectral and spatial. Spectral methods [4]–[9] proceed by projecting both input graph signals and convolutional filters using the Fourier transform, and achieve convolution in the Fourier domain, prior to back-project the resulting convolved signal in the input domain. These projections rely on the eigen-decomposition of graph Laplacians whose complexity scales polynomially with the size of the input graphs [16], and this makes spectral GCNs clearly intractable. Spatial methods [10]–[15] instead rely on message passing, via attention matrices, before applying convolution. While spatial GCNs have been relatively more effective compared to spectral ones, their success is highly reliant on the accuracy of the attention matrices that capture context and node-to-node relationships [17]. With multi-head attention, GCNs are more accurate but computationally more demanding, so lightweight variants of these models should instead be considered.

Several methods have been proposed in the literature in order to design lightweight yet effective deep convolutional networks [18]–[21]. Some of them build efficient networks from scratch while others pretrain heavy networks prior to reduce their time and memory footprint using distillation [22]–[24] and pruning [25], [26]. Pruning methods, either unstructured or structured, allow removing connections whose impact on the classification performance is the least perceptible. Unstructured pruning [27], [28] consists in cutting connections individually using different criteria, including weight magnitude, prior to fine-tuning. In contrast, structured pruning [29], [30] aims at removing groups of connections, channels or entire sub-networks. Whereas structured pruning may reach high speed-up on dedicated hardware resources, its downside resides in the rigidity of the class of learnable lightweight networks. On another side, unstructured pruning is more flexible, but may result into topologically inconsistent sub-networks (i.e., either partially or completely disconnected), and this may lead to limited generalization especially at very high pruning rates.

The goal of this thesis subject, is to devise novel approaches for very lightweight GCN design that gathers the advantage of both structured and unstructured pruning, and discards their inconvenient; i.e., the proposed methods should impose constraints on the structure of the learned sub-networks (namely their topological consistency) while also ensuring their flexibility at some extent. The proposed solutions may consider network connections using different criteria (highest magnitudes, connectivity and predefined topologies, etc.) while guaranteeing their accessibility (i.e., their reachability from neural network inputs) and their co-accessibility (i.e., their actual contribution in the evaluation of neural network outputs) [31]. Hence, only topologically consistent sub-networks should be considered when selecting network connections. Applications include image classification and human action recognition in large video collections. We consider both raw videos and already extracted skeleton data described with graphs, where nodes correspond to human joints, and edges to their spatial and temporal relationships.

**Keywords.** Deep machine learning, convolutional neural networks and transformers, lightweight graph convolutional networks, visual scene recognition, image classification, video action recognition.

**Thesis Director:** Hichem SAHBI, CNRS Researcher, HDR, LIP6 Lab, Sorbonne University  
**Contact:** [hichem.sahbi@lip6.fr](mailto:hichem.sahbi@lip6.fr)

**PhD Student Background.** We are seeking a highly motivated PhD candidate, with a preferred background in applied mathematics or computer science with more emphasis on statistics, machine learning and/or image processing, computer vision, and familiar with existing machine learning tools and programming platforms.

## Related bibliography

- [1] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in NIPS, vol. 60, pages 1097–1105, 2012.
- [2] J. Bruna, W. Zaremba, A. Szlam, Y. LeCun. Spectral networks and locally connected networks on graphs. arXiv preprint arXiv:1312.6203 (2013)
- [3] M. Henaff, J. Bruna, Y. LeCun. Deep convolutional networks on graph structured data. arXiv preprint arXiv:1506.05163 (2015)
- [4] TN. Kipf, M. Welling. Semi-supervised classification with graph convolutional networks. In ICLR, 2017
- [5] R. Levie, F. Monti, X. Bresson, M.M. Bronstein. Caylennets: Graph convolutional neural networks with complex rational spectral filters. IEEE Transactions on Signal Processing 67(1), 97–109 (2018)
- [6] R. Li, S. Wang, F. Zhu, J. Huang. Adaptive graph convolutional neural networks. In AAAI, 2018.
- [7] M. Defferrard, X. Bresson, P. Vandergheynst. Convolutional Neural Networks on graphs with Fast Localized Spectral Filtering. In NIPS, 2016
- [8] A. Mazari and H. Sahbi. "MLGCN: Multi-Laplacian graph convolutional networks for human action recognition." The British Machine Vision Conference (BMVC). 2019.
- [9] H. Sahbi. "Learning laplacians in chebyshev graph convolutional networks." Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021.
- [10] M. Gori, G. Monfardini, F. Scarselli. A new model for learning in graph domains. In IEEE IJCNN, vol. 2, pp. 729–734, 2005.
- [11] H. Sahbi. "Kernel-based Graph Convolutional Networks." 2020 25th International Conference on Pattern Recognition (ICPR). IEEE, 2021.
- [12] A. Micheli. Neural network for graphs: A contextual constructive approach. IEEE TNN 20(3), 498-511 (2009)
- [13] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, P.S. Yu. A comprehensive survey on graph neural networks. ArXiv:1901.00596 (2019).
- [14] H. Sahbi, "Lightweight Connectivity In Graph Convolutional Networks For Skeleton-Based Recognition." 2021 IEEE International Conference on Image Processing (ICIP). IEEE, 2021.
- [15] W. Hamilton, Z. Ying, J. Leskovec. Inductive representation learning on large graphs. In NIPS. pp. 1024–1034 (2017)
- [16] Chung, Fan RK, and Fan Chung Graham. Spectral graph theory. No. 92. American Mathematical Soc., 1997.
- [17] Knyazev, Boris, Graham W. Taylor, and Mohamed Amer. "Understanding attention and generalization in graph neural networks." Advances in neural information processing systems 32 (2019).
- [18] Gao Huang, Shichen Liu, Laurens van der Maaten, and Kilian Q. Weinberger, "Condensenet: An efficient densenet using learned group convolutions," in CVPR, 2018.
- [19] M. Sandler, A. G. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in CVPR, 2018.
- [20] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," CoRR, vol. Abs/1704.04861, 2017.
- [21] M. Tan and Q. V. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," in ICML. 2019, vol. 97, PMLR.
- [22] G. E. Hinton, O. Vinyals, and J. Dean, "Distilling the knowledge in a neural network," CoRR, vol. Abs/1503.02531, 2015.
- [23] A. Romero, N. Ballas, S. Ebrahimi Kahou, A. Chassang, C. Gatta, and Y. Bengio, "Fitnets: Hints for thin deep nets," in ICLR, 2015.
- [24] S.-I. Mirzadeh, M. Farajtabar, A. Li, N. Levine, A. Matsukawa, and H. Ghasemzadeh, "Improved knowledge distillation via teacher assistant," in AAAI, 2020.
- [25] Y. LeCun, J. S. Denker, and S. A. Solla, "Optimal brain damage," in NIPS, 1989
- [26] B. Hassibi and D. G. Stork, "Second order derivatives for network pruning: Optimal brain surgeon," in NIPS, 1992.
- [27] S. Han, J. Pool, J. Tran, and W. J. Dally, "Learning both weights and connections for efficient neural network," in NIPS, 2015.
- [28] S. Han, H. Mao, and W. J. Dally, "Deep compression: Compressing deep neural network with pruning, trained quantization and huffman coding," in ICLR, 2016.

- [29] H. Li, A. Kadav, I. Durdanovic, H. Samet, and H. P. Graf, "Pruning filters for efficient convnets," in ICLR, 2017.
- [30] Z. Liu, J. Li, Z. Shen, G. Huang, S. Yan, and C. Zhang, "Learning efficient convolutional networks through network slimming," in ICCV. 2017, IEEE Computer Society.
- [31] H. Sahbi. Lightweight graph convolutional networks with topologically consistent magnitude pruning. In arxiv, march 2022