

EU-DOCS for SmUCS

CO-SUPERVISED SUBJECT PROPOSAL FOR A DOCTORAL CONTRACT

Title of the thesis project Towards Fair and Explainable Lightweight Multimodal Learning Models for Effective Document Understanding	
La Rochelle University Research Unit L3i, Computer Science Department	Partner university South East Technological University Cotutelle research unit: CompuCore-Lab, Department of Computing
LRUniv supervisors Antoine Doucet, Cyrille Suire	Co-supervisor Rejwanul Haque
Non-academic partner SindiceTech (Siren) Company, Bordeaux, France	
Main research field: Document Understanding Secondary research field(s): Explainable AI, Fairness and Bias in Multimodal Learning	
Keywords: antimicrobial, biomaterials, microbial resistance, artificial intelligence, deep learning, graph neural networks	
Scientific description of the research project The proposed research project aims to develop lightweight, generalizable, and multimodal learning models for document analysis. The integration of deep learning (DL) has greatly advanced the field, allowing for the analysis of complex documents by incorporating vision, text, and layout information. Multimodal learning has become an essential strategy for understanding various document types such as legal, medical, administrative, and historical archives. Despite these advancements, current models face limitations in terms of size, computational efficiency, generalizability, and adaptability to different domains. Additionally, addressing social biases, ensuring fairness, and providing explainability in these models remain significantly challenging. The main objective of this research is to create multimodal, multitask learning models that are lightweight and can effectively process multimodal data while preserving fairness and transparency. The focus will be on developing innovative compression and quantization techniques to reduce model size, ensuring that the models can be deployed in environments with limited resources. The project will explore knowledge distillation methods to transfer knowledge from large, complex teacher models to smaller, efficient student models. This research introduces an innovative paradigm by enabling resource-efficient AI systems to handle complex medical, administrative, and legal data, where multimodal processing, fairness, and interpretability are critical for accuracy, ethical compliance, and real-world applicability. Scientific challenges in this project include enhancing the generalizability of models so that they perform well across various document types without overfitting or requiring excessive retraining. For instance, medical documents often include dense, domain-specific terminologies, structured tables, and diagnostic imagery, while legal documents are characterized by long, text-heavy clauses with complex semantic structures. Adapting a single model to excel across these diverse formats requires innovative approaches to avoid performance degradation in one domain while optimizing for another. The project will tackle modality biases and improve cross-modal interactions between vision, text, and layout information, which are critical for accurate document analysis. Another key challenge is adapting models to diverse document domains, such as legal or medical documents, with minimal fine-tuning. Compression and quantization will be explored to develop lightweight models suitable for fast and adaptive inference, ensuring computational efficiency without sacrificing accuracy. This is particularly	

critical in real-time medical triage systems or portable legal aid tools, where rapid responses are essential, and computational resources may be constrained. Finally, ensuring fairness by addressing social biases and enhancing explainability will be pivotal, allowing users to trust the model's decisions and insights. For example, when analyzing loan applications, ensuring that the model does not unfairly disadvantage applicants based on gender or ethnicity is essential. Similarly, providing clear rationales for extracted insights from medical records or legal agreements can foster trust and compliance in highly sensitive and regulated environments.

The state-of-the-art methods chosen for this research include a comprehensive combination of model compression and quantization strategies such as pruning, weight-sharing, and mixed-precision representations. These techniques aim to maintain accuracy while significantly reducing model size. Knowledge distillation will be utilized to train smaller student models that replicate the capabilities of larger teacher models. For the multimodal learning component, the project will design unified models that process visual, textual, and layout data cohesively, using shared parameter spaces and cross-modal attention mechanisms to facilitate seamless integration of information. Techniques like adversarial training and transfer learning will be employed for domain adaptation, ensuring the model's ability to adapt to new document types. Meta-learning approaches will be incorporated to enhance few-shot and zero-shot learning, boosting generalizability with minimal data. To ensure fairness and interpretability, the project will integrate metrics and loss functions that detect and mitigate social biases during training. Explainable AI tools, such as attention visualization and layer-wise relevance propagation (LRP), will be used to make the model's decision-making process transparent. Counterfactual fairness algorithms will also be explored to guarantee that the model provides unbiased results across different demographics.

The expected outcomes of this project include scalable and efficient models that achieve state-of-the-art results with significantly reduced size, enabling deployment in real-world, resource-constrained environments. The research aims to produce a multimodal learning model capable of generalizing across diverse document types with minimal retraining while ensuring fairness and transparency. The project will also contribute to the development of multitask learning frameworks that can handle multiple related tasks, such as machine translation, content summarization, docVQA, etc. within a single unified system. By leveraging knowledge distillation, smaller models will effectively inherit the capabilities of their larger counterparts, providing practical solutions without sacrificing performance.

This research has the potential to transform the field of document analysis by creating models that are lightweight, fair, and adaptable to various domains. Such advancements will benefit industries dealing with vast quantities of documents, including legal, healthcare, and administrative sectors, by offering AI solutions that are both cost-effective and trustworthy. Furthermore, by emphasizing fairness and transparency, the project will set new benchmarks for ethical AI practices in document understanding, promoting broader adoption and trust in AI technologies.

PhD student profile and skills required

The ideal candidate for this PhD project will have a strong academic background and demonstrated motivation for research at the intersection of artificial intelligence, machine learning, and document analysis. The student should be able to combine technical expertise with an interest in addressing ethical, societal, and practical challenges in AI.

Educational Background

- Master's degree (or equivalent) in Computer Science, Artificial Intelligence, Machine Learning, Data Science, or a closely related field.
- Solid understanding of deep learning techniques and neural architectures, with prior coursework or research experience in natural language processing (NLP), computer vision, or multimodal learning.

Technical Skills

- Strong programming skills in Python, with experience using ML/DL frameworks such as PyTorch or TensorFlow.
- Knowledge of multimodal architectures (vision, text, layout) and related methods such as attention mechanisms and transformers.
- Experience with data preprocessing, annotation, and large-scale dataset handling.
- Ability to use high-performance computing resources, and an interest in deploying models on resource-constrained devices.

Research and Analytical Skills

- Strong problem-solving abilities and a capacity to work on open-ended research problems.
- Familiarity with scientific writing, experimentation protocols, and reproducible research practices.
- Interest in explainable AI, bias detection/mitigation, and fairness evaluation in machine learning.

Soft Skills

- Motivation to work independently while collaborating effectively within a multidisciplinary research team.
- Strong communication skills in English, both written and oral, for research dissemination (publications, conferences, seminars).
- Openness to interdisciplinary collaboration, particularly with domains such as law, healthcare, and administrative sciences, where document analysis is crucial.
- Curiosity, adaptability, and commitment to advancing both the scientific and societal impact of AI.

Desired but not Mandatory

- Experience with transfer learning or meta-learning approaches.

Interest in sustainable computing, energy-efficient AI, or applications of AI for societal good.

This PhD offers the candidate an opportunity to contribute to cutting-edge AI research while addressing real-world societal and environmental challenges. The student will develop strong expertise in multimodal learning, model efficiency, and ethical AI, positioning them as a future leader in both academic and applied AI research.

Scientific alignment with EU-DOCs for SmUCS objectives

The proposed thesis aligns with the scientific orientations of the EU-DOCs for SmUCS objectives and La Rochelle Université's focus on "Smart Urban Coastal Sustainability" (LUDI) by leveraging advanced AI technologies to address societal and environmental challenges. Although the project primarily focuses on the analysis of medical, legal, and administrative documents, it significantly contributes to sustainable digital transformation and data-driven decision-making.

By developing lightweight, explainable, and fair multimodal learning models, the thesis enhances the accessibility of AI solutions for institutions with limited computational resources. These models are optimized for deployment on small devices, contributing to the decarbonization of digital infrastructures by reducing energy consumption and promoting the use of efficient AI frameworks. This focus on computational efficiency aligns with the EU's green transition goals, ensuring that AI technologies are not only powerful but also environmentally responsible.

The thesis aims to develop robust models capable of processing and analyzing diverse and complex documents, such as medical reports, legal texts, and administrative records. These capabilities are crucial for improving the efficiency of public and private sector institutions. By providing AI-driven tools to better understand regulatory frameworks, enhance compliance monitoring, and optimize administrative processes, the project directly supports the development of resilient urban coastal areas.

Furthermore, the emphasis on explainability and fairness aligns with the EU-DOCs' commitment to ethical AI. Ensuring transparency and unbiased decision-making fosters public trust, especially in sensitive sectors like healthcare and legal affairs. By integrating cross-modal information from text, vision, and document layouts, the project not only enhances document understanding but also supports interdisciplinary approaches. This bridges the gap between technology, policy, and societal needs, aligning with LUDI's mission to bring science and society closer together.

Societal and economic challenges and contributions

The societal and economic challenges addressed by this project are multifaceted, focusing on the integration of advanced document analysis and AI technologies into the fabric of modern society and the economy. A major challenge lies in mitigating social bias and ensuring fairness in AI models, which is essential for preventing discrimination and promoting equitable decision-making across various sectors. Transparency and trust are also critical issues, as current AI models often lack sufficient explainability, making it difficult for users to fully understand and trust the decisions these systems generate. Furthermore, ensuring accessibility and inclusivity is key, as it enables organizations of all sizes, including those with limited resources, to adopt and benefit from these technologies. This project also aims to address the challenge of adapting AI models to handle multilingual and complex document types, such as historical archives and domain-specific records, thereby broadening their applicability and impact.

From a societal standpoint, this project addresses critical policy and industry needs, including compliance with regulations such as the EU Artificial Intelligence Act and GDPR, which emphasize fairness, transparency, and accountability in AI systems, especially for sensitive applications in healthcare, legal, and administrative domains. The project aligns with industry demands for efficient, explainable AI solutions that can process and analyze complex documents while minimizing risks of bias and ensuring equitable outcomes. It also addresses the growing need for tools that support multilingual and culturally diverse document analysis to meet the globalized nature of modern enterprises. Furthermore, by developing lightweight, resource-efficient models, the project caters to the demand for sustainable and cost-effective AI systems, supporting organizations in adopting advanced technologies without requiring substantial computational infrastructure.

This project aligns closely with the objectives and capabilities of our non-academic partner as well (i.e., Siren) by addressing key challenges in data analysis and investigation through advanced AI-driven document understanding. The development of lightweight, multitask, and multimodal learning models in this project complements Siren's focus on integrating diverse data sources—open, vendor-supplied, and classified—by enabling efficient processing and analysis of complex, multimodal data such as text, images, and structured layouts at enterprise scale. The emphasis on fairness and explainability ensures that insights derived from sensitive investigations, such as those in national security, fraud detection, and compliance, are transparent and trustworthy, which is essential for decision-making in high-stakes environments. Moreover, the focus on resource efficiency aligns with Siren's need to handle vast data volumes without compromising performance, scalability, or accessibility for organizations with varying levels of computational resources. These synergies could significantly deliver cutting-edge, AI-driven solutions to safeguard people, assets, and networks.

Partnership context

Research: La Rochelle University (La Rochelle, France), SETU (Carlow, Ireland)

Socio-economic: SindiceTech (Siren) Company (Bordeaux, France)

The project takes place in the context of the European University alliance EU-CONEXUS, as a first concrete step to bridge only two computer science laboratories in the alliance that work on topics related to natural languages processing and multimodal analysis. This can help foster future Horizon Europe proposals, as both groups are very active in collaborative projects, in a field that is very active and with numerous industrial and societal applications with the emergence of generative AI and large language models. The partnership with SindiceTech (Siren) will be especially useful to this end, as the company is specialized in data investigation, focused on research and development on methods using AI, automation and advanced search, Siren links data from open source, vendors and classified sources allowing investigators to surface and analyze risks and threats. This partnership will be very beneficial for the student, as (s)he will be able to spend a substantial amount of time (3 months) in the company on a concrete subject dealing related to the PhD work, with a renowned actor in the field.



HR EXCELLENCE IN RESEARCH