



SIFED 2022

Rennes – 14 octobre



Déroulement de la journée

Lieu : IRISA, Rennes ; espace conférences - INRIA

Horaires	Sessions	Session chair	Lieu
8h15-8h45	Café d'accueil		Hall de l'espace conférences
8h45-8h50	Ouverture du symposium SIFED		Amphi
8h50-10h10	Présentations Session 1	Aurélie Joseph	Amphi
10h10-10h30	Pause-café		Hall
10h30-11h30	Présentations Session 2	Frédéric Rayar	Amphi
11h30-11h45	Présentations des sponsors	Yann Soullard	Amphi + stands hall
11h45-12h15	Assemblée générale du GRCE		Amphi
12h15-13h15	Repas		Salles Petri-Turing - Markov
13h15-14h	Posters et démonstrations	Aurélie Lemaître Nathalie Girard	Salles Petri-Turing - Markov
14h-15h	Présentation invitée : Elisa Fromont	Nicholas Journet	Amphi
15h-15h20	Pause-café		Hall
15h20-16h20	Présentations Session 3	Mickaël Coustaty	Amphi
16h20-16h40	Clôture du symposium SIFED		Amphi

Présentation invitée : Elisa FROMONT

Titre : Explainable Time Series Classification

Résumé : Après avoir brièvement redonné les motivations inhérentes à l'explication des méthodes d'apprentissage automatique en général et d'apprentissage profond en particulier; ainsi qu'un état de l'art des méthodes actuelles, je proposerai quelques contributions que nous avons pu avoir sur le sujet, notamment pour la classification de séries temporelles univariées et multivariées.

Bio : Elisa est professeur des universités à l'université Rennes 1 depuis 2017 et membre junior de l'institut universitaire de France (2019-2024). Elle travaille à l'institut de recherche IRISA au sein de l'équipe INRIA LACODAM (« Large Scale Collaborative Data Mining »). Ses domaines de recherche portent sur l'**IA explicable** (interprétabilité des modèles d'apprentissage machine / exploration de données et décisions explicables) et l'**exploration de données pour l'analyse de séries temporelles** (incluant la fraude et la détection d'anomalies).

Session Orale 1

8h50 – 9h10	
Titre de la communication	Handwriting Recognition using Multi-Hypothesis based Semi-Supervised Learning for adaptation to specific documents
Nom(s) du/des auteur(s)	Alexandre Chapin, Yann Soullard, Aurélie Lemaître, Bertrand Coüasnon
Résumé de la contribution	Handwritten text recognition in particular documents is still an unsolved problem due to the difficulty to get annotated images. This is especially the case for ancient documents due to the ancient language, untidy handwriting which is hard to read or degraded documents. In addition, it is not possible to get large training sets of specific documents for which only a few hundreds of pages have to be transcribed. This requires to train deep models with a very limited number of examples of the specific document. To overcome these difficulties, semi-supervised learning offers an interesting framework to benefit from unlabeled examples and improve the robustness of models. In this paper, we propose a multi-hypothesis semi-supervised approach where multiple network architectures are jointly use to produce label hypotheses, also called pseudo-labels, on unlabeled data which can then be used in training. This approach can be seen as a democratic co-learning method where predictions of various network architectures are used to strengthen the models. We experiment our approach using two different network architectures on the ICFHR 2018 READ dataset.
Mot clé 1	handwritten text recognition
Mot clé 2	semi-supervised learning
Mot clé 3	specific documents
Nom	Soullard
Prénom	Yann
Affiliation	IRISA
Adresse e-mail	yann.soullard@univ-rennes2.fr

9h10 – 9h30	
Titre de la communication	Vision-Language Contrastive Pre-Training Model for Cross-Modal Document Classification
Nom(s) du/des auteur(s)	Souhail Bakkali, Zuheng Ming, Mickael Coustaty, Marçal Rusinol, Oriol Ramos Terrades
Résumé de la contribution	Multimodal learning from document data has achieved great success lately as it allows to pre-train semantically meaningful features as a prior into a learnable downstream task. In this paper, we approach the document classification problem by learning cross-modal representations through language and vision cues, considering intra- and inter-modality relationships. Instead of merging features from different modalities into a joint representation space, the proposed method exploits high-level interactions and learns relevant semantic information from effective attention flows within and across modalities. The proposed learning objective is devised between intra- and inter-modality alignment tasks, where the similarity distribution per task is computed by contracting positive sample pairs while simultaneously contrasting negative ones in the joint representation space. Extensive experiments on public document classification datasets demonstrate the effectiveness and the generality of our model on low-scale and large-scale datasets.
Mot clé 1	Multimodal Document Representation Learning
Mot clé 2	Document Classification
Mot clé 3	Contrastive Learning
Mot clé 4	Self-Attention
Nom	Bakkali
Prénom	Souhail
Affiliation	L3i, La Rochelle Université
Adresse e-mail	souhail.bakkali@univ-lr.fr

9h30 – 9h50	
Titre de la communication	EXO-POPP : Extraction d'entités nommées dans des actes de mariage manuscrits
Nom(s) du/des auteur(s)	Thomas CONSTUM, Pierrick TRANOUEZ, Thierry PAQUET
Résumé de la contribution	<p>Le projet EXO-POPP a pour but de développer une base de données de 300 000 actes de mariage de Paris et de sa banlieue entre 1880 et 1940. Ces certificats de mariage fournissent un grand nombre d'informations sur les mariés, leurs parents et les témoins de mariage. Construire une telle base de données à la main nécessiterait au moins 50 000 heures de travail. Mais grâce aux récents développements en matière d'apprentissage profond, il est désormais possible de construire d'énormes bases de données avec des systèmes de lecture automatisés comportant la reconnaissance de l'écriture manuscrite et la reconnaissance d'entités nommées. Afin d'éviter les étapes de segmentation habituellement nécessaires pour effectuer la reconnaissance d'écriture, nous utilisons l'architecture DAN. Cette architecture (Coquenot-2022) a été développée par le LITIS et effectue de la reconnaissance d'écriture à partir d'un document complet pouvant contenir une ou deux pages. Le DAN fonctionne sans aucun label de segmentation grâce à l'utilisation de documents synthétiques. Ces documents doivent reproduire la structure des documents réels tout en comportant une diversité suffisante pour éviter le surapprentissage, notamment en ce qui concerne le modèle de langage appris implicitement par le modèle. L'adaptation du DAN aux données EXO-POPP a nécessité la conception d'un nouveau type de données synthétiques reproduisant la structure des actes de mariage. La génération de ces données a été enrichie de manière itérative afin de reproduire les cas avérés complexes pour le modèle. Dans cette présentation, nous expliquerons succinctement le fonctionnement de l'architecture DAN, puis nous présenterons les résultats de reconnaissance obtenus sur les données d'EXO-POPP. Enfin, nous détaillerons les premières pistes de recherche concernant la modification du DAN en vue d'effectuer conjointement la reconnaissance d'écriture et d'entités nommées. 1: Coquenot, Denis & Chatelain, Clement & Paquet, Thierry. (2022). DAN: a Segmentation-free Document Attention Network for Handwritten Document Recognition. https://arxiv.org/abs/2203.12273</p>
Mot clé 1	reconnaissance d'entités nommées
Mot clé 2	reconnaissance d'écriture
Mot clé 3	modèle sans segmentation
Mot clé 4	architecture transformer
Nom	CONSTUM
Prénom	Thomas
Affiliation	LITIS (université de Rouen)
Adresse e-mail	thomas.constum1@univ-rouen.fr

9h50 – 10h10	
Titre de la communication	Localisation d'une fraude dans des documents
Nom(s) du/des auteur(s)	Aurélie Ling, Cuong van Kieu, Camille Kurtz, Nicole Vincent
Résumé de la contribution	Dans un cadre de détection de fraude, des documents passent d'un utilisateur à un autre, par exemple un échange de contrat pour signature. Suite à cette dernière, le partenaire a besoin de vérifier la présence d'une fraude éventuelle (par exemple, le montant du salaire, l'ajout d'une marque ou d'un surlignage), c'est-à-dire un changement entre le document (initial) référant et le document renvoyé. Nous proposons une méthode pour la comparaison de deux pages de documents en s'appuyant sur une étude locale statistique concernant la couleur, aussi bien que la forme, et produisant la localisation des différences. Les premières expériences menées sur des documents administratifs, comportant différentes classes de modification, ont montré l'intérêt de cette approche. Ce travail a été réalisé dans le cadre d'une collaboration entre le LIPADE (Université Paris Cité) et la société IMDS Software Inc (Montréal, Canada).
Mot clé 1	Fraude
Mot clé 2	Comparaison de documents
Mot clé 3	Tests statistiques
Mot clé 4	Alignement par le contenu
Nom	Vincent
Prénom	Nicole
Affiliation	LIPADE (Université Paris Cité)
Adresse e-mail	nicole.vincent@u-paris.fr

Session Orale 2

10h30 – 10h50	
Titre de la communication	Détection et association des queues de bulles : cas des multi-bulles
Nom(s) du/des auteur(s)	Christophe Rigaud(L3i), Jean-Christophe Burie (L3i), Samuel Petit (Geo Comix)
Résumé de la contribution	Cet exposé présente une nouvelle approche de détection de queue de bulle à partir du contour des bulles, notamment celles composées de plusieurs bulles interdépendantes, dans les images de bandes dessinées contemporaines. Un indicateur de confiance d'association automatique entre les queues des bulles et les personnages sera aussi exposé. Cette association d'éléments peut permettre d'associer textes et graphiques (e.g. visage) dans la recherche d'émotion, d'améliorer la pré-translation ou la synthèse vocale en caractérisant le texte en fonction du personnage qui parle.
Mot clé 1	analyse de document
Mot clé 2	détection de forme
Mot clé 3	évaluation de qualité
Nom	Rigaud
Prénom	Christophe
Affiliation	L3i/GeoComix - La Rochelle
Adresse e-mail	christophe.rigaud@univ-lr.fr

10h50 – 11h10	
Titre de la communication	Open Source MAIF pour une IA transparente & Maitrisée
Nom(s) du/des auteur(s)	Laure BONAFOUS – Sébastien BIDAULT
Résumé de la contribution	<p>Les grandes lignes de ce que l'on présente :</p> <ul style="list-style-type: none"> Pourquoi la MAIF se porte sur l'Open Source & LabelIA Le cycle de vie d'un modèle. <ul style="list-style-type: none"> L'entraînement & l'explicabilité avec Shapash L'industrialisation & le suivi de la dérive avec Eurybia Le tout illustré dans un cas usage (lien avec la démo de midi)
Mot clé 1	Open Source
Mot clé 2	Entraînement de modèle
Mot clé 3	Explicabilité
Nom	BONAFOUS
Prénom	Laure
Affiliation	Data Scientist, MAIF
Adresse e-mail	Laure.bonafous@maif.fr & Sebastien.bidault.marketing@maif.fr

11h10 – 11h30	
Titre de la communication	Arkindex pour l'expérimentation en traitement de documents numérisés
Nom(s) du/des auteur(s)	Mélo die Boillet, Solène Tarride et Christopher Kermorvant
Résumé de la contribution	Arkindex est une plateforme de traitement de documents numérisés développée par TEKLI A. L'objectif de cette plateforme est de fournir un environnement complet pour la réalisation de projets de reconnaissance de documents numérisés à large échelle. Elle permet la gestion , la consultation, l'annotation de documents, l'application de modèles de reconnaissance automatique (DLA, HTR/OCR, classification, NER) et leur évaluation. Bien qu'originale ment développée pour les besoins internes de TEKLI A, la plateforme a été conçue avec un objectif de modularité et d'ouverture: utilisation de IIF, accès complet par API, client python, client en ligne de commande, import/export. Elle n'a cependant pas été développée pour une utilisation grand public comme Transkribus, mais plutôt pour des professionnels du traitement de documents. Nous présenterons l'architecture et les grandes fonctionnalités d'Arkindex que nous illustrerons par deux cas pratiques d'utilisation sur des bases publiques avec des exemples de code en python. La plateforme est disponible sur demo.arkindex.org (edited)
Mot clé 1	plateforme de traitement de document
Mot clé 2	analyse de document
Mot clé 3	reconnaissance d'écriture
Nom	Kermorvant
Prénom	Christopher
Affiliation	TEKLI A
Adresse e-mail	kermorvant@tekli a.com

Session Posters et démonstrations

13h15 – 14h – Salles Petri-Turing – Markov

POSTER	
Titre de la communication	Leveraging Domain Knowledge for Document Information Extraction
Nom(s) du/des auteur(s)	Arthur HEMMER
Résumé de la contribution	Business documents contain a multitude of structured values such as dates, amounts, quantities, percentages, invoices numbers, etc. In the context of automated decision making, high accuracy and confidence estimation of the recognition of these values are crucial. While state of the art OCR technology has greatly improved on such documents, there still remain many conditions that can add noise to the text recognition process, especially when it comes to handwritten text. My thesis will focus on exploiting the intrinsic structure as well as the interdependent semantic structure of such values in order to improve both the accuracy and confidence estimation.
Mot clé 1	Business documents
Mot clé 2	Information extraction
Mot clé 3	Document understanding
Mot clé 4	Confidence estimation
Nom	Hemmer
Prénom	Arthur
Affiliation	L3I, La Rochelle & Shift Technology, Paris
Adresse e-mail	arthur.hemmer@shift-technology.com

POSTER	
Titre de la communication	Reconnaissance et segmentation de l'écriture manuscrite d'enfants
Nom(s) du/des auteur(s)	Simon Corbillé, Élisabeth Fromont, Éric Anquetil
Résumé de la contribution	<p>Cette contribution cherche à résoudre une problématique de reconnaissance et de segmentation d'écriture manuscrite d'enfant. Le challenge est double, car il faut concevoir un système à la fois performant en reconnaissance et en segmentation. Un cas applicatif de ce système est : à partir d'un mot écrit sur une tablette avec stylet par un enfant, reconnaître le mot écrit et segmenter le mot en lettre pour indiquer en temps réel à l'élève s'il a commis une erreur. Ces travaux s'intéressent à des mots d'enfant en cours d'apprentissage de l'écriture. Cela apporte une difficulté supplémentaire, car l'écriture est dégradée et comporte des fautes d'orthographe comparées à une écriture d'adulte. L'idée de cette contribution est d'utiliser une hybridation entre un modèle performant en reconnaissance (réseau Seq2Seq) et un modèle performant en segmentation (réseau Mask R CNN) pour répondre à la problématique. Les mots sont représentés par un signal en ligne converti en image comme entrée des modèles. Un autre axe d'amélioration vise à intégrer les informations contenues dans le signal en ligne dans le système hybride.</p>
Mot clé 1	reconnaissance
Mot clé 2	segmentation
Mot clé 3	écriture manuscrite d'enfant
Nom	Corbillé
Prénom	Simon
Affiliation	équipe IntuiDoc, laboratoire IRISA, Université Rennes 1
Adresse e-mail	simon.corbille@irisa.fr

POSTER	
Titre de la communication	A comparative study for classification of French language corporate documents on imbalanced datasets
Nom(s) du/des auteur(s)	Tri-Cong PHAM, Aurélie Joseph, Muriel Visani, Nicolas Sidere, Mickael Coustaty
Résumé de la contribution	Companies need to manage and process large amounts of documents every day. Depending on the age of the company, the number of documents can be very large both in terms of the number of types of documents and the number of documents of each type. The first step in this processing chain is document classification. If the classification is wrong, it will lead to heavy cost in the next step. Traditional classification tasks require a static and balanced distribution of data to ensure the learner to be trained to generalize over all classes. Obviously, in real-world business documents datasets, spontaneous new classes or intra-class variations may appear and the number of instances vary substantially among classes. Incremental learning is particularly suitable for this task where the dataset is constantly evolving with the addition of new instances and categories of documents. In this study, we evaluate the effectiveness of some NLP and Deep Learning based document classification algorithms on public and private imbalanced datasets. We evaluate the performance of not only the static learning scenario, but also incremental learning scenarios when the number of new document categories is added.
Mot clé 1	Document Classification
Mot clé 2	French Document
Mot clé 3	Imbalanced Dataset
Mot clé 4	Incremental Learning
Nom	PHAM
Prénom	Tri-Cong
Affiliation	La Rochelle Université, L3i Avenue Michel Crepeau, 17042 La Rochelle, France.
Adresse e-mail	cong.pham@univ-lr.fr

POSTER	
Titre de la communication	Transformers pour la reconnaissance de documents anciens
Nom(s) du/des auteur(s)	Killian Barrere, Yann Soullard, Aurélie Lemaitre, Bertrand Couïasnon
Résumé de la contribution	Le but de ces travaux est d'appliquer des réseaux de neurones basé sur des couches types transformer sur la reconnaissance de documents anciens. Ces couches peuvent potentiellement apporter un fort bénéfice pour des écrits complexes à reconnaître. Cependant, la quantité de données disponible pour ce type de documents est assez vite limitante, et requiert l'utilisation de données synthétiques pour palier à ce manque et permettra l'apprentissage de tels réseaux.
Mot clé 1	Transformer
Mot clé 2	Documents anciens
Mot clé 3	Réseaux de neurones profonds
Mot clé 4	Reconnaissance de texte
Nom	Barrere
Prénom	Killian
Affiliation	IRISA
Adresse e-mail	killian.barrere@irisa.fr

POSTER	
Titre de la communication	Classification of Untranscribed Handwritten Notarial Documents by Textual Contents
Nom(s) du/des auteur(s)	José Ramon Prieto Fontcuberta
Résumé de la contribution	
Nom	Prieto Fontcuberta
Prénom	José Ramon
Affiliation	PRHLT
Adresse e-mail	joprfon@prhlt.upv.es

POSTER	
Titre de la communication	Détection de fraude de tickets de caisse en prenant en compte la sémantique
Nom(s) du/des auteur(s)	Beatriz Martínez Tornés
Résumé de la contribution	<p>Nous nous intéressons à la tâche de détection de fraude à partir du contenu textuel de documents. Comme ce type de tâche est principalement abordé avec des techniques de vision par ordinateur, il repose sur la détection de modifications graphiques. Cependant, étant donné l'abondance et l'hétérogénéité des documents numérisés qui peuvent facilement être générés et modifiés, des méthodes basées sur le contenu et conscientes de la sémantique pourraient permettre la détection d'incohérences. Nous proposons d'aborder cette tâche avec des techniques récentes de traitement du langage naturel. Nous la traitons donc comme une approche basée sur la régression, en tirant parti d'un modèle de langage pré-entraîné afin de représenter le contenu textuel, et en enrichissant la représentation avec des entités et des relations basées sur une ontologie. Nous utilisons un jeu de données librement disponible de faux tickets de caisse et nous fournissons une analyse approfondie de nos résultats en ce qui concerne l'efficacité de nos méthodes et les biais existants dans les documents falsifiés.</p>
Mot clé 1	Fraude
Mot clé 2	Sémantique
Nom	Martínez Tornés
Prénom	Beatriz
Affiliation	L3i, La Rochelle Université
Adresse e-mail	beatriz.martinez_tornes@univ-lr.fr

POSTER	
Titre de la communication	Transformation digitale d'images de documents techniques d'ingénierie et de plans industriels.
Nom(s) du/des auteur(s)	Mathieu FRANCOIS
Résumé de la contribution	<p>La dématérialisation de la documentation papier est devenu un nouvel enjeu de l'entreprise actuelle : elle ne se limite pas à scanner des documents, ni à simplement stocker de gros volumes de documents numérisés, mais elle consiste à capturer, annoter, reconnaître et sécuriser des informations, des données, et des documents provenant de sources diverses, en circulation dans l'entreprise et se présentant sous de multiples formats. Dans ce contexte, l'objectif du travail de thèse est d'aboutir à une solution de reconnaissance de texte et de symboles sur les documents d'ingénierie visant la construction et le remplissage de bases de données. Il s'agit d'un travail de recherche centré sur la transformation numérique des cartes et plans contenant des informations textuelles et symboliques (textes, symboles et schémas) peu structurés. Ce travail concerne différentes étapes de la transformation numérique des contenus: l'automatisation de la rétro-ingénierie de pattern de données non-structurées, l'extraction des tags (entités nommées et abréviations textuelles) en environnement non contraint (non-structuré) et la classification de ces tags selon leur contenu et leur contexte ; la reconnaissance de symboles graphiques simples et complexes (apprentissage de représentation et classification en environnements semi-supervisés) ; la reconnaissance conjointe contextualisée et couplée des tags et des symboles sur les documents scannés et l'analyse de la structure des schémas incluant des mécanismes de désambiguïsation de la reconnaissance ; une représentation complète et enrichie des plans et de leurs données symboliques et textuelles (visualisation des liaisons entre symboles), visant une annotation sémantique des documents et leur compréhension ; la mise en place d'une plateforme de recherche par le contenu conçue pour interroger des bases de données de plans à partir de requêtes graphiques ou textuelles</p>
Mot clé 1	Text detection/recognition
Mot clé 2	Correction Post-OCR
Mot clé 3	Symbol recognition
Mot clé 4	Image/Text Association
Nom	François
Prénom	Mathieu
Affiliation	Liris
Adresse e-mail	francois.mathieu@orinox.com

POSTER	
Titre de la communication	Document Forgery Detection in the Context of Double JPEG Compression
Nom(s) du/des auteur(s)	Théo Taburet, Kais Rouis, Mickaël Coustaty, Petra Gomez-Krämer, Nicolas Sidère, Saddok Kébairi, Vincent Poulain d'Andecy
Résumé de la contribution	In this paper, we propose a strategy to train a CNN to detect document manipulations in JPEG documents under data scarcity scenario. As it comes to scanned PDF documents, it is common that the document consists of a JPEG image encapsulated into a PDF. Indeed, if the document before tampering was a JPEG image, its manipulation will lead to double compression artefacts within the resulting tampered JPEG image. In contrast to related methods that are based on handcrafted histograms of DCT coefficients, we propose a double compression detection method using a one-hot encoding of the DCT coefficients of JPEG images. We can use accordingly a CNN model to compute co-occurrence matrices and avoid handcrafted features such as histograms. Using simulated frauds on Perlin noise, we train our network and then test it on textual images against a state-of-the-art CNN algorithm trained on natural images. Our approach has shown an encouraging generalization on both the database used in the paper and on a stream of synthetic frauds on real documents used in the company Yooz.
Mot clé 1	Jpeg
Mot clé 2	Documents
Mot clé 3	Tampering
Nom	TABURET
Prénom	Théo
Affiliation	L3i
Adresse e-mail	theo.taburet@gmail.com

DEMO	
Titre de la communication	Geo Comix Translator: a unique tool to speed-up translation process for comics, mangas and webtoons.
Nom(s) du/des auteur(s)	Christophe Rigaud, Jean-Christophe Burie
Résumé de la contribution	L'entreprise Rochelaise GeoComix propose depuis moins d'un an un premier outil de traduction assistée par ordinateur aux éditeurs de BD. Il a été développé en étroite collaboration avec le laboratoire de recherche L3i depuis plus de 10 ans. Une présentation de l'outil et les étapes clés qui ont permis son industrialisation pourront être discutées pour faire le lien entre la recherche et l'industrie sur ce secteur.
Mot clé 1	analyse de document
Mot clé 2	traduction assistée par ordinateur
Nom	Rigaud
Prénom	Christophe
Affiliation	L3i/GeoComix - La Rochelle
Adresse e-mail	christophe.rigaud@univ-lr.fr

DEMO	
Titre de la communication	Arkindex pour l'expérimentation en traitement de documents numérisés
Nom(s) du/des auteur(s)	Mélodie Boillet, Solène Tarride et Christopher Kermorvant
Résumé de la contribution	Arkindex est une plateforme de traitement de documents numérisés développée par TEKLIÀ. L'objectif de cette plateforme est de fournir un environnement complet pour la réalisation de projets de reconnaissance de documents numérisés à large échelle. Elle permet la gestion, la consultation, l'annotation de documents, l'application de modèles de reconnaissance automatique (DLA, HTR/OCR, classification, NER) et leur évaluation. Bien qu'originellement développée pour les besoins internes de TEKLIÀ, la plateforme a été conçue avec un objectif de modularité et d'ouverture: utilisation de IIF, accès complet par API, client python, client en ligne de commande, import/export. Elle n'a cependant pas été développée pour une utilisation grand public comme Transkribus, mais plutôt pour des professionnels du traitement de documents. Nous présenterons l'architecture et les grandes fonctionnalités d'Arkindex que nous illustrerons par deux cas pratiques d'utilisation sur des bases publiques avec des exemples de code en python. La plateforme est disponible sur demo.arkindex.org (edited)
Mot clé 1	plateforme de traitement de document
Mot clé 2	analyse de document
Mot clé 3	reconnaissance d'écriture
Nom	Kermorvant
Prénom	Christopher
Affiliation	TEKLIÀ
Adresse e-mail	kermorvant@teklià.com

DEMO	
Titre de la communication	Open Source MAIF pour une IA transparente & Maitrisée
Nom(s) du/des auteur(s)	Laure BONAFOUS – Sébastien BIDAULT
Résumé de la contribution	<p>Les grandes lignes de ce que l'on présente :</p> <p>Pourquoi la MAIF se porte sur l'Open Source & LabelIA</p> <p>Le cycle de vie d'un modèle.</p> <p>L'entraînement & l'explicabilité avec Shapash</p> <p>L'industrialisation & le suivi de la dérive avec Eurybia</p> <p>Le tout illustré dans un cas usage</p>
Mot clé 1	Open Source
Mot clé 2	Entraînement de modèle
Mot clé 3	Explicabilité
Nom	BONAFOUS
Prénom	Laure
Affiliation	Data Scientist, MAIF
Adresse e-mail	Laure.bonafous@maif.fr & Sebastien.bidault.marketing@maif.fr

Session Orale 3

15h20 – 15h40	
Titre de la communication	Toward Deep neural network for pen trajectory reconstruction from kinematic sensors
Nom(s) du/des auteur(s)	Imbert Florent, Wassim Swaileh, Yann Soullard, Romain Tavenard, Eric Anquetil
Résumé de la contribution	Handwriting with connected pen become one of the major human-computer easy interaction methods. In comparison with traditional touch screen handwriting systems, the pen-based interaction method has the advantage of producing the online handwriting signal without surface constraints. Indeed, someone can write on a paper and obtain the corresponding pen trajectory coordinates that represent the online handwriting signal. In this work, we introduce a neural network based approach that reconstruct the handwriting traces of the digital stylus Digipen of STABILO that is equipped with a wireless trajectory tracking system based on kinematic sensors. We use deep neural network model in order to reconstruct the pen-down (touching) and pen-up (hovering) parts of the writing. The proposed model receives 13 input channels provided by IMUs signals of two accelerometers, a gyroscope, a compass and press sensor. The model performance is evaluated on dataset collected by different scriptures.
Mot clé 1	DeepLearning
Mot clé 2	Temporal Neural Networks
Mot clé 3	Trajectory reconstruction
Mot clé 4	Inertial Measurement Unit (IMU)
Nom	Imbert
Prénom	Florent
Affiliation	Laboratoire IRISA UMR 6074, CNRS - INSA de Rennes
Adresse e-mail	florent.imbert@irisa.fr

15h40 – 16h00	
Titre de la communication	EyCon: Computational Approaches to the Circulation of photographs in late 19th-/Early 20th-Century periodicals and beyond
Nom(s) du/des auteur(s)	Julien Schuh (Université Paris Nanterre, CSLF EA 1586I), Marina Giardinetti (Université Paris Cité, LARCA UMR 8225), Soumik Mallick (Université Paris Cité, LARCA UMR 8225) and Daniel Foliard (Université Paris Cité, LARCA UMR 8225)
Résumé de la contribution	In this paper, we examine machine-learning methods in the analysis of late 19c and early 20c photo-illustrated French and British periodicals. We demonstrate how the application of deep-learning tools to pre-classify page layouts not only paves the way for more reliable image extraction, but also for a better understanding of the history of photographic news. It allows for a detailed and diachronic analysis of the development and esthetics of photographic illustration in the news. We suggest that the evolution of layout designs and image placement in periodicals cannot be disconnected from the visual culture materialized by amateur albums. The article shows that machine learning algorithms can foster the analysis of visual historical big data and that they should be a full part of the hermeneutics of the history of modern media culture.
Mot clé 1	Layout-parser
Mot clé 2	Machine learning
Mot clé 3	Historical document image document analysis
Mot clé 4	Deep learning
Nom	Mallick
Prénom	Soumik
Affiliation	
Adresse e-mail	soumikbsnl@gmail.com

16h00 – 16h20	
Titre de la communication	Reconnaissance de texte dans les bandes dessinées par Smartphone
Nom(s) du/des auteur(s)	Clément Charrier, Frédéric Daubignard (Entreprise ALGONA), Rémy Leconge (Laboratoire PRISME), Frédéric Rayar (Laboratoire LIFAT), Sylvie Treuillet (Laboratoire PRISME)
Résumé de la contribution	<p>Ce projet a pour but d'améliorer l'accessibilité aux bandes dessinées (BD) aux personnes rencontrant des difficultés de lecture ou fournir un outil ludique d'apprentissage de la lecture aux personnes voulant s'initier au français, comme langue étrangère par exemple. L'idée est de développer une application sur terminal mobile offrant une expérience de lecture augmentée sous la forme d'une lecture vocale des bulles permettant de conserver le plaisir du contact avec le livre. L'application souhaitée doit être simple à utiliser, notre objectif est donc de faire en sorte que l'utilisateur n'ait qu'à prendre en photo une case/page, sélectionner une bulle et le texte devra être lu dans un délai relativement court. Notre projet se fonde sur une base logicielle, réalisée lors d'une étude précédente[1]. La détection de texte présents dans des bulles de bandes dessinées a déjà fait l'objet de différents travaux comme ceux de Christophe Rigaud [2][3 que nous avons essayé d'appliquer dans notre projet. Sa méthode de détection de bulle repose sur la binarisation de l'image puis la recherche des zones blanches fermées contenant de petites structures noires alignées et proches (texte). Cependant, les résultats ne furent pas très probants, nos images sont acquises par un simple smartphone sans autre équipement ni contrainte additionnelle, ce qui implique des problématiques supplémentaires (luminosité, torsion du support, cadrage non optimal,...) conduisant à une détection des bulles plus délicates. Cette problématique ne s'applique pas qu'à la détection des bulles, la détection de texte est aussi impactée, d'autant que certains textes sont difficiles à lire, l'écriture manuscrite, par exemple. Pour pallier à ce problème, nous avons suivi une autre approche. Partant du constat que la détection ou reconnaissance du texte est bien souvent partielle, nous avons décidé de créer un fichier de données composé de tous les textes contenus dans la bande dessinée ciblée. L'idée est de retrouver la bonne chaîne de caractères dans ce fichier à partir de la reconnaissance totale ou partielle du texte photographié. La reconnaissance du texte est réalisée sur la photo grâce à l'OCR Google ML –Kit. De plus, certains morceaux de textes reconnus par ML-Kit peuvent s'agréger ou fusionner selon la pagination, il est donc nécessaire de les réorganiser afin de lire de façon fidèle la bulle sélectionnée. Cette méthode est rapide et les premiers résultats sont bons. Testé sur un album d'Astérix comprenant 725 bulles sur 388 images, on atteint une reconnaissance de 97 % pour un temps de traitement moyen inférieur à 2 secondes, ces résultats devront être confirmés par d'autres tests sur d'autres bandes dessinées afin de finaliser l'application. [1 Étude comparative de reconnaissance de texte dans les bandes dessinées. (2022) Florian Le Meur, Frédéric Rayar, Sylvie Treuillet, Frédéric Daubignard. 22e Conférence francophone sur l'Extraction et la Gestion des Connaissances, Blois, France. fhal-03597262 [2 Text-independent speech balloon segmentation for comics and manga (2017).</p>

	Christophe Rigaud, Jean-Christophe Burie, Jean-Marc Ogier Laboratoire L3i, Université de La Rochelle International Workshop on Graphics Recognition [3 Text block segmentation in comic speech bubbles (2021). Christophe Rigaud, Nhu-Van Nguyen, Jean-Christophe Burie. ICPR International Workshops and Challenges, pp.250-261 ff10.1007/978-3-030-68780-9_22ff. fhal-03281488f
Mot clé 1	Bande dessinée
Mot clé 2	Android
Mot clé 3	TextToSpeech
Mot clé 4	MLKIT
Nom	CHARRIER
Prénom	Clément
Affiliation	Laboratoire PRISME
Adresse e-mail	clement.charrier@etu.univ-orleans.fr