Post-doctorate 3IA

INRIA Sophia Antipolis, STARS group
2004, route des Lucioles, BP93
06902 Sophia Antipolis Cedex – France
http://www-sop.inria.fr/members/Francois.Bremond/

1. Title: research project (2 years)

Deep Learning for View Invariant Activity Recognition in Videos

2. Scientific context

STARS group works on video analytics. The "SUP" ("Scene Understanding Platform") framework developed in STARS, detects mobile objects, tracks their trajectory and recognizes related behaviors predefined by experts or learnt through video databases. This framework contains several techniques for the recognition of human postures and activities using RGB cameras. However there are scientific challenges in people monitoring when dealing with real word scenes: cluttered scenes, handling wrong and incomplete person segmentation, handling static and dynamic occlusions, low contrasted objects, moving contextual objects (e.g. chairs), viewpoints...

Several investigations have been carried out to model activities of daily living (ADLs) to monitor older adults at home. For instance, Deep Convolutional Neural Network CNN algorithms [3] have been applied with great success to videos, related to monitoring applications. However, existing work has either focused on simple activities in real-life scenarios, or the recognition of more complex (in terms of visual variabilities) activities in hand-clipped videos from a frontal viewpoint. We still lack research on generic methods that can recognize complex activities in a long video recorded in the wild, from any viewpoint. In particular, methods that can handle real-world settings tend to ignore the 3D geometry, the temporal and composite relations of activities and learn short-term activity models directly from 2D pixel data. Hence, latter model cannot recognize long-term or composed activities. In addition, current state-of-the-art algorithms focus on some specific actions (with low intra class variation) like for instance "chopping". Hence, more generic actions like "cooking" can mean either "chopping" or "mixing". Current methods do not perform well on recognizing activities viewed from different viewpoints.

Typical situations that we would like to monitor are Eating and Drinking (how much? how often?) or Cooking (detect behavior that might lead to dangerous situations or non-completion of the task).

The algorithm we want to develop will help senior people and their relatives to feel more comfortable at their home, since scene understanding intends to help at recognizing potentially dangerous situations and reporting to caregivers if necessary.

3. General objectives of the Post-doctorate

This work consists in proposing novel Deep Learning algorithms for view invariant activity recognition in videos, which are efficient in real-world settings, meaning that they can be applied in situations where only small training datasets are available (or using only generic training datasets recorded for different sites) and can dynamically adapt to changing environment.

The performance of the current system strongly relies on the camera angle used: provided that the camera angle used in testing is the same (or close to) as the camera angle used in training, the system performs well. On the contrary, the performance drops when a different camera is used. In this work package, we aim at improving the performance by using 3D human pose information. For the extraction of the pose information CMU's openpose (https://github.com/CMU-Perceptual-Computing-Lab/openpose) but any other more performing software can be used. To generate extra views, Generative Adversarial Network (GAN) [2, 4] can be used together with the human pose information to complete the training dataset. Another promising solution would be to train a GAN using sparse representations of human poses [6]. Then, the newly generated poses can be transformed to human images by means of a pose transfer algorithm [7]. The validation dataset will include NTU RGB+D [5] and smart-home [1] datasets.

4. Bibliography

1. *S. Das, R. Dai, M. Koperski, L. Minciullo, L. Garattoni, F. Bremond and G. Francesca. Toyota Smarthome: Real-World Activities of Daily Living with supplementary. In Proceedings of the 17th International Conference on Computer Vision, ICCV 2019, in Seoul, Korea, October 27 to November 2, 2019.*
2. *Y. Wang, A. Dantcheva,... and F. Bremond. ImaGINator: Conditional Spatio-Temporal GAN for Video Generation. In Proceedings of the IEEE Winter Conference on Applications of Computer Vision, WACV 2020, Snowmass village, Colorado, March 2-5, 2020.*
3. *R. Dai, L. Minciullo, L. Garattoni, G. Francesca and F. Bremond. Self-Attention Temporal Convolutional Network for Long-Term Daily Living Activity Detection. In Proceedings of the 14th IEEE International Conference on Advanced Video and Signal-Based Surveillance, AVSS 2019, in Taipei, Taiwan, 18-21 September 2019.*
4. *Lakhal, M. I., Lanz, O., & Cavallaro, A. View-LSTM: Novel-View Video Synthesis Through View Decomposition. In Proceedings of the IEEE International Conference on Computer Vision ICCV (pp. 7577-7587), 2019.*
5. *A. Shahroudy, J. Liu, T.-T. Ng, and G. Wang. NTU-RGB+D: A large scale dataset for 3d human activity analysis. In the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2016*
6. *A. Ben Tanfous, H. Drira, B. Ben Amor. Coding Kendall's Shape Trajectories for 3D Action Recognition. In the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2018*
7. *Z. Zhu, T. Huang, … and X. Bai. Progressive Pose Attention Transfer for*

Person Image Generation. In the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2019

The application file should include:

- a detailed CV,
- a letter of motivation from the candidate related to the subject of the post-doc topic,
- the PDF of the thesis manuscript and some articles published as the first author,
- all documents deemed useful for evaluating the candidate's profile (letters of recommendation, etc.).

Contact:

François Brémond <francois.bremond@inria.fr>
Sophia-Antipolis,
Research Director
Head of INRIA STARS Team
http://www-sop.inria.fr/members/Francois.Bremond/