A propos du centre ou de la direction fonctionnelle

The Inria Sophia Antipolis - Méditerranée center counts 34 research teams as well as 7 support teams. The center's staff (about 500 people including 320 Inria employees) is made up of scientists of different nationalities (250 foreigners of 50 nationalities), engineers, technicians and administrative staff. 1/3 of the staff are civil servants, the others are contractual agents. The majority of the center's research teams are located in Sophia Antipolis and Nice in the Alpes-Maritimes. Four teams are based in Montpellier and two teams are hosted in Bologna in Italy and Athens. The Center is a founding member of Université Côte d'Azur and partner of the I-site MUSE supported by the University of Montpellier.

Contexte et atouts du poste

Assignments:

- With the help of F. Bremond, the recruited person will develop a new algorithm.
- For a better knowledge of the proposed research subject:
  - A state of the art, bibliography and scientific references are available at the following URL, do not hesitate to log in: http://www-sop.inria.fr/members/Francois.Bremond/

Mission confiée

STARS group works on automatic sequence video interpretation. The “SUP” (“Scene Understanding Platform”) Platform developed in STARS, detects mobile objects, tracks their trajectory and recognises related behaviors. The supervisor of the first step is to extract meaningful mid-level features that can be further processed thanks to more long-term reasoning based on TCN or Transformers or even ontology-based knowledge elicitation (infusion) using student-teacher paradigm, contrastive learning and co-training. Several levels of ground truth (GT) supervision (e.g. weak-supervision) will be used to train the model.

Existing work has either focused on simple activities in real-life scenarios, or the recognition of more complex (in terms of visual variability) activities in hand-clipped videos with well-defined temporal boundaries. We still lack research on methods that can retrieve several instances of complex activity in a continuous video (untrimmed) flow of data. Existing methods that perform in online scenarios that can reason about the temporal and composite relations characterizing complex activities generally cannot handle uncertainty and tend to underperform in real life scenarios. Moreover, they have difficulties to distinguish similarly looking activities.

On the other hand, these methods are mostly dedicated to action detection and ignore the emotion component.

An emotion is a mental state that arises spontaneously and is often accompanied by cognitive, physical and physiological changes. Due to the complexity of human reactions, recognizing emotions is still limited to our knowledge and remains the target of many relevant scientific researches. Literature, the recognition of human behaviours, especially from facial expressions, often rely on the interpretation of dynamic scenes observed by video cameras. The accuracy of computer vision (CV) algorithms, as in the case of CNN, is typically limited by the identification of real emotion. A person may be happy even if she is not smiling and people differ widely in how expressive they are in showing their inner emotions. Recent multimodal sentiment analysis approaches focus on deep neural networks and propose multi-sensor data fusion methods. As emotions are complex set of reactions with multiple components (a), the idea is to compare/infuse/combine salient information from different modalities, coming from video cameras and biosensors. To lift the ambiguity, bio-signals (or Galvanic Skin Conductance (GSC) or electrodermal activity (EDA), ECG, EMG, Respiration Rate, etc.) with be used as ground truth (GT) for emotion.

Principales activités

This work consists in the improvement of Emotion Recognition/Detection algorithms using RGB video cameras at test time, but using multi-modalities at training time.

The objective is to develop and test a model on multiple datasets with various modalities to identify specific emotions, such as stress, anxiety, joy. The approach will consist of advanced Deep Learning methods for combining multimodal inputs, comparing various strategies such as multi-task learning, Knowledge Elicitation (infusion) using Student-Teacher paradigm, contrastive learning and co-training or Transformer. Several levels of ground truth (GT) supervision (e.g. weak-supervision) will be used to trained the model.

Typical pipeline can combine CNNs for 3D pose, eye-gaze and facial expression estimation, depending on the emotions to detect. Short temporal aspects of the actions can be handled through RNN or 3DCNN. The objective of the first step is to extract meaningful mid-level features that can be further processed thanks to more long-term reasoning based on TCN or Transformers or even ontology-based reasoning.

A challenge will be to propose an approach to leverage the knowledge acquisition process and the long-term reasoning with a weakly supervised setting.

This work aims at reducing the supervision in order to conceive a general and robust algorithm enabling the detection of the emotions of an individual (together with his/her facial expressions) living in an unconstrained environment and observed through a limited number of sensors (restricting to a single video camera).

To validate the work, we will assess the proposed approaches on videos from a set of applications in

Informations générales

- Thème/Domaine: Vision, perception et interprétation multimedia
- Ville: Sophia Antipolis
- Centre Inria: CRI Sophia Antipolis - Méditerranée
- Date de prise de fonction souhaitée: 2021-10-01
- Durée de contrat: 3 ans
- Date limite pour postuler: 2021-05-15

Contacts

- Equipe Inria: STARS
- Directeur de thèse: Brémond François / Francois.Bremond@inria.fr

A propos d'Inria

Inria est l'institut national de recherche dédié aux sciences et technologies du numérique. Il emploie 2600 personnes. Ses 200 équipes-projets agiles, en général composées de 50 personnes, travaillent en partenariat avec de nombreux scientifiques et chercheurs du monde entier.

L'essentiel pour réussir

Computer Vision, Strong background in C++/Python programming, Linux, Deep Neural Network frameworks (PyTorch, TensorFlow, Keras).

Knowledge on the following topics is a plus:

- Machine learning
- Probabilistic Graphical Models and Optimization techniques
- Mathematical (Geometry, Graph theory, Optimization),
- Artificial intelligence
- Image processing and 3D Vision

Consignes pour postuler

Before applying, it is strongly recommended that you contact the Scientific manager beforehand.

Sécurité défense:

Ce poste est susceptible d'être affecté dans une zone à régime restrictif (ZRR), telle que définie dans le décret n°2011-1425 relatif à la protection du potentiel scientifique et technique de la nation (PPST). L'autorisation d'accès à une zone est délivrée par le chef d'établissement, après avis ministériel favorable, tel que défini dans l'arrêté du 03 juillet 2012, relatif à la PPST. Un avis ministériel défavorable pour un poste affecté dans une ZRR aurait pour conséquence l'annulation du recrutement.

Politique de recrutement:

Dans le cadre de sa politique diversité, tous les postes Inria sont accessibles aux personnes en situation de handicap.

Attention: Les candidatures doivent être déposées en ligne sur le site Inria.

Le traitement des candidatures adressées par d'autres canaux n'est pas garanti.
collaboration with Nice Hospital, such as the ones related to the monitoring of patients (e.g. autistic, dementia, depressed) with behavioral disorders.

1st year:
- Study the limitations of existing emotion recognition/detection algorithms.
- Depending on the targeted emotions, data collection might need to be carried out.
- Propose an original algorithm that addresses current limitations on detection.
- Evaluate the proposed algorithm on benchmarking datasets.
- Write a paper

2nd year:
- Investigation of feasibility/appropriateness of the framework in practical situations.
- Propose an algorithm to address model learning task in weakly-supervised settings.
- Write a paper

3rd year:
- Optimize proposed algorithm for real-world scenarios.
- Write a paper, and
- PhD Manuscript

Avantages
- Subsidized meals
- Partial reimbursement of public transport costs
- Leave: 7 weeks of annual leave + 10 extra days off due to RTT (statutory reduction in working hours) + possibility of exceptional leave (sick children, moving home, etc.)
- Possibility of teleworking (after 6 months of employment) and flexible organization of working hours
- Professional equipment available (videoconferencing, loan of computer equipment, etc.)
- Social, cultural and sports events and activities
- Access to vocational training
- Social security coverage

Rémunération
Duration: 36 months
Location: Sophia Antipolis, France
Gross Salary per month: 1982€ per month (year 1 & 2) and 2085€ per month (year 3)