



Offre n°2025-08731

PhD Position F/M pre-doc position / Deep Neural Networks for Analyzing Non-Verbal Behavior during Clinical Interactions

Le descriptif de l'offre ci-dessous est en Anglais

Type de contrat : CDD

Niveau de diplôme exigé : Bac + 5 ou équivalent

Fonction : Chercheur contractuel

A propos du centre ou de la direction fonctionnelle

The Inria centre at Université Côte d'Azur includes 42 research teams and 9 support services. The center's staff (about 500 people) is made up of scientists of different nationalities, engineers, technicians and administrative staff. The teams are mainly located on the university campuses of Sophia Antipolis and Nice as well as Montpellier, in close collaboration with research and higher education laboratories and establishments (Université Côte d'Azur, CNRS, INRAE, INSERM ...), but also with the regional economic players.

With a presence in the fields of computational neuroscience and biology, data science and modeling, software engineering and certification, as well as collaborative robotics, the Inria Centre at Université Côte d'Azur is a major player in terms of scientific excellence through its results and collaborations at both European and international levels.

Contexte et atouts du poste

Inria, the French National Institute for computer science and applied mathematics, promotes “scientific excellence for technology transfer and society”. Graduates from the world’s top universities, Inria's 2,700 employees rise to the challenges of digital sciences. With its open, agile model, Inria is able to explore original approaches with its partners in industry and academia and provide an efficient response to the multidisciplinary and application challenges of the digital transformation. Inria is the source of many innovations that add value and create jobs.

Team

The STARS research team combines advanced theory with cutting edge practice focusing on cognitive vision systems.

Team web site : <https://team.inria.fr/stars/>

Mission confiée

The Inria STARS team is seeking for a pre-doc researcher with strong background in computer vision, deep learning and machine learning.

“Actions speak louder than words”. Humans are complex beings, and they often convey a wealth of information not through their words but through their actions and demeanor. Non-verbal behaviors can offer crucial insights into their emotional state, pain level, or anxiety, often more eloquently than words alone. The analysis of non-verbal communication is of critical importance in the diagnostic landscape. Decoding non-verbal cues in a clinical setting requires healthcare professionals to be astute observers, picking up on nuances that may be subtle yet critical. The challenge lies in accurately interpreting these cues, as they can vary greatly from one individual to another.

To address this challenge, automated systems capable to detect non-verbal behaviors and their corresponding meanings can assist healthcare providers. Such technology is not to replace medical experts but rather to act as their supportive tool.

The primary objective of this technical internship is to lead the development of an advanced AI model for Human Behavior Understanding to identify non-verbal cues expressed by patients, and then interpreting the cues to derive critical insights about their health. Traditionally, computer vision methodologies encompassing skin color analysis, shape analysis, pixel intensity examination, and anisotropic diffusion were used to identify body parts and trace their activities. However, these algorithms provided limited flexibility because of their domain-specific nature. Deep learning methods can be used to deal with this issue as they offer more training flexibility, and better performance results. The overarching goal is to provide a real-time, data-driven analysis of non-verbal cues exhibited by patients during clinical interactions, thereby delivering invaluable insights to healthcare practitioners.

Principales activités

With our vision of evidence-based diagnosis, we will develop explainable methods for biomarker detection from audiovisual and physiological data. Generally, AI models are based on machine learning concepts that find intrinsic correlations between multiple input channels and the true labels. To be able to model the complex action patterns, we need to go beyond deep learning by incorporating some semantic modeling within the deep learning pipeline, which today consists of a combination of CNN and transformers. These complex action patterns include composite actions and concurrent actions occurring in long untrimmed videos. Existing methods have mostly focused on modeling the variation of visual cues across time locally or globally within a video. However, these methods consider the temporal information without any further semantics. Videos may contain rich semantic information such as objects, actions, and scenes. Real-world videos contain also many complex actions with inherent relationships between action classes at the same time steps or across distant time steps. Modeling such class-temporal relationships can be extremely useful for locating actions in those videos. Therefore, semantic relational reasoning can help determine the action instance occurrences and locate the actions in the video, especially for complex actions in the video.

Going beyond classical deep CNNs, our first attempts will be to extract the relevant semantics using large language-vision models (LVMs). However, large foundation models work really well and have almost pixel-level attention, although they are not scalable. Their monstrous size makes it hard to fine-tune. What we will do is instead of learning temporal relations from scratch we will exploit the optical flow of attention maps and its information of motion on a feature level, which does not require much processing to classify actions. This optical flow is obtained using the attention maps from processed frames of videos using image foundation models. Adapters have shown to work well and provide a downsampled embedding of the hidden layers of the base model which is easy to work with. We intend to move towards the direction of designing plugin architectures that makes large transformer models more efficient by omitting fine-tuning of the whole models and other additions.

Compétences

Candidates must hold a Master degree or equivalent in Computer Science or a closely related discipline by the start date.

The candidate must be grounded in the basics of computer vision, have solid mathematical and programming skills.

Preferably in Python, OpenCV, deep learning framework Pytorch or Tensorflow.

The candidate must be committed to scientific research and strong publications.

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 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
- Contribution to mutual insurance (subject to conditions)

Rémunération

Gross Salary per month: 2200€ brut per month

Informations générales

- **Thème/Domaine** : Vision, perception et interprétation multimedia
Biologie et santé, Sciences de la vie et de la terre (BAP A)
- **Ville** : Sophia Antipolis
- **Centre Inria** : [Centre Inria d'Université Côte d'Azur](#)
- **Date de prise de fonction souhaitée** : 2025-07-01
- **Durée de contrat** : 12 mois
- **Date limite pour postuler** : 2025-04-19

Contacts

- **Équipe Inria** : [STARS](#)
- **Recruteur** :
Balazia Michal / michal.balazia@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 215 équipes-projets agiles, en général communes avec des partenaires académiques, impliquent plus de 3900 scientifiques pour relever les défis du numérique, souvent à l'interface d'autres disciplines. L'institut fait appel à de nombreux talents dans plus d'une quarantaine de métiers

différents. 900 personnels d'appui à la recherche et à l'innovation contribuent à faire émerger et grandir des projets scientifiques ou entrepreneuriaux qui impactent le monde. Inria travaille avec de nombreuses entreprises et a accompagné la création de plus de 200 start-up. L'institut s'efforce ainsi de répondre aux enjeux de la transformation numérique de la science, de la société et de l'économie.

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.

Consignes pour postuler

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.