Offre n°2023-06755

PhD Position F/M PhD Position Computer Vision / Deep Learning: Video Generation

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

**Type de contrat :** CDD

**Niveau de diplôme exigé :** Bac + 5 ou équivalent

**Fonction :** Doctorant

A propos du centre ou de la direction fonctionnelle

The Inria Université Côte d'Azur center counts 36 research teams as well as 7 support departments. 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

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/](https://team.inria.fr/stars/)

Mission confiée

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

The candidate is expected to conduct research related to generative models, including the development of computer vision algorithms for image and video generation.

Principales activités

Despite remarkable progress in generative models, a pretrained network is currently limited in being able to generate only a single training subject / object within a single scenario the training data was pertained to.

This Ph.D. thesis aims at bringing video generation to the next level by proposing strategies to generalize the generation ability of generative models by disentangling **appearance** and **motion** in the latent space and further disentangling **motion** in primary directions, applicable to any subject in any setting. This carries the premise of allowing for more complex settings incorporating interaction of subjects / objects.

Context:

Generative adversarial networks (GANs) [1] have witnessed increased interest from academia and industry, due to exceptional capacity in generating highly realistic images [2, 3, 4, 5, 6, 7]. Videos signify more complex data, due to the additional temporal dimension. While some research works showed early results in video generation [8-11], there are many open questions in the field.

1. Model architecture

The thesis firstly will investigate, how to design model architecture for generator and discriminator in
generative models. We will explore traditional model architectures such as CNN and RNN, as well as
Transformer-based generators. Our objective will be to explore whether we can design a unified model
architecture that generalizes over categories, such as human bodies and faces. We will study how to
connect different architectures, in order to create such a general system for cross-category generation.

2. 3D-aware generation

Learning 3D-aware models from 2D data has become a popular research topic in image generation. In
this thesis, we will go one step further in this direction to explore novel view synthesis in video
generation. We intend to combine jointly state-of-the-art novel view synthesis techniques with video
generation, aiming at creating 3D-aware video generation. Our idea is to explore implicit representation
(e.g., NeRF), explicit representation (e.g., 3D representation), as well as hybrid (implicit-explicit)
representation in video generation models. One objective will be to design an efficient and effective
representation for novel-view synthesis in video generation.

3. Generalizability

Finally, we will aim to design a universal model which is able to generate videos across categories. Most
of current models focus on generating single category (e.g., faces, sky...). Currently, there is no models,
which are able to generate complex multi-category videos (e.g. Kinetics-600). We plan to increase the
complexity of video generative models and design a large-scale video GAN. The objective is to study
whether big generative models are able to capture the distribution of complex video datasets and create
semantic meaningful videos.

2680.
4401–4410.
et al., “Photo-realistic single image super-resolution using a generative adversarial network.” in CVPR,
2017.

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: 2082€ brut per month (year 1 & 2) and 2190€ brut/month (year 3)
Informations générales

- **Thème/Domaine**: Vision, perception et interprétation multimedia
- **Ville**: Sophia Antipolis
- **Centre Inria**: Centre Inria d'Université Côte d'Azur
- **Date de prise de fonction souhaitée**: 2022-12-01
- **Durée de contrat**: 3 ans
- **Date limite pour postuler**: 2023-11-30

Contacts

- **Équipe Inria**: STARS
- **Directeur de thèse**: Dantcheva Antitza / antitza.dantcheva@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.