

CURRICULUM VITAE DETAILLE
&
DOSSIER ANALYTIQUE

La Rochelle Université
Section CNU : 27 – Informatique

Souhail BAKKALI
27-Informatique
61-Génie informatique, automatique et traitement du signal

Enseignant-Chercheur Contractuel
Laboratoire L3i, La Rochelle Université
Avenue Michel Crépeau
17042 La Rochelle cedex
Souhail.bakkali@univ-lr.fr
+33 (0) 7 53 42 00 36

I. Curriculum Vitae Détaillé

1. Informations Générales

Situation Personnelle	27 ans, Marié, de nationalité Marocaine
Courriel	souhail.bakkali@univ-lr.fr
Contact	+33 (0) 7 53 42 00 36
Adresse Personnelle	34, Rue du Général Guillaumat, Résidence Beaugé, PTE 2 RDC, 17000, La Rochelle, France
Situation Actuelle	Chercheur Post-Doctoral au Département Informatique, Laboratoire L3i, La Rochelle Université (LR-UNIV)
Qualification	Section 27- Informatique Section 61- Génie Informatique, automatique, et traitement du signal

2. Objectif

Candidature pour l'obtention d'un poste de Maître de Conférences au département informatique de l'IUT de La Rochelle Université

3. Parcours Universitaire

2019-2022 DOCTORAT EUROPEEN

Doctorat obtenu en Informatique et Applications, Département Informatique, Laboratoire L3i – La Rochelle Université, France

Label Européen obtenu au Département Computer Science, Computer Vision Center (CVC), Universitat Autònoma de Barcelona.

Intitulé : Apprentissage Inter-Modal et Unifié de la Vision et de la Langue pour la Compréhension Multi-modale des Documents.

<i>Jury</i> : Président	:	Elisa Barney Smith – Boise State University
Rapporteurs	:	Simone Marinai – Università degli Studi di Firenze Ergina Kavallieratou – University of The Aegean
Examineurs	:	C. V. Jwahaar – IIIT Hyderabad Andreas Fischer – University of Applied Sciences and Arts Western Switzerland
Directeur de Thèse	:	Mickael Coustaty – La Rochelle Université
Encadrants	:	Oriol Ramos Terrades – Universitat Autònoma de Barcelona Zuheng Ming – Université Sorbonne Paris Nord Marçal Rusiñol – Universitat Autònoma de Barcelona

2016-2019

DIPLOME D'INGENIEUR

Ingénierie des télécommunications et technologies de l'information

– Institut National des Postes et Télécommunications (INPT), Rabat, Maroc

Option: Management des Systèmes d'Information

Mémoire: "Face Authentication: Deep Learning to merge biometric anti-fraud methods with the face verification algorithm by automatic recognition of L3I's facial expression methods"

2013-2016

CPGE

Classes Préparatoires aux Grandes Ecoles d'Ingénieurs – Al-Qalam, Agadir, Maroc

Option : Mathématiques – Physiques

2013

Baccalauréat

Lycée Al-Qalam, Agadir, Maroc

Option : Sciences Mathématiques

4. Expériences Professionnelles

2023-Présent

POST-DOCTORAT / VACATAIRE

Département Informatique, Laboratoire L3i, La Rochelle Université

Encadrants : Mickael Coustaty (Maître de Conférences, La Rochelle Université)

Apprentissage inter-modal, unifié et profond pour l'analyse et la compréhension multi-modale (vision et langue) des documents (classification des images de documents, récupération des documents basée sur le contenu, classification des documents à quelques exemples) ; vision par ordinateur ; traitement du langage naturel ; approches de fusion multi-modales.

2019-2022

DOCTORAT / VACATAIRE

Doctorat en Informatique et Applications au Département Informatique, Laboratoire L3i, La Rochelle Université, avec un Label Européen obtenu au laboratoire Computer Vision Center (CVC) de Universitat Autònoma de Barcelona.

Intitulé : "Apprentissage Inter-Modal et Unifié de la Vision et de la Langue pour la Compréhension Multi-modale des Documents".

Encadrants : Mickael Coustaty (Maître de Conférences – La Rochelle Université)

Zuheng Ming (Maître de Conférences – Université Sorbonne Paris Nord)

Oriol Ramos Terrades (Maître de Conférences, HDR – Universitat Autònoma de Barcelona)

Marçal Rusiñol (Maître de Conférences – Universitat Autònoma de Barcelona)

2018-2019

INGÉNIEUR EN MANAGEMENT DES SYSTEMES D'INFORMATIONS

Mémoire de Recherche – Département Informatique, Laboratoire L3i, La Rochelle Université

Intitulé : “Face Authentication: Deep Learning to merge biometric anti-fraud methods with the face verification algorithm by automatic recognition of L3I’s facial expression methods”

Encadrants : Jean-Christophe Burie (Professeur – La Rochelle Université)
Zuheng Ming (Maître de Conférences – Université Sorbonne Paris Nord)
Muhammad Muzzamil Luqman (Ph.D. – La Rochelle Université)
Zineb El Akkaoui (Ph.D. – INPT)
Mostafa Bellafkih (Ph.D. – INPT)

2017-2018 INGÉNIEUR EN SCIENCES DES DONNEES

Stage Technique – Inlog Solutions, Rabat, Maroc

Intitulé : “Fingerprint verification and recognition with deep learning”
“Building Data Mining models & making descriptive data analysis for knowledge extraction & highlight concrete indicators: Client
Segmentation - Recommendation of new possible offers.”

2016-2017 DEVELOPPEUR EN INFORMATIQUE

Stage d’Initiation à Lotus IT, Tanger, Maroc

Intitulé : “Développement Web & ERP Odoo”

5. Activités de Recherche

- Apprentissage inter-modal, unifié et profond pour l’analyse et la compréhension multi-modale (modalités de vision et de la langue) des documents (classification des images de documents, récupération des documents basée sur le contenu, classification des documents à quelques exemples) ; vision par ordinateur ; traitement du langage naturel ; approches de fusion multi-modales ; apprentissage profond basé sur les modèles d’attention (supervisé / auto-supervisé, non-supervisé), apprentissage multitâche, apprentissage métrique.
- **5 Publications :**
 - **2** articles de Revues Internationales, dont **1 PR – SJR rang Q1, IF 8.518** et **1 IJDAR – SJR rang Q1, IF 3.870**.
 - **2** articles de Conférences Internationales, dont **1 ICIP – Core rang B**, et **1 Core rang A*** soumis à **ICCV-2023** en cours de relecture.
 - **2** articles de Colloques Internationaux de qualité dont **1 CVPRW** et **1 ICDARW**.

6. Activités d'Enseignement

- Aide à la Préparation des sujets des CMs, TDs, et TPs.
- Plus de 216 heures équivalents TD d'enseignement en informatique allant du S1 au S9 à la FLASH, FDSPM, ILUDI, FST et COLLEGIUM (enseignement inter-disciplinaire).

7. Compétences Informatiques

- Développement : Python, RStudio, Shell, PHP/HTML/CSS, Tableau, SPSS Modeler, Enterprise Architect
- Base de Données : PostgreSQL, MySQL
- Framework : Tensorflow/ Keras, Pytorch, OpenCV, Scikit-learn, NLTK
- Environnement : Git, Visual Studio Code, Pycharm, Jupyter Notebook
- OS : Linux, Windows
- Bureautique : LaTeX, Microsoft Office, LibreOffice

8. Langues

- Français : parlé, lu, écrit
- Anglais : parlé, lu, écrit
- Arabe : langue maternelle

II. Résumé de mes activités pédagogiques et projet d'enseignement

1. Activités pédagogiques

Mon statut de doctorant ingénieur d'études embauché sur un projet ANR, m'a permis de transmettre des connaissances à des étudiants à travers l'enseignement en tant que vacataire. J'ai effectué ainsi plus de 216 heures équivalents TD depuis 2020.

Les enseignements ont été réalisés au sein de différentes structures tout au long de ce cursus. J'ai ainsi découvert et participé à la vie des départements : Informatique du pôle Sciences et Technologies de La Rochelle Université ; Département de Droit, Science politique et Gestion de La Rochelle Université, et aussi à la Faculté des Lettres, Langues, Arts et Sciences Humaines. Enfin, j'ai effectué des vacances au sein de l'école Niort-Tech de Niort. Mon implication dans ces diverses unités m'a permis de découvrir des filières variées et de mêler des enseignements exigeants à la fois la vulgarisation de l'informatique et aussi la transmission des techniques spécifiques en lien avec mon domaine d'études et de recherche, permettant de découvrir des profils d'étudiants différents.

Voici ci-dessous, un récapitulatif synthétique des enseignements en question :

Période	Signalétique E.C	Intitulé de l'enseignement	Niveau d'enseignement	Type d'enseignement	Durée
2022/2023	Faculté de Droit, de Science Politique et de Management	Technologies de lutte contre la cybercriminalité	Master 2 - tiers de confiance droit du numérique	TD	4.5
	Institut Littoral Urbain Durable Intelligent	Mise en œuvre des systèmes big data – Niort	Master 2 - Informatique (parcours Architecte logiciel)	CM TP	1.5 3
		Mise en œuvre des systèmes big data – La Rochelle	Master 2 - Informatique (parcours Architecte logiciel)	CM TP	1.5 7.5

		Réseaux et stockage sécurisés – Niort	Master 1 - Informatique (parcours Architecte logiciel)	CM TD TP	4.50 3 9
		Réseaux et stockage sécurisés – La Rochelle	Master 1 - Informatique (parcours Architecte logiciel)	CM TD TP	4.50 1.5 9
		Projet professionnel et informatique	Master 1 LCAI & MCI	TP	52.5
	Pôle Licences Collegium	Fonctions avancées des SGBD	Licence 3 Informatique	TP	12
2021/2022	Faculté de Droit, de Science Politique et de Management	Technologies de lutte contre la cybercriminalité	Master 2 - tiers de confiance droit du numérique	TD	4.5
	Institut Littoral Urbain Durable Intelligent	Fouille de Données – La Rochelle	Master 2 – Informatique	CM TP TEA	1.5 3 6
		Mise en œuvre des systèmes big data – Niort	Master 2 – Informatique	CM TP TEA	1.5 3 1.5
		C2i Niveau 2	Master 1 LCAI	TP	24
	Pôle Licences Collegium	Informatique d'usage	Licence 1	TP	82.5
		Génie Logiciel 2	Licence 3	TP	3
2020/2021	Faculté des Lettres, Langues, Arts et Sciences Humaines	Informatique d'usage	Licence 1	TP	45
	Pôle sciences et technologies	Informatique d'usage	Licence 1	TP	15
				CM	15
				TD	13.5
				TP	268.5
				TEA	7.5
				Total	304.5

2. Détails des enseignements

La présentation qui suit détaille les enseignements que j'ai assurés tout au long des 3 dernières années de thèse (période du 2020 – 2022), et aussi en tant qu'ingénieur de recherche (depuis début 2023). J'ai eu l'occasion d'encadrer des CMs, TDs, et TPs fournis par les enseignants responsables d'UE, mais j'ai également eu l'opportunité de prendre la responsabilité de quelques enseignements. Lorsque j'ai apporté une contribution plus importante, celle-ci est explicitement spécifiée.

L'ensemble des enseignements que j'ai pu mener au cours de ces trois dernières années m'a permis d'enseigner au sein d'environnements variés de La Rochelle (pôle Sciences et Technologies de l'université, pôle Droit, Science Politique et Gestion, département Informatique de l'université, faculté des Lettres, Langues, Arts et Sciences Humaines), dans divers domaines (informatique, sciences juridiques, politiques et gestion, management et commerce international, langues étrangères appliquées, sciences humaines), et enfin auprès du publics de niveaux différents. L'ensemble du contenu des enseignements réalisés est récapitulé ci-dessous.

Fouille de Données - Apprentissage profond - et Mise en œuvre des systèmes Big Data

Cet enseignement, destiné aux étudiants du Master 2 Informatique, option Architecte Logiciel de La Rochelle Université et de Niort-Tech, permet de découvrir différents algorithmes d'apprentissage automatique et profond dans un cadre Big Data afin de permettre aux étudiants de mieux comprendre la chaîne complète de traitement des données massives « big data » dans des applications pratiques aux niveaux industriel et recherche. Trois applications concrètes ont été adoptées dans cet enseignement : traitement d'images de scènes naturelles, traitement automatique du langage, et le traitement automatique multimodal (vision-langage) pour illustrer les techniques de base et montrer aux étudiants les tendances actuelles pour analyser des données dans un système big data. Mon implication dans cet enseignement est attachée au troisième cas d'étude : traitement automatique multimodal (vision-langage) où j'ai participé à la création et la rédaction de 2 supports de cours, et 2 sujets de TP afin d'assurer (3h CM, 6h TP, et 7.5h TEA) en 2021/2022, et (3h CM, 10.5h TP) en 2022/2023 à La Rochelle Université et aussi à Niort-Tech. Cet enseignement est directement lié à mon activité professionnelle (doctorant au L3i), ce qui m'a permis de transmettre des connaissances tout en m'appuyant sur mes activités de recherche.

Réseaux et Stockage Sécurisés

Cet enseignement, dispensé aux étudiants du Master 1 Informatique parcours Architecte Logiciel de la Rochelle Université et de Niort-Tech, permet aux étudiants de découvrir comment les informations sont transmises sur un réseau. Au travers de cet enseignement, ils découvrent les notions d'architectures réseaux : modèles OSI et TCP/IP, les couches réseaux et leurs différents protocoles. En TP, les étudiants configurent différentes structures réseau afin d'assurer la communication entre différentes machines (VMs dédiées). Ensuite, dans une deuxième partie du cours, ils découvrent les notions de la sécurité par le chiffrement et en réseau permettant de sécuriser les flux d'informations circulant dans un réseau sous ses différents protocoles. Ces enseignements couvrent les aspects suivants :

- Historique des réseaux et des moyens de communication
- Couche Accès au Réseau : contrôle d'erreurs, contrôle de flux (asservissement du débit de l'émetteur vers le récepteur), multiplexage, concentration et contrôle d'accès aux supports (câbles électriques, ondes lumineuses, ondes électromagnétiques).
- Couche Internet : IPv4, IPv6, datagramme, routage statique et dynamique, adressages, ponts et passerelles
- Couche Transport : TCP (mode connecté), UDP (mode non connecté)
- Couche Application : technologies d'internet, services et protocoles HTTP, DNS, FTP, SSH, etc.
- Cryptologie symétrique, asymétrique et hybride, et protocoles cryptographiques (SSL/TLS, Ipsec, S/MIME, PGP, SSH)

Dans cet enseignement, j'ai pu assurer un total de 9h de CMs, 4.5h de TDs, et 18h de TPs à La Rochelle et Niort.

Fonctions Avancées des SGBD

Cet enseignement, réservé aux étudiants de Licence 3 Informatique de La Rochelle Université porte sur l'utilisation des systèmes de gestion de bases de données PostgreSQL et son interface graphique PgAdmin. Principalement, ce cours vise à faire découvrir les bases de la programmation, des types et structures des données, et à mettre en application ces concepts via 12h de TPs que j'ai assurés, tout en incluant :

- Fonctions, règles et triggers
- Stockage et indexation
- Transactions et sécurité

- Performances et optimisation

Technologies de lutte contre la cybercriminalité

Cet enseignement, réservé aux étudiants de Licence 3 Informatique de La Rochelle Université permet aux étudiants juristes en droit, sciences politiques et gestion, de découvrir différentes techniques de lutte contre la cybercriminalité (fraude des documents hybrides). Bien que ces étudiants soient du domaine « non-informatique », il était très intéressant de leur faire découvrir des notions de base sur les algorithmes de hachage utilisés pour sécuriser les informations, pour ensuite vérifier leur intégrité par le destinataire (si l'information envoyée par l'émetteur n'a pas été falsifiée durant la transmission). Comme l'une des activités du L3i porte sur la sécurité et la détection de la fraude dans des documents numériques et/ou scannés, qui est aussi une activité principale des juristes (assurer la traçabilité des documents papier/électroniques), j'ai mis en place un cas d'étude concret où on sécurise un document, le modifie, et ensuite on vérifie son intégrité en utilisant des lignes de commandes mis en place, afin de leur montrer qu'on peut utiliser des techniques d'informatique simples et efficaces pour les aider dans leur quotidien. Dans cet enseignement, j'ai participé à la mise à jour des supports du cours, et à la création des supports de TDs et applications. Enfin, les étudiants ont montré leur intérêt et étaient ravis de pouvoir manipuler des lignes de commandes pour la première fois.

Cet enseignement inclut :

- Interopérabilité des systèmes d'informations, fonctionnement des systèmes d'informations avec une étude de cas : ENT de La Rochelle Université.
- Traçabilité des documents hybrides, avec une étude de cas (documents administratifs).

C2i Niveau 2 & Informatique d'usage

La préparation au C2i comportait des enseignements en tableur avancé, en bases de données, en création des sites web ou en droit et déontologie. Principalement, ces cours sont destinés à des non-informaticiens, visant à dispenser les notions de base de l'informatique bureautique, de l'internet, et plus généralement des technologies de l'information et de la communication afin de préparer l'obtention du certificat informatique et internet (C2i) ou la certification PIX.

Projet professionnel et informatique

Cet enseignement, destiné aux étudiants master 1 de la Faculté des Lettres, Langues, Arts et Sciences Humaines, options management et commerce international, et langues étrangères appliquées comportait des enseignements en tableur avancé, en bases de données, en création des sites web ou en droit et déontologie. Principalement, ce cours est destiné à des non-informaticiens, visant à dispenser les notions de base de l'informatique bureautique, de l'internet, et plus généralement des technologies de l'information et de la communication afin d'accumuler le maximum d'outils informatiques pour leur servir à la réalisation de leur projets professionnels. Cet enseignement fait l'objet d'un projet final où les étudiants doivent utiliser le maximum d'outils informatiques (création des sites web, communication interne avec google collab, utiliser des tableurs pour préparer des fiches clients ainsi que les fiches budgétaires, créer des flyers, des posters, montages de vidéo pour la publicité, éditer les images avec GIMP, utiliser ChatGPT pour les annonces publicitaires, DALL-E pour générer des logos etc.) afin de réaliser le projet (création d'une marque/entreprise).

III. Résumé de mes activités de recherche et projet d'intégration au L3i

1. Activités de Recherche

i. Contexte et sujet d'étude

Mes travaux de recherche ont été réalisés au sein du laboratoire L3i (dans l'équipe IC), du département informatique de La Rochelle Université entre décembre 2019 et décembre 2022. Mes travaux s'intègrent principalement dans le domaine de l'analyse et la compréhension des documents administratifs. Ces documents peuvent être des documents papier et/ou numériques produites par les grandes institutions publiques ou privées, intégrant différents types de contenus très hétérogènes (structurées et non-structurées). En effet, ces contenus se présentent souvent sous diverses formes, sous forme de graphiques dans des rapports techniques, de diagrammes dans des articles scientifiques et de conceptions graphiques dans des bulletins. Afin de prendre des décisions sur des sujets d'intérêt tels que la science, l'industrie, la santé, etc., l'être humain peut traiter efficacement les informations visuelles et textuelles contenues dans ces documents. Toutefois, comprendre et analyser manuellement de grandes quantités de données à partir de documents prend généralement du temps et coûte cher.

En général, les données de document sont souvent présentées dans des mises en page (layout) complexes en raison des différentes manières d'organiser chaque document. Contrairement aux images générales de scènes naturelles, les documents sont très difficiles compte tenu de leurs propriétés structurelles visuelles et de leur contenu textuel hétérogène. Dans ces conditions, le développement d'outils informatiques capables de comprendre et d'extraire automatiquement des informations structurées précises à partir d'une grande variété de documents reste crucial, d'une manière qui conduit à effectuer d'importantes applications administratives et/ou commerciales.

Il existe aujourd'hui plusieurs applications utilisées pour comprendre automatiquement les données des documents administratifs et commerciaux telles que : la classification des documents, la récupération de documents basée sur le contenu et la classification de documents à quelques exemples. Par conséquent, la clé de la compréhension automatisée des documents réside dans l'intégration efficace des signaux provenant de plusieurs modalités de données. Étant donné que les documents sont nativement multimodaux, il est important de tirer parti des informations multi-modales du langage et de la vision. Contrairement à d'autres formats de données tels que les images ou leur texte brut extrait à partir d'une reconnaissance optique de caractères (OCR), les documents combinent des informations visuelles et linguistiques, complétées par la mise en page du document. En outre, d'un point de vue pratique, de nombreuses tâches liées à la compréhension des documents sont rares. Un modèle qui peut apprendre à partir de documents non étiquetés (c.-à-d. un pré-entraînement), effectuer un ajustement du modèle pour des applications de documents spécifiques est plus préféré que celui qui nécessite des données d'entraînement entièrement annotées (c.-à-d. entraînés dans un mode d'apprentissage entièrement supervisé).

Le propre de mes travaux de recherche actuelle s'inscrit dans le cadre de la compréhension et l'analyse des images de documents administratifs (courriels, factures, publicités, articles, rapports, etc.), qui a été largement adoptée dans diverses applications de traitement d'images de documents. Mes travaux de recherche se concentrent principalement sur les interactions inter-modales entre les informations visuelles et textuelles dans les images de documents, visant la conception d'un environnement d'apprentissage efficace. En effet, le processus de conception de tels systèmes implique l'étude des avantages des interactions inter-modales dans l'apprentissage multi-modal. De tels systèmes encouragent l'apprentissage inter-modal entre les caractéristiques visuelles et textuelles des modalités visuelles et langagières afin d'améliorer leur distribution dans l'espace de représentation commun. Encore, les modèles développés sont le résultat d'un processus itératif d'analyse et de synthèse entre les théories existantes et nos études réalisées. Le propre de ma recherche actuelle part alors du fait d'étudier l'apprentissage inter-modal pour la compréhension contextualisée sur les composantes du document à travers le langage et la vision. L'idée principale est de tirer parti des informations multi-modales des images de documents dans un espace sémantique commun. Le principe consiste à extraire automatiquement des informations du contenu présenté dans les systèmes d'information (scan des documents, informations structurées et non structurées). Ensuite, comprendre les interactions entre données visuelles et

textuelles, réorganiser l'espace de recherche, et enfin trouver un espace sémantique commun pour réaliser les applications requises.

Dans l'ensemble, mon projet de recherche se concentre sur l'avancement de la recherche sur l'apprentissage inter-modal et apporte des contributions sur quatre fronts : (i) proposer une approche inter-modale avec un réseau neuronal bicéphale profond capable d'apprendre simultanément le contenu textuel et l'information visuelle des images de documents numérisés. En effet, l'objectif est d'exploiter conjointement les informations du langage et de la vision dans un espace de représentation sémantique commun pour effectuer et faire automatiquement des prédictions sur les documents multimodaux (c'est-à-dire le sujet dont ils traitent); (ii) étudier des stratégies concurrentielles pour s'attaquer aux tâches de classification inter-modale des documents, de classification de documents à quelques exemples, et de récupération de documents basée sur le contenu; (iii) résoudre les problèmes liés aux données comme l'apprentissage lorsque les données ne sont pas annotées, en proposant un réseau qui apprend des représentations génériques à partir d'une collection de documents non étiquetés ; enfin (iv) à exploiter les paramètres d'apprentissage lorsque les données ne contiennent que quelques exemples.

2. Résumé des contributions

Les modèles développés dans ma thèse sont le résultat d'un processus itératif d'analyse et de synthèse entre les théories existantes et nos études réalisées. Plus spécifiquement, nous avons étudié l'apprentissage inter-modal pour la compréhension contextualisée sur les composantes des documents à travers le langage et la vision. Nous avons limité nos explorations à deux des modalités bien étudiées dans la littérature documentaire, qui sont la vision et le langage. D'une part, en tant qu'application de la compréhension de documents, nous avons abordé le problème de la classification des images de documents en proposant des modèles qui trouvent un espace de représentation sémantique commun pour les modalités de vision et de langage en utilisant des réseaux profonds bien connus comme principaux piliers. La classification des images de documents est ensuite effectuée dans une méthodologie de fusion de caractéristiques précoce. Ensuite, nous avons amélioré l'espace de représentation sémantique en alignant les prédictions et en permettant aux deux modalités de transférer des informations pertinentes et des connaissances positives d'une modalité à une autre de manière à fusionner les caractéristiques intermédiaires. La classification des images de documents est ensuite effectuée dans différents contextes expérimentaux. D'autre part, nous avons également pris en compte les nouvelles avancées de l'apprentissage automatique pour intégrer ses stratégies. Nous avons choisi de concevoir des modèles pré-entraînés indépendants des tâches et des domaines pour valider l'hypothèse selon laquelle une grande généralisation intra-ensemble de données conduit à une grande généralisation inter-ensembles de données. Cette tâche est effectuée sous un paradigme de pré-entraînement puis d'ajustement. De plus, nous avons abordé le problème du manque de disponibilité des données de documents annotés par l'homme et amélioré l'apprentissage de la représentation sémantique en encourageant l'interaction multi-modale au sein des modalités de langage et de vision de manière auto-supervisée. Nous avons réalisé différentes études d'ablation démontrant l'efficacité de notre approche sur la tâche bien établie de classification de documents. Ainsi, réduire l'écart avec les méthodes état de l'art qui s'appuient sur la vision, le langage et les informations de mise en page (i.e. layout). De plus, nous avons effectué de nouvelles expériences sur deux nouvelles tâches que nous avons introduites comme référence dans la littérature sur la compréhension des documents. Ces tâches sont la classification de documents à peu d'exemples et la récupération de documents basée sur le contenu.

Enfin, mes activités de recherche sont toujours liées au développement des méthodes fondées sur l'apprentissage automatique en évoluant vers l'apprentissage profond / l'apprentissage multi-modal profond et en les appliquant aux différentes applications de la compréhension multi-modale des documents. Au cours de mes recherches, j'ai toujours eu une grande autonomie et j'ai pu développer mes propres innovations en collaboration avec d'autres collègues. J'ai également eu l'opportunité de faire un séjour de recherche de 4 mois (Label Européen) au Computer Vision Center (CVC) de l'Université Autonome de Barcelone. J'ai pu proposer un séminaire et initier une collaboration avec un doctorant, qui a fait l'objet d'une soumission à une conférence internationale ICCV 2023 Core rang A* en cours de relecture, et.

Au cours de mes recherches, j'ai toujours eu une grande autonomie et j'ai pu développer mes propres innovations en collaboration avec d'autres collègues. J'ai été aussi très intéressé par les problématiques que rencontrent les chercheurs dans le milieu industriel toujours dans le contexte de la compréhension et l'analyse des documents. De ce fait, et en début 2023, j'ai initié des collaborations qui sont en cours d'élaboration avec des collègues du L3i, afin d'encourager la recherche en apprentissage profond, développer des méthodes et approches capables de résoudre conjointement les problématiques rencontrées dans les deux milieux recherche et industrie dans le cadre de la compréhension et analyse des documents; proposer mes innovations pour résoudre des problématiques liées à la détection de la fraude dans des documents (exemple : tickets de caisse), et à la reconnaissance d'entités nommées dans des factures (exemple : montant total à payer, nom client, etc.). Je suis également l'un des administrateurs de serveur de calcul GPUs du L3i, ce qui me permet d'assister les nouveaux arrivants et les gens pour une meilleure utilisation des ressources GPUs pour leurs recherches sur l'apprentissage profond.

3. Rayonnement scientifique

L'ensemble de ces travaux de recherche a été valorisé dans cinq publications internationales acceptés :

- **Articles dans des revues internationales avec comités de relectures (3) :**

[J1] **Bakkali, S.**, Ming, Z., Coustaty, M., Rusiñol, M., & Terrades, O. R. (2023). VLCDoC: Vision-Language contrastive pre-training model for cross-Modal document classification. *Pattern Recognition*, 139, 109419. (**SJR rang Q1, IF 8.518**)

[J2] **Bakkali, S.**, Ming, Z., Coustaty, M., & Rusiñol, M. (2021). EAML: ensemble self-attention-based mutual learning network for document image classification. *International Journal on Document Analysis and Recognition (IJ DAR)*, 24(3), 251-268. (**SJR rang Q1, IF 3.870**)

- **Articles dans des conférences internationales avec comités de relectures (1) :**

[C1] **Bakkali, S.**, Biswas, S., Ming, Z., Coustaty, M., Rusiñol, M., Terrades, O. R. & Llados Josep. TransferDoc: A Self-Supervised Transferable Document Representation Learning Model Unifying Vision and Language. Submit to ICCV 2023. (**Core rang A***)

[C2] **Bakkali, S.**, Ming, Z., Coustaty, M., & Rusiñol, M. (2020, October). Cross-modal deep networks for document image classification. In *2020 IEEE International Conference on Image Processing (ICIP)* (pp. 2556-2560). IEEE. (**Core rang B**)

- **Articles dans des workshops internationaux avec comités de relectures (2) :**

[W1] **Bakkali, S.**, Ming, Z., Coustaty, M., & Rusiñol, M. (2020). Visual and textual deep feature fusion for document image classification. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops* (pp. 562-563).

[W2] **Bakkali, S.**, Luqman, M. M., Ming, Z., & Burie, J. C. (2019, September). Face detection in camera captured images of identity documents under challenging conditions. In *2019 International Conference on Document Analysis and Recognition Workshops (ICDARW)* (Vol. 4, pp. 55-60). IEEE.

- **Thèse :**

Souhail Bakkali. Multimodal Document Understanding with Unified Vision and Language Cross-Modal Learning. Laboratoire L3i de La Rochelle Université, Thèse, 2022.

- **Rapport de recherche M2 :**

[T1] **Souhail Bakkali**, Zuheng Ming, Muhammad Muzzamil Luqman, Jean-Christophe Burie, Zineb El Akkaoui, Mostafa Bellafkih. Face Authentication: Deep Learning to merge biometric anti-fraud methods with the face verification algorithm by automatic recognition of L3I's facial expression methods. Laboratoire L3i de La Rochelle Université.

Bien que l'ensemble de ces travaux listés ait déjà été validé et publié, 1 article de conférence internationale est en cours d'évaluation dans cette thématique.

4. Collaborations, responsabilités collectives et administratives

- **Collaborations :**

- Laboratoire L2TI, Université Sorbonne Paris Nord (USPN), France

Je collabore étroitement avec Dr. Zuheng Ming chercheur en IA et Vision au sein de L2TI. Nous travaillons régulièrement sur la compréhension et analyse multi-modale des documents. Nous avons déjà publié 5 articles dans des revues, conférences et colloques internationaux, et 1 article de conférence internationale en cours de relecture.

- Computer Vision Center (CVC, Universitat Autònoma de Barcelona, Espagne)

Je collabore également avec Dr. Marçal Rusinol, Dr. Oriol Ramos Terrades, chercheurs en IA et Vision au sein de CVC. Nous travaillons aussi et régulièrement sur la compréhension et analyse multimodale des documents. Nous avons déjà publié 4 articles dans des revues, conférences et colloques internationaux.

Je collabore également avec des doctorants et ingénieurs de recherche du L3i, et un doctorant du CVC autour des thématiques liés à l'analyse des documents.

- **Organisation de colloques & autres :**

Membre du comité d'organisation de :

- DAS 2022. 15TH IAPR International Workshop on Document Analysis Systems

Membre du Bureau de l'ADOCS 2020-2021 (Association des Doctorants et jeunes chercheurs de l'université de La Rochelle).

- **Symposium & Colloques :**

- Symposium International Francophone sur l'Écrit et le Document (SIFED 2022) :
 - Présentation Orale : Multimodal Document Understanding for Cross-Modal Document Image Classification.
- Participation au colloque des doctorants des 2^{ème} année (2022).
- Symposium International Francophone sur l'Écrit et le Document (INNOV & SIFED 2021).
- Ecole d'été EU-CONEXUS : *Professional and Scientific Communication and Networking in Multidisciplinary Environment*. Université de Zadar, Croatia, 2021.
- Symposium International Francophone sur l'Écrit et le Document (SIFED 2020) :
 - Présentation Orale : Cross-Modal Document Image Classification.

- **Activités de relecture :**

Depuis 2021, je suis activement reviewer pour Neurocomputing et aussi reviewer pour ICDAR2023 :

- ICDAR 2023
- SN Computer Science (2023) Springer Nature
- SN Computer Science (2022) Springer Nature
- SN Computer Science (2021) Springer Nature

- **Séminaires et présentations :**

- Multimodal Document Understanding with Unified Vision and Language Cross-Modal Learning, L3i, La Rochelle Université, La Rochelle, France, 2022. Ph.D. Thèse.
- Improving Semantic Multimodal Document Representation Learning, L3i, La Rochelle Université, La Rochelle, France, 2022.
- Multimodal Document Understanding, Computer Vision Center (CVC), Barcelone, Espagne.
- ICDAR 2021 | 16th International Conference on Document Analysis and Recognition (ICDAR-IJDAR), Lausanne, Suisse. **(Présentation Orale)**
- ICIP 2020 – 27th IEEE International Conference on Image Processing, conférence virtuelle. **(Présentation Orale)**
- CVPR 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition conférence virtuelle. **(Présentation Orale)**