Caroline Mazini Rodrigues, MSc

PhD candidate in Computer Science

Expected degree 2nd Semester 2024

+33 0753823009 | carollmazini@gmail.com | Paris, France
linkedin.com/in/caroline-mazini-rodrigues-41787427 | github.com/carolmazini | carolmazini.github.io/

Skills

Programming languages: Python, C / C++, Java.
Machine learning: Deep learning, explainable artificial intelligence (xAI), interpretability of neural networks, supervised learning, unsupervised learning, computer vision, generative AI.
Machine learning tools: Pytorch, tensorflow, hugging face, keras, captum, scikit-learn, pytorch-lightening, wandB.
Image processing tools: OpenCV, scikit-image, pillow.
Computational Theory and Mathematics: Signal mathematics, algorithms and complexity, rational languages' theory.
Information Retrieval: Content-based image retrieval, text and image representation.
Research: Presentation, planning, creative problem-solving, teamwork, active listening, adaptability, analytical thinking.
Peer Review: Critically evaluate and provide constructive feedback on academic work.

Experience

Research and teaching associate, *Université Gustave Eiffel* **France** *09/2023 - 09/2024*

- Improved the **interpretability of CNN** decisions by devising a method that employs hierarchical segmentation to identify important image components for the model, resulting in a **human-based bias detection** rate increase of at least 20% points than baseline methods.
- Teaching **improved communication proficiency and knowledge** in topics such as Python/C programming, Image processing, and Databases to undergrad and master's students.

Research in xAI, PhD fellowship at LRE, *EPITA* **France** *09/2020 - 08/2023*

- **Developed two methods within the field of explainable AI (xAI)**. The first method, published in Pattern Recognition Letters, aimed to improve the interpretability of gradient-based xAI techniques by enhancing the visualization of important features using artificial class distancing with support regression networks. This approach successfully **reduced visualization complexity** in 5 out of 4 tested methods **without compromising sensitivity** to the model's knowledge, thereby increasing interpretability while maintaining accuracy. The second method, published in Information Sciences, introduced a framework for globally explaining CNNs through **concept decomposition**, which **improved concept localization** in images and facilitated **bias detection for both AI experts and non-experts** in 77% of the cases.
- Teaching courses to undergraduate and master's students **increased proficiency in complex topics** such as signal mathematics, algorithm complexity, rational languages theory, and Python for databases.

Data Scientist, *Neuralmind* **Brazil** *02/2020 - 08/2020*

- Designed applications using deep learning models for tasks such as **image and text recognition and processing**. In the image domain, employed deep learning models for object identification and bounding box attribution. In the text domain, utilized **Transformers models** to extract contextual information from unstructured texts.

Research in Complex Data Analysis, *University of Campinas (Unicamp)* **Brazil** *02/2018 - 08/2020*
FAPESP – São Paulo Research Foundation – Master's fellowship at RECOD.

- Contributed to a **big data project** by developing a method to retrieve the most representative images from a collection of **social media images** related to forensic events. This work was an integral part of the DéjàVu project's pipeline, which aimed to understand forensic events using data from diverse sources. The contribution enhanced the project's forensic capabilities by improving its ability to analyze and interpret data effectively. The methodology successfully retrieved small sets of key images from unbalanced datasets with high precision and recall.
- **Curated 3 datasets** for benchmarking purposes: two focused on forensic events and one on a general event.
- Tutored in the Complex data mining extension course, focusing on information retrieval, supervised learning, and unsupervised learning. Additionally, tutored the undergraduate course in algorithms and computer programming.

Education

PhD in Computer Science, *Université Gustave Eiffel* **France** *09/2020 - 09/2024*
Supervisors: Professor Laurent Najman (LIGM) and Dr. Nicolas Boutry (LRE).

MSc in Computer Science, *University of Campinas (Unicamp)* **Brazil** *02/2018 - 08/2020*
Supervisors: Professor Zanoni Dias (LOCo) and Professor Anderson Rocha (RECOD).

BSc in Computer Science, *São Paulo State University (UNESP)* **Brazil** *02/2013 - 08/2017*
Exchange program participation: *University of Glasgow* **Scotland** *01/2016 - 06/2016*

Publications

Complete Journal Articles

- RODRIGUES, CAROLINE MAZINI; BOUTRY, NICOLAS; NAJMAN, LAURENT. Reasoning with trees: interpreting CNNs using hierarchies. Arxiv, 2024.
- RODRIGUES, CAROLINE MAZINI; BOUTRY, NICOLAS; NAJMAN, LAURENT. Unsupervised discovery of Interpretable Visual Concepts. Information Sciences, 2024.
- RODRIGUES, CAROLINE MAZINI; BOUTRY, NICOLAS; NAJMAN, LAURENT. Transforming gradient-based techniques into interpretable methods. Pattern Recognition Letters, 2024.
- RODRIGUES, C. M.; SORIANO-VARGAS, A.; BAHRAM, L.; ROCHA, A.; DIAS, Z.. Manifold Learning for Real-World Event Understanding. IEEE Transactions on Information Forensics and Security. 2021.
- PADILHA, R.; RODRIGUES, C. M.; ANDALO, F. A.; BERTOCCO, G.; DIAS, Z.; ROCHA, A. . Forensic Event Analysis: From Seemingly Unrelated Data to Understanding. IEEE SECURITY & PRIVACY. 2020.

Conference Proceedings

- DOH, M.; RODRIGUES, C. M.; BOUTRY, N.; NAJMAN, L.; MANCAS, M.; BERSINI, H. Bridging Human Concepts and Computer Vision for Explainable Face Verification. BEWARE-23 Joint Workshop @ AIxIA, 2024.
- RODRIGUES, C. M.; BOUTRY, N.; NAJMAN, L. . Gradients Intégrés Renforcés. Explain'AI Conférence Francophone sur l'extraction et la gestion des connaissances (EGC). 2023.
- RODRIGUES, C. M.; PEREIRA, L. ; ROCHA, A. R. ; DIAS, Z. . Image Semantic Representation for Event Understanding. 2019 IEEE International Workshop on Information Forensics and Security (WIFS). 2019.
- RODRIGUES, C. M.; PITERI, M. A.; ARTERO, A. O.; ELER, D. M.; SILVA, F. A.; PEREIRA, D. R. . Facial Recognition in Digital Images using Local Binary Pattern Methods. XIII Workshop de Visão Computacional (WVC). 2017. v. 1.

Awards

Best presentation – PhD day MSTIC (2021).
Academic Merit, São Paulo State University – UNESP (2017).
Honorable Mention by presenting the work: Neper Number Origin Based on its Derivative – UFU (2015).

Events participation

MLx Generative AI – Oxford Machine Learning Summer School (OxML). 2024. (Summer school).
EuADS Data Science for Explainable and Trustworthy AI. 2023. (Summer school).
Explain'AI (EGC) – Presentation "Gradients Intégrés Renforcés". 2023. (Workshop).
Oxford Machine Learning Summer School (OxML). 2022. (Summer school).
École Jeune chercheu/r/se/s en Informatique Mathématique – Presentation "Visual xAI techniques". 2022. (Summer school).
Latin American Meeting In Artificial Intelligence (KHIPU) – Presentation "Complex Data Relevance Analysis for Event Detection". 2019. (Meeting).

Online Courses & Certifications

Speaking to inform: Discussing complex ideas (Feb. 2022) – University of Washington – Coursera.
Introduction to Public Speaking (Dec. 2021) – University of Washington – Coursera.
Practical Peer Review (May 2021) – Publons Academy.

Languages

Portuguese [Native] – **English** [Advanced] – **French** [Advanced] – **Spanish** [Basic] – **German** [Basic – Learning]