



PhD Position

Multimedia, Multimodal and Explainable Fact-Checking

Keywords

Computer Vision, Multimodality, Artificial Intelligence, Data Journalism

Context

This thesis falls within the Fact Checking project carried out since 2021 between the LIFAT laboratory and the PRIM team [1][2][3]. Fact checking consists in verifying the veracity of facts presented in the media by public figures or other organisations. A great deal of research has been and is being carried out [4][5][6] to address the major challenge of disinformation, especially since the advent of deep-fake and generative AI technologies. This phenomenon of massive disinformation can be observed in several areas, for instance the political sphere ("political discourse"), climate change or the health sector ("COVID-19 misinformation"). Over the last decade, significant progress has been made in the fields of Computer Vision (CV) and Natural Language Processing (NLP), largely thanks to the success of neural networks. Thus, it is quite natural that the multimodal aspect (joint learning of modalities) combining CV and NLP have been studied recently, with success in tasks such as visual question answering (VQA), image or video captioning, etc.

Goals

This thesis aims at the design and evaluation of new **multimodal** detection techniques (video, image, audio, text) in **multimedia** content (video, audio, text). To do this, the thesis will explore the use of AI concepts at different levels: Audio-visual Active Speaker identification (ASD), Multimodal Named Entity recognition (MNER), Retrieval-Augmented Generation (RAG), etc., and the exploitation of an existing database of verified facts (database from [1], [3] or existing benchmarks).

Furthermore, given the societal impact of this work, a strong emphasis will be placed on the **explainability** of the proposed multimodal algorithms, by aggregating different solutions: retrieval of third-party information via information search, text generation, visualisation of salient elements in images/videos, isolation of sequences of interest in audio/video tracks, etc.

Finally, during this research work, locks of the citizens' acceptance to acknowledge and use such a tool will also be taken into account, via UX design and ergonomics considerations. These aspects will be addressed via a collaboration with researchers in information and communication sciences, journalists and fact-checkers.

Duration: 3 years, starting from October 2025

Affiliation : Computer Science Laboratory of Université de Tours (<u>LIFAT</u>), Pattern Recognition and Image Analysis Group (<u>RFAI</u>)

Grant: around 1600€-1700€/month





Supervisors:

- Nicolas RAGOT https://www.univ-tours.fr/annuaire/m-nicolas-ragot
- Frédéric RAYAR <u>http://frederic.rayar.free.fr/</u>
- Mathieu DELALANDRE http://mathieu.delalandre.free.fr/

Skills

- M2 in Computer Science,
- good experience in data analysis and machine learning (theory and practice) is required,
- good experience in programming and software development,
- good communication and writing skills (French knowledge is not mandatory but preferable),
- curiosity and ability to work in collaboration with scientists from other fields,
- autonomy and good organisation skills.

How to candidate:

Send the following documents by e-mail to **frederic [dot] rayar [at] univ-tours [dot] fr** before 30th of March:

- a resume,
- a motivation letter,
- a short description of your experiences related to computer vision, natural language processing, machine/deep learning,
- an internship report or scientific publication,
- references from academics.

Bibliography (selection)

[1] <u>Rayar F</u>. FrenchFacts: A French Dataset of Fact-Checked Claims. In Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING). **2024** (Accepted).

[2] <u>Rayar F</u>. Fact-Checked Claim Detection in Videos Using a Multimodal Approach. In Proceedings of the 19th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISAPP), 614-620. **2024**.

[3] <u>Rayar F.</u>, <u>Delalandre M</u>., and Le V-H. A large-scale TV video and metadata database for French political content analysis and fact-checking. In Proceedings of the 19th International Conference on Content-based Multimedia Indexing (CBMI), 181–185. **2022**

[4] Guo, Z., Schlichtkrull, M., and Vlachos, A. A survey on automated fact-checking. Transactions of the Association for Computational Linguistics, 10:178. **2023**.

[5] Akhtar M., Schlichtkrull M., Guo, Z., Cocarascu, O., Simper E., and Vlachos, A. Multimodal automated factchecking: A survey. **2023**

[6] Kotonya, N. and Toni, F. Explainable automated fact-checking: A survey. CoRR, abs/2011.03870. 2020