

Graphs in Computer Vision then and now: how Deep Learning has reinvigorated Structural Pattern Recognition

Donatello Conte

donatello.conte@univ-tours.fr

Laboratoire d'Informatique Fondamentale et Appliquée de Tours (LIFAT - EA 6300)

Université de Tours

Tours, France

ABSTRACT

Computer Vision Problems, such as object detection, object tracking, action recognition and so on, have been, in the past, usually addressed through Statistical Pattern Recognition techniques. SVM, Regression or Neural Networks, are some examples of classical statistical techniques that have been used, quite effectively, in many application contexts of computer vision.

Nevertheless, some attempts have been proposed using more complex data structures (notably graphs) for solving Computer Vision Tasks. However, in terms of performances, their use did not have the same success as techniques based on vector representations. First part of this talk will present some of these proposals, in the context of object tracking ([1]), people re-identification ([3]) and action recognition ([2]). An graph representation is proposed in [1] to deal with occlusion problem. The representation is based on a graph pyramid, namely, each moving region is represented at different levels of resolution using a graph for each level. The algorithm compares the topmost levels of each pyramid in the association phase between moving objects in two consecutive frames. If the comparison outcome is sufficient to assign a label to each node the tracking algorithm stops. Instead, if some ambiguities arise (as it is the case when two objects overlap), the algorithm is repeated using the next levels of the pyramids, until either a consistent labelling is found. The purpose of re-identification (re-id) is to identify people coming back into the field of view of a camera or to recognize an individual through different cameras in a distributed network. At the heart of the process there is a comparison between signatures given probe and gallery sets. In [3] graphs are used to represent people appearance and comparison is done by means of Graph Kernels. Finally, action recognition is a classification problem in which each video representing an action has to be classified with the correct action label. In [2] we proposed to represent videos using graph sequences and proposed a model inspired from bag-of-words techniques to classify a sequence.

Recently, graphs have gained a lot of attention in the Computer Vision community thanks to the use of this kind of data within deep learning techniques. Graph Neural Networks have demonstrated their effectiveness in solving Computer Vision problems, and in

some cases recent proposals have bridged the gap between statistical and structural pattern recognition. Second part of the talk will be devoted to illustrate some of these examples ([4–6]). Starting from the already mentioned applications in Computer Vision (object tracking, action recognition), we will discuss the new proposals based on Deep Learning with graphs and the open problems in this context.

CCS CONCEPTS

• **Computing methodologies** → **Computer vision problems; Machine learning approaches.**

KEYWORDS

Computer Vision, Structural Pattern Recognition, Graph Neural Networks

ACM Reference Format:

Donatello Conte. 2022. Graphs in Computer Vision then and now: how Deep Learning has reinvigorated Structural Pattern Recognition. In *Companion Proceedings of the Web Conference 2022 (WWW '22 Companion)*, April 25–29, 2022, Virtual Event, Lyon, France. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/3487553.3526096>

REFERENCES

- [1] Donatello Conte, Pasquale Foggia, Jean-Michel Jolion, and Mario Vento. 2006. A graph-based, multi-resolution algorithm for tracking objects in presence of occlusions. *Pattern Recognition* 39, 4 (2006), 562–572.
- [2] Xavier Cortés, Donatello Conte, and Hubert Cardot. 2018. Bags of graphs for human action recognition. In *Joint IAPR International Workshops on Statistical Techniques in Pattern Recognition (SPR) and Structural and Syntactic Pattern Recognition (SSPR)*. Springer, 429–438.
- [3] Amal Mahboubi, Luc Brun, and Donatello Conte. 2018. A structural approach to Person Re-identification problem. In *2018 24th International Conference on Pattern Recognition (ICPR)*. IEEE, 1616–1621.
- [4] Akshay Ranges, Pranav Maheshwari, Mez Gebre, Siddhesh Mhatre, Vahid Ramezani, and Mohan M Trivedi. 2021. Trackmpnn: A message passing graph neural architecture for multi-object tracking. *arXiv preprint arXiv:2101.04206* (2021).
- [5] Zonghan Wu, Shirui Pan, Fengwen Chen, Guodong Long, Chengqi Zhang, and S Yu Philip. 2020. A comprehensive survey on graph neural networks. *IEEE transactions on neural networks and learning systems* 32, 1 (2020), 4–24.
- [6] Sijie Yan, Yuanjun Xiong, and Dahua Lin. 2018. Spatial temporal graph convolutional networks for skeleton-based action recognition. In *Thirty-second AAAI conference on artificial intelligence*.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

WWW '22 Companion, April 25–29, 2022, Virtual Event, Lyon, France

© 2022 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-9130-6/22/04.

<https://doi.org/10.1145/3487553.3526096>

AUTHOR



D. Conte received his Ph.D. degree in 2006 by a joint supervision between LIRIS laboratory of the INSA of Lyon (France) and MIVIA laboratory of the University of Salerno (Italy).

He has been an Assistant Professor from 2006 to 2013, in Italy at the University of Salerno. From 2013 to date, he is Associate Professor at the Computer Science Laboratory of the University of Tours.

He is currently head of the Computer Science Department at Polytech Tours School of Engineering.

Currently he is co-head of the RFAI team at the Computer Science Laboratory and he participates, as member and sometimes as local coordinator, to several regional projects on image and video analysis.

His main research fields are: structural pattern recognition (graph matching, graph kernels, combinatorial maps), video analysis (objects detection and tracking, trajectories analysis, behavioral analysis, etc.), and affective computing (emotion recognition, multi-modality analysis for affective analysis, physiological measures by video analysis, etc.).

He is the author of more than 70 publications and reviewers in the main journals in his research field (PAMI, PR, CVIU, TIP, etc.). He is member of the Editorial Board of the Elsevier Journal Internet of Things, MDPI Journal of Imaging and he is Guest Editor for Elsevier Pattern Recognition Letters, IEEE Transactions on Emerging Topics in Computing and Elsevier Signal Processing: Image Communication.

He has been co-chair of the International Workshop on Graph-based Representation in Pattern Recognition (Gbr2019) that was held in France in June 2019. He has been co-chair of the Video Processing for Human Behavioral Analysis (VP-HBA) Track at the 35th ACM Symposium on Applied Computing (SAC 2020) and it is co-chair of the Graph Models for Learning and Recognition (GMLR) Track at the 37th ACM Symposium on Applied Computing (SAC 2022).

Since 2016 he is member of the Governing Board of the French Association for Research in Technical Aids for the Disability (IFRATH). Since 2016 he is member of the Governing Board of the French Association for Pattern Recognition (AFRIF) and he has been association secretary since 2018. Since 2021, he is chairman of the International IAPR Technical Committee 15 (dedicated to the promotion of graphs in the Pattern Recognition).