Sujet de thèse en Informatique

Titre de la thèse : **Apprentissage pour la détection de pathologies grâce à la capillaroscopie**

- Unité de recherche : Limics - Laboratoire d'Informatique Médicale et d'Ingénierie des Connaissances en e-Santé (UMR_S 1142)
- Discipline : Informatique
- Directeur de thèse : Robin DHOTE, PU-PH et Thomas GRÉGORY, PU-PH.
- Co-encadrants : Mélanie COURTINE, MCF Informatique et Julien CULERRIER, Chef de clinique.
- Contact : melanie.courtine@univ-paris13.fr
- Domaine de recherche : Intelligence artificielle, Aide à décision, Images médicales.
- Mots clés : Apprentissage profond, Traitement d'images, Reconnaissance d'objets, Images médicales, Médecine interne.

Contexte médical

La **capillaroscopie** (NFC) est un examen traditionnel d'évaluation de santé, qui fournit un moyen direct et non invasif d'évaluer la structure microvasculaire des capillaires, au niveau du pli des ongles des mains. L'intégrité microvasculaire joue un rôle important dans le maintien de la microcirculation et facilite les échanges dans l'organisme. Des études antérieures montrent que la capillaroscopie est un examen prometteur dans la détection précoce des pathologies comme le diabète, la polyarthrite rhumatoïde et la sclérose systémique, mais aussi les affections cardiovasculaires et rhumatologiques auto-immunes [EbadiJalal2025, Ozturk2025].

Classiquement, les images de capillaroscopie sont analysées visuellement par un médecin, qui va décrire les **informations observées** sous forme binaire, catégorielle ou numérique et ensuite utiliser ces informations et des référentiels d'expertise, comme la classification de Cutolo [Cutolo2000], pour établir son diagnostic. Cette approche manuelle pose plusieurs problèmes. Elle demande aux médecins beaucoup de temps et une grande attention pour identifier les spécificités des capillaires et différencier les différentes architectures rencontrées. Chacun forme subtil ou complexe dans les images NFC peuvent révéler des problèmes de santé sous-jacents.

Contexte scientifique

Des travaux sur l'analyse automatique des images de capillaroscopie ont été réalisés au cours des dernières années [Venkatapathiah2023]. Il s'agit d'approches basées sur **le traitement des images** pour segmenter les capillaires et calculer leurs densités et sur l'**apprentissage supervisé** (classique ou profond) pour différencier les sujets sains des sujets pathologiques ou classer les images en fonction de leurs annotations. Toutes les images utilisées dans ces projets ont été annotées manuellement et souvent vérifier par plusieurs experts afin d'harmoniser les annotations. Les performances de ces modèles peuvent être considérées comme faible, dans le contexte de la santé, puisque la précision varie entre 0.60 et 0.95 et le rappel ne dépasse jamais les 85%.

Les dernières approches d'**apprentissage profond** (*Deep Learning*) ont montré leurs performances sur de nombreuses problématiques liées à l'analyse d'images médicales [Ker2017]. Mais elles ont aussi montré leurs capacités d'annotation, avec l'**apprentissage auto-supervisé**. Ces technologies ouvrent de nouvelles portes, qui peuvent se révéler la clé dans le cadre de l'analyse des images de capillaroscopie.

Problématique

Comment reconnaître sur des images de capillaroscopie les capillaires et les hémorragies ? Comment en extraire des descripteurs ? Comment associer ces données apprises avec les référentiels d'expertise ? Comment déduire un diagnostic avec une probabilité d'erreurs très faible ?

Défis scientifiques et axes de recherche de la thèse

Cette thèse a pour objectif la conception, l'optimisation et le déploiement d'un **modèle complexe pour la détection d'anomalies capillaires et des pathologies induites**. Les approches simples ont montré leur limite dans leur capacité à gérer des données hétérogènes (images, données binaires, catégorielles et numériques). Une approche plus complexe mélangeant différents modèles basés sur des architectures optimisées pour chacune des données et co-habitant dans un même système doit permettre d'améliorer les performances.

Cette thèse s'articulera autour de plusieurs axes scientifiques :

1) Développer une méthode de **traitement d'images** pour prétraiter les images et vérifier leur **qualité** ;
2) Définir un modèle basé sur l'**apprentissage auto-supervisé** pour **générer des descripteurs** (type morphologie, diamètre, visibilité, distribution, longueur, densité, présence de microhémorragies …). L'apprentissage auto-supervisé permettra de pré-entraîner un encodeur sur les images de capillaroscopie non étiquetées en générant des pseudo-étiquettes (ou descripteurs), améliorant ainsi la représentation des données. Le modèle sera ensuite adapté via un *fine-tuning* sur des données labellisées pour renforcer sa capacité à détecter les descripteurs des anomalies ;
3) Utiliser l'**apprentissage non supervisé** pour classer les patients (décrits par les images et les descripteurs induits) grâce à un ou des réseaux de neurones convolutifs pré-entrainés ou non. Le(s) modèle(s) proposera l'affectation d'une classe à chaque patient et nous étudierons les relations existantes entre ses résultats et les annotations définies par les experts dans nos bases annotés ;
4) Si plusieurs modèles complémentaires sur les données se révèlent performants, concevoir une approche basée sur le **mélange des modèles** pour optimiser les performances générales ;
5) **Optimiser et évaluer** le modèle obtenu sur de nouveaux jeux de données.

L'innovation dans cette thèse ne sera pas forcément dans les algorithmes utilisés à chaque étape, mais plutôt dans l'association de différents modèles optimisés dans une tâche donnée afin d'améliorer les performances globales du système dans la prédiction précoce de pathologies. De plus, les référentiels d'expertise devront être pris en compte dans ce processus pour enrichir les données.

Données

Le service de médecine interne de l'AP-HP Avicenne, dirigé par le Professeur Robin DHOTE, disposent déjà d'une base d'images de capillaroscopie (433 sains et 278 atteints d'anomalies). Une demande au comité d'éthique de l'AP-HP Hôpitaux Universitaires Paris Seine-Saint-Denis est en cours.

Ces données seront complétées par des données libres ayant permis de démontrer les liens entre les capillaires et la détection précoce de maladie, comme Capi-detect[1] et Rheumatology_RHE_22-2222[2].

Bibliographie

[Cutolo2000] Cutolo M, Sulli A, Pizzorni C, Accardo S (2000) *Nailfold videocapillaroscopy assessment of microvascular damage in systemic sclerosis*. J Rheumatol, 27:155—60.

[EbadiJalal2025] Ebadi Jalal M, Emam OS, Castillo-Olea C et al. (2025) *Abnormality detection in nailfold capillary images using deep learning with EfficientNet and cascade transfer learning*. Sci Rep 15, 2068. Doi: 10.1038/s41598-025-85277-8.

[Ker2017] Ker J, Wang L, Rao J, Lim T (2017) *Deep Learning Applications in Medical Image Analysis*. IEEE Access, Doi: 10.1109/access.2017.2788044.

[Ozturk2025] Ozturk L, Laclau C, Boulon C, Mangin M, Braz-ma E, Constans J, Dari L, Le Hello C (2025) *Analysis of nailfold capillaroscopy images with AI.* Microvascular Research, *Vol 157, Doi : 0.1016/j.mvr.2024.104753.*

[Venkatapathiah2023] Venkatapathiah K, Nanda P, Shetty M, Swamy V, Awasthi K (2023) *Deep learning based object detection in nailfold capillary images*. IAES International Journal of Artificial Intelligence (IJ-AI). 12. 931. Doi : 10.11591/ijai.v12.i2.pp931-942.

[1] https://github.com/capi-detect

[2] https://gitlab.com/manchester_nailfold/papers/rheumatology_rhe_22-2222