INSTITUT NATIONAL DES SCIENCES APPLIQUÉES RENNES

Demande de Contrat Doctoral Ordinaire - INSA RENNES 2024

FICHE DESCRIPTIVE (3 pages maximum*) – annexes autorisées (pdf)

Catégorie demande	X contrat doctoral 100 %
Si demande effectuée sur l'appel Contrats Doctoraux Ordinaires (ED SPI.bzh, S3M, Matisse) indiquez-le	☐contrat doctoral 100 %
Si cofinancement nature du cofinancement**	**(partenaire/organisme, montant cofinancement, nature du contrat, durée, calendrier)**
Titre du sujet de thèse	Apprentissage auto-supervisé pour la reconnaissance d'écriture manuscrite
Composante/équipe	X IRISA
Direction de la thèse/Codirection (HDR)	(au moins un personnel INSA ou hébergé à l'INSA) – souligner le porteur scientifique du projet Bertrand Couasnon (Maitre de conférences HDR)
Co-encadrement	Yann Soullard (IRISA, Maitre de conférences univ Rennes 2), Denis Coquenet (IRISA, Maitre de conférences univ Rennes)
Partenaires impliqués	
Cotutelle / co-direction internationale envisagée* ?	**(si oui joindre lettre engagement partenaire en annexe de la demande)**
Description des objectifs scientifiques du projet de thèse	**(1 page maximum)** Cette thèse s'inscrit dans la thématique de la reconnaissance d'écriture manuscrite. On s'intéressera plus particulièrement à reconnaître du texte manuscrit présent dans des images de pages de documents. L'objectif est ainsi de faciliter l'accès à la transcription automatique de documents divers, potentiellement anciens. Les documents anciens présentent en effet des difficultés accrues liés à la langue ancienne, le style d'écriture, la dégradation des ouvrages ou les numérisations de faibles qualités qui rendent souvent difficile et fastidieux la tâche de transcription d'un ouvrage même pour un expert. Les systèmes traditionnels, combinant successivement une tâche d'extraction de lignes de texte, un modèle optique reconnaissant les caractères dans les images, et un modèle de langues venant apporter des corrections linguistiques [Soullard et al., 2019], sont progressivement remplacés par des systèmes dit end-to-end réalisant ces différentes tâches simultanément [Yousef and Bishop, 2020, Coquenet et al., 2022, 2023]. Ces systèmes s'appuient sur des architectures totalement convolutives [Yousef and Bishop, 2020] ou sur des Transformers [Coquenet et al., 2023] qui s'imposent dans de nombreux domaines comme celui de la reconnaissance d'écriture grâce à l'efficacité des mécanismes d'attention qu'elles incluent [Barrere et al., 2024, Kang et al., 2022]. Ces systèmes ont ainsi de grandes capacités à analyser des images et reconnaître du texte. Bien que très performants sur le ou les types de documents sur lesquels ils ont été appris, ces systèmes peuvent néanmoins souffrir d'un manque de capaciter à généraliser à de nouveaux documents. Ceci est notamment lié au manque de données supervisées et variés pouvant être exploitées pour l'entraînement de ces modèles.

	L'objectif de cette thèse sera de s'intéresser à de l'apprentissage auto-supervisé afin de pouvoir tirer profit d'exemples non supervisés pour l'apprentissage de ces modèles. Ces dernières années, de nombreux travaux se sont intéressés aux approches auto-supervisées afin de pré-entraîner les réseaux. Partant d'une tâche prétexte ne nécessitant pas d'annotations manuelles, le modèle acquiert tout de même de fortes capacités de modélisation. Ces approches sont à la base de tous les modèles de fondation de la litérature, qu'ils soient basés sur du texte (GPT-3 [Brown et al., 2020]), de l'image (DINO [Caron et al., 2021]) ou les deux (CLIP [Radford et al., 2021]). Les tâches prétextes peuvent être diverses : prédiction du token suivant pour GPT-3, apprentissage contrastif pour CLIP et DINO. Plus récemment, des travaux [He et al., 2022] ont montré des résultats remarquables pour la tâche prétexte de reconstruction d'images partiellement masquée, dans le cadre de réseaux Transformers. Ainsi, on s'intéressera à réaliser de l'apprentissage auto-supervisé pour une tâche de reconnaissance d'écriture manuscrite dans des images. L'objectif est ici double : a) améliorer les capacités en généralisation d'un système générique entraîné sur de nombreux exemples variés ; b) permettre de spécialiser un système sur un corpus non étiqueté pour renforcer les capacités de transcription de ce corpus. L'apprentissage auto-supervisé devra également permettre de réaliser ces objectifs de généralisation et de spécialisation en minimisant les quantités de données annotées nécessaires.
Positionnement thématique scientifique INSA	☐ **Numérique et communications sécurisés et responsables**
Adéquation du projet avec le critère 1	**(visibilité de la recherche)** **(1/2 page maximum)** Ces travaux de recherche s'inscrivent dans les axes de recherche de l'équipe SHADoc de l'IRISA sur l'analyse de documents et la reconnaissance automatique d'écriture manuscrite. Ils s'inscrivent dans la continuité d'une thèse récemment soutenue (Killian Barrere), co-encadrée par 2 des 3 co-encadrants de cette thèse, et pour laquelle un projet de valorisation des travaux scientifiques est en discussion avec la SATT suite à l'intérêt d'entreprises. Ce travail présente donc un potentiel de valorisation à terme.
Adéquation du projet avec le critère 2	**(Qualité de l'encadrement et de l'environnement de recherche)** **(1/2 page maximum)** L'équipe SHADoc est spécialisé sur l'analyse de documents et la reconnaissance automatique d'écriture manuscrite. Bertrand Couasnon travaille sur ce sujet depuis de nombreuses années au sein de l'équipe (anciennement ImaDoc et IntuiDoc). Il encadre ou a encadré 3 thèses récentes en lien avec la thématique du sujet de thèse. Yann Soullard, depuis son recrutement en 2019, a intégré l'équipe et collabore sur ce domaine sur lequel il a acquis une expertise depuis près de 8 ans. Yann Soullard encadre ou a encadré 3 thèses récentes en lien avec la thématique du sujet de thèse et souhaite préparer l'obtention d'une HDR. Denis Coquenet a réalisé sa thèse au LITIS sur ce domaine. Nouvellement recruté au sein de l'équipe en septembre dernier, il vient apporter son expertise pour mener à bien ce sujet.
Adéquation du projet avec le critère 5	**(Lien avec les objectifs de responsabilités sociétales et développement durable)** **(1/2 page maximum)** **ODD Nations Unies**