Offre n°2024-08245

M2 Internship: Zero-shot deepfake detection

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

Type de contrat : Convention de stage

Niveau de diplôme exigé : Bac + 5 ou équivalent

Fonction : Stagiaire de la recherche

A propos du centre ou de la direction fonctionnelle

The Inria center at the University of Rennes is one of eight Inria centers and has more than thirty research teams. The Inria center is a major and recognized player in the field of digital sciences. It is at the heart of a rich ecosystem of R&D and innovation, including highly innovative SMEs, large industrial groups, competitiveness clusters, research and higher education institutions, centers of excellence, and technological research institutes.

Contexte et atouts du poste

Deepfakes are synthetic media created using typically deep learning techniques, to manipulate or generate realistic audio, video, or images of people. While deepfakes have potential applications in entertainment and creative industries, they also pose serious threats. In case of videos, manipulation of people is particularly studied. Face swapping aims to replace the face in an original image or video with a selected one. It thus changes the identity (ID-Replaced). Face reenactment and lip synching are other types of video manipulation that aim at modifying the facial movements and expressions of a target person according to another one for example, or even according to some audio record. In this case the identity is not changed (ID-Remained).

While there is an increasing number of approaches for detecting whether an image or video has been the result of AI-based synthesis or manipulation, a key issue that detectors face is the generalization to unseen generative architectures. A significant hurdle in the development of robust deepfake detection systems is the limited availability of diverse datasets for training and evaluation purposes.

A possible solution to gain generalization and robustness is to shift to a different paradigm, where models are only trained on real videos, with the goal to detect manipulated videos based on their anomalous behavior [1]. Recent work has focused on verifying audio-visual consistency with different strategies [2-5].

On the other hand, few studies explore an approach that modelize not just real videos, but more precisely the real videos of a particular individual [6-9]. However those approaches have not been evaluated on same datasets, and do not provide sufficient details on performance for different types of manipulation (ID-replaced or ID-remained), robustness to degradation like compression or blurring, and false positives, that are a key issue of synthetic media detection methods because they compromise the trust of journalists and fact-checkers on their results.

Mission confiée

The idea is to combine these 2 approaches, i.e. the possibility of modeling a person, and proposing a solution to detect manipulations independent of the modification technique. The aim of the internship is to answer the following questions:

- To what extent are ID-Remained cases are more difficult than ID-Replaced ones?
- Are the behavioral characteristics proposed in the literature really relevant, discriminative and robust?
- How robust are these methods to low quality and high compression?
- What is the minimum number of videos needed to learn the model of an individual?

Principales activités

- Identify in the literature the most promising features and zero-shot methods
- Set up a robust evaluation framework to assess the various aspects of performance
- Evaluate the performances of selected methods, either existing ones or proposed ones
- Results analysis to pinpoint difficulties of both the dataset and the methods

Compétences

- Master in Computer Sciences, with proficiency in python and its libraries for deep learning
- General background in computer vision and machine learning
- Understanding of deep learning methodologies and techniques;
- Proficiency in data handling, particularly in video processing

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

Brut mensuel de 650 euros.

Informations générales

- **Thème/Domaine :** Sécurité et confidentialité
- **Ville :** Rennes
- **Centre Inria :** Centre Inria de l'Université de Rennes
- **Date de prise de fonction souhaitée :** 2024-03-01
- **Durée de contrat :** 4 mois
- **Date limite pour postuler :** 2024-11-17

Contacts

- **Équipe Inria :** ARTISHAU
- **Recruteur :**
  Kijak Ewa / Ewa.Kijak@irisa.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.

L'essentiel pour réussir

Candidates are invited to submit a clear CV together with:

- a letter detailing their education, experience (internships, etc.), and explaining how they match the profile sought
- their Master's transcripts (list of marks, even preliminary)

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.