



UNIVERSITÉ DE BOURGOGNE EUROPE

Ecole Supérieurs d'Ingénieurs Numérique et Matériaux

Polytech Dijon Tel : 03 80 39 60 09
<https://polytech.ube.fr/>



Rapport de fin de stage 4A

Mise en place d'algorithmes de Deep Learning pour la segmentation automatique de structures rénales

Présenté par :
Marel Sèdégbé AGONGLO

—

Tuteur de stage : Patrick BARD
Enseignant référent : Albert DIPANDA

Année Académique : 2024-2025

Sommaire

Remerciements	ii
Résumé	iii
Abstract	iv
Introduction	1
1 L'entreprise	4
2 Revue de Littérature	6
3 Bilan du stage	14
Conclusion	16
Bibliographie	18
Table des matières	19

Remerciements

Je tiens à exprimer ma profonde reconnaissance à **Monsieur Patrick Bard**, tuteur de stage au sein du laboratoire LEAD de l'Université de Bourgogne, pour m'avoir offert cette opportunité de stage. Son encadrement attentif, sa disponibilité et la qualité de son accompagnement tout au long de ce projet méritent également toute ma reconnaissance.

Je souhaite également adresser mes remerciements à **Monsieur Dandana Nasser** pour sa supervision et ses orientations pertinentes, qui ont permis de guider avec justesse et efficacité les choix méthodologiques et stratégiques de ce travail.

Enfin, j'exprime ma gratitude à l'ensemble des membres du LEAD pour leur accueil chaleureux et pour les échanges constructifs qui ont enrichi mon expérience tant professionnelle que personnelle.

Résumé

Ce projet de stage s'inscrit dans le cadre d'une collaboration entre le Laboratoire d'Étude de l'Apprentissage et du Développement (LEAD) de l'Université de Bourgogne et le service de néphrologie du Centre Hospitalier Universitaire (CHU) de Dijon. Il vise à améliorer la segmentation automatique de structures rénales dans des images histologiques, étape cruciale pour l'analyse et le diagnostic en néphrologie.

Le travail se positionne dans la continuité d'un projet précédent ayant utilisé l'architecture Mask R-CNN, qui, bien que performante, présentait certaines limites en termes de précision et de robustesse, notamment pour les structures complexes ou de petite taille. L'objectif principal est donc de sélectionner et d'implémenter une architecture de Deep Learning plus adaptée aux spécificités des images histologiques rénales.

Dans cette optique, le Segment Anything Model (SAM), et plus particulièrement son adaptation au domaine médical via Medical-SAM-Adapter, a été étudié et testé. Le projet comprend l'analyse de l'état de l'art, la préparation des données, l'entraînement du modèle et l'évaluation des performances selon des métriques standard telles que l'Intersection over Union (IoU) et le coefficient de Dice.

Abstract

This final-year project was conducted as part of a collaboration between the Laboratoire d'Étude de l'Apprentissage et du Développement (LEAD) at the University of Burgundy and the Nephrology Department of Dijon University Hospital (CHU Dijon). The aim is to enhance the automatic segmentation of renal structures in histological images, a critical step for analysis and diagnosis in nephrology.

Building on previous work that implemented the Mask R-CNN architecture, this study addresses its limitations in terms of accuracy and robustness, especially for complex or small-scale structures. The main objective is to select and implement a Deep Learning architecture better suited to the specific challenges of renal histology images.

To this end, the Segment Anything Model (SAM), and more specifically its medical adaptation—Medical-SAM-Adapter—was explored and tested. The project includes a state-of-the-art review, data preparation, model training, and performance evaluation using standard metrics such as Intersection over Union (IoU) and the Dice coefficient.

Introduction Générale

L'imagerie médicale joue un rôle central dans le diagnostic, le suivi et l'orientation thérapeutique de nombreuses pathologies. Parmi ses nombreuses applications, l'analyse d'images histopathologiques occupe une place particulière. Elle consiste à examiner au microscope des échantillons de tissus biologiques afin d'en observer la structure cellulaire et tissulaire. Ces images, obtenues à l'aide de techniques de coloration spécifiques (par exemple l'hématoxyline-éosine-safran, ou HES), offrent une vision fine des différentes structures anatomiques, permettant de détecter d'éventuelles anomalies morphologiques.

Dans le domaine de la néphrologie, l'histologie rénale est utilisée pour analyser des biopsies de reins. Ces biopsies sont essentielles pour diagnostiquer certaines maladies rénales, en évaluer la sévérité et guider les décisions thérapeutiques. Les structures rénales d'intérêt incluent principalement :

- les **glomérules**, responsables de la filtration du sang,
- les **tubules**, impliqués dans la réabsorption et la sécrétion,
- les **artères et capillaires**, permettant la vascularisation,
- les **zones fibreuses ou inflammatoires**, indicatrices de lésions.

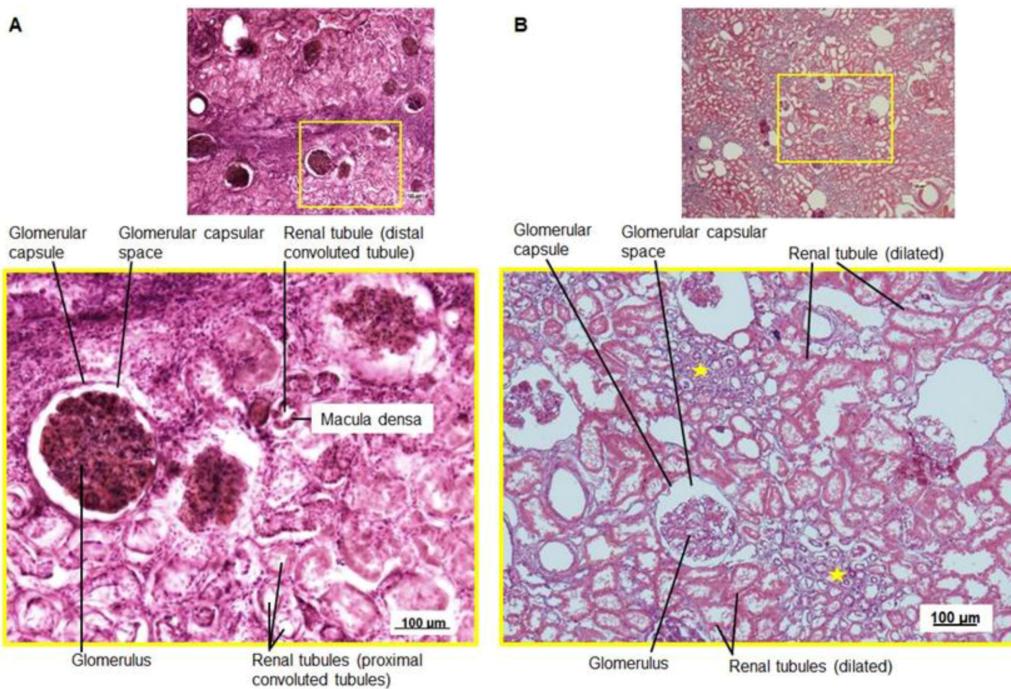


Figure 1: Exemples de structures rénales observables dans une image histopathologique (glomérules, tubules, vaisseaux).

La qualité de l'analyse histopathologique dépend en grande partie de la capacité à identifier précisément ces structures dans les images, ce qui rend la **segmentation** particulièrement cruciale. La segmentation est le processus qui consiste à diviser une image en régions homogènes, chacune correspondant à un objet ou une structure identifiable.

Il existe plusieurs types de segmentation (Figure 2) :

- **Segmentation binaire** : distinguer seulement l'objet d'intérêt de l'arrière-plan.
- **Segmentation sémantique** : classer chaque pixel selon une catégorie (glomérule, tubule, etc.), mais sans distinguer les instances individuelles.
- **Segmentation d'instances** : identifier chaque occurrence d'un objet donné (par exemple, chaque glomérule individuellement), ce qui est l'objectif de ce projet.

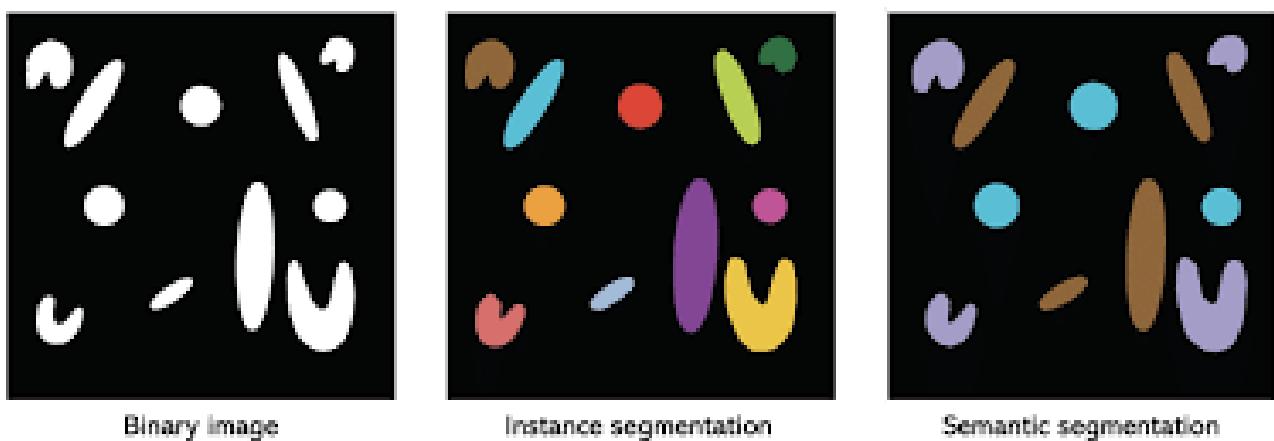


Figure 2: Comparaison entre les différents types de segmentation : binaire, sémantique et d'instances.

Depuis quelques années, l'essor du **Deep Learning** et plus spécifiquement des réseaux de neurones convolutionnels (CNN) a profondément transformé ce domaine, offrant des performances nettement supérieures aux approches traditionnelles. Des architectures spécialisées comme **U-Net**, **Mask R-CNN** ou, plus récemment, le **Segment Anything Model (SAM)**, ont montré un potentiel remarquable pour extraire automatiquement des structures complexes à partir d'images médicales.

L'objectif de ce projet est donc de développer une approche performante de **segmentation d'instances appliquée aux structures rénales**, afin d'améliorer l'analyse histopathologique et de contribuer à une meilleure assistance au diagnostic en néphrologie.

L'entreprise

Introduction

1.1 - Présentation de l'entreprise

Le présent projet de fin d'études a été mené au Laboratoire d'Étude de l'Apprentissage et du Développement (LEAD) de l'Université de Bourgogne, en étroite collaboration avec le Centre Hospitalier Universitaire (CHU) de Dijon, et plus particulièrement avec le service de néphrologie. Ce partenariat associe les compétences en intelligence artificielle et traitement d'images du LEAD à l'expertise médicale du CHU, créant un environnement pluridisciplinaire propice à l'innovation.

Le LEAD, reconnu pour ses travaux en modélisation et en analyse de données complexes, met à disposition ses ressources techniques et son savoir-faire en conception d'algorithmes d'apprentissage profond. De son côté, le service de néphrologie du CHU apporte un accès privilégié à des données médicales réelles (images histopathologique de biopsies rénales), ainsi qu'une expertise indispensable pour l'annotation, la validation et l'interprétation des résultats.

Cette synergie entre recherche scientifique et pratique clinique vise un objectif commun : améliorer l'efficacité et la fiabilité de l'analyse des images histopathologique rénales, et par conséquent, optimiser la prise en charge des patients.

1.2 Travaux antérieurs et motivations pour une nouvelle approche

Ce projet s'inscrit dans la continuité des travaux menés au sein du LEAD par d'anciens étudiants de Polytech Dijon, notamment **Adrien Jaugey** (2021), qui avaient développé une méthode de segmentation automatique basée sur l'architecture **Mask R-CNN**. Cette approche, largement reconnue dans la communauté scientifique pour ses performances en segmentation d'instances, avait permis d'obtenir des résultats encourageants sur les images histopathologiques de biopsies rénales fournies par le CHU de Dijon.

Cependant, plusieurs limites avaient été identifiées :

- **Difficulté à segmenter des structures de petite taille ou présentant des contours irréguliers** (ex. glomérules partiellement sectionnés).
- **Sensibilité à la qualité des annotations** et à la **variabilité inter-échantillons**, rendant la généralisation difficile.
- **Temps de calcul élevé**, en particulier pour des images histopathologiques en très haute résolution, ce qui limite les perspectives d'utilisation clinique en routine.

Le domaine du Deep Learning appliqué à l'imagerie médicale évolue extrêmement rapidement. Depuis la mise en place de Mask R-CNN en 2021, de nouvelles architectures plus performantes (basées sur les Transformers, les modèles fondationnels comme SAM, ou des approches hybrides comme Med-SA) ont émergé. Il est donc essentiel de poursuivre cette démarche exploratoire afin d'identifier et d'adapter les modèles les mieux adaptés à la segmentation des structures rénales dans un contexte médical.

Chapitre 2

Revue de Littérature

2.1 Le contexte

La segmentation d'images médicales constitue un enjeu central dans le diagnostic assisté par ordinateur. Dans le cas particulier des images histopathologiques rénales, cette tâche est rendue complexe par plusieurs facteurs : la **forte variabilité inter-échantillons**, la **ressemblance visuelle entre certaines structures** (par exemple entre **tubules sains** et **tubules pathologiques**), ainsi que la **taille réduite des bases de données annotées** disponibles.

Dans ce contexte, le choix de l'architecture de réseau de neurones est déterminant : un modèle trop complexe risque de sur-apprendre (*overfitting*) avec peu de données, tandis qu'un modèle trop simple manquera de précision pour délimiter correctement les structures fines.

2.2 Métriques d'évaluation des modèles de segmentation d'instances

L'évaluation de la performance d'un modèle de segmentation repose sur la comparaison entre les masques prédits par le réseau et les masques de référence (annotations manuelles fournies par des experts). Deux métriques sont couramment utilisées : l'Indice de Dice et l'Intersection-over-Union (IoU).

2.2.1 Indice de Dice

L'indice de Dice, aussi appelé coefficient de similarité de Sørensen-Dice, mesure le degré de recouvrement entre deux ensembles (ici, les pixels prédits comme appartenant à une structure et les pixels annotés comme tels). Il est défini par :

$$Dice = \frac{2 \times |A \cap B|}{|A| + |B|}$$

où A représente le masque prédit et B le masque de référence. Une valeur de $Dice = 1$ indique une segmentation parfaite, tandis qu'une valeur proche de 0 traduit une absence de recouvrement.

2.2.2 Intersection-over-Union (IoU)

L'IoU, aussi appelé Jaccard Index, mesure le rapport entre l'intersection et l'union des deux ensembles :

$$IoU = \frac{|A \cap B|}{|A \cup B|}$$

Comme pour Dice, une valeur élevée traduit une bonne concordance entre prédiction et vérité terrain. Cependant, l'IoU est généralement plus strict que le Dice car il ne double pas la pondération de l'intersection.

2.2.3 Lien entre Dice et IoU

Ces deux métriques sont fortement corrélées mais apportent des nuances complémentaires. Dans la pratique, le Dice est souvent privilégié en imagerie médicale car il est plus sensible aux petits objets, tandis que l'IoU fournit une évaluation plus stricte du recouvrement. C'est pourquoi la plupart des travaux de segmentation rapportent les deux scores pour comparer les performances des modèles.

2.3 Choix du réseau

Historiquement, les réseaux de neurones convolutionnels (CNN) ont dominé la segmentation médicale, avec des architectures phares comme **U-Net** [1], qui introduit des connexions de saut (*skip connections*) permettant de mieux préserver l'information spatiale.

Par la suite, des architectures plus avancées comme **Mask R-CNN** ont été utilisées pour la segmentation d'instances. Les travaux précédents menés à l'ex-étudiant de l'ESIREM avaient montré l'efficacité de ce modèle sur les biopsies rénales. Néanmoins, cette approche présentait plusieurs limites : difficulté à segmenter des structures petites ou irrégulières, sensibilité à la qualité des annotations, et temps de calcul élevé.

L'émergence des modèles basés sur les **Transformers** (*TransUNet*, *Swin-UNet*) a marqué un tournant dans la segmentation médicale, en exploitant le mécanisme d'attention pour mieux capturer les dépendances à longue distance dans l'image. Ces approches ont permis des gains significatifs en précision et en robustesse.

Plus récemment, un modèle fondationnel appelé **Segment Anything Model (SAM)** a été proposé par Meta AI [2]. SAM se distingue par son interface à base de *prompts* (points, boîtes, masques) et sa capacité à générer des segmentations dans un grand nombre de contextes. Toutefois, il a été démontré que ses performances chutent significativement en imagerie médicale, principalement en raison du manque de connaissance spécifique du domaine (faible contraste, frontières floues, petites lésions).

Pour pallier cette limite, une nouvelle approche baptisée **Medical SAM Adapter (Med-SA)** a été proposée [3]. Cette méthode repose sur une adaptation paramètre-éfficace (*Parameter Efficient Fine-Tuning*, PEFT), qui ne modifie que 2% des paramètres du modèle SAM, tout en intégrant des connaissances spécifiques au domaine médical. Cette approche présente plusieurs avantages :

- **Moindre coût computationnel** qu'un ré-entraînement complet.
- **Capacité à traiter plusieurs modalités médicales** (IRM, scanner, histologie, etc.).

- **Performances supérieures** aux modèles classiques (U-Net, nnU-Net, Swin-UNet) comme le montrent les résultats de la littérature.

C'est sur la base de ces éléments que nous avons choisi d'implémenter le **Med-SA** dans le cadre de mon projet de segmentation des structures histopathologiques rénales.

2.3.1 Comparaison des performances dans la littérature

La Table 2.1 issue de la publication originale de Med-SA illustre la comparaison entre ce modèle et plusieurs méthodes de référence en segmentation médicale (*ResUNet*, *nnUNet*, *TransUNet*, *Swin-UNet*, etc.) sur différentes modalités (fond d'œil, IRM, échographie, etc.).

Model	Optic-Disc		Optic-Cup		Brain-Tumor			Thyroid Nodule		Melanoma	
	Dice	IoU	Dice	IoU	Dice	IoU	HD95↓	Dice	IoU	Dice	IoU
ResUNet	92.9	85.5	80.1	72.3	78.4	71.3	18.71	78.3	70.7	87.1	78.2
BEAL	93.7	86.1	83.5	74.1	78.8	71.7	18.53	78.6	71.6	86.6	78.0
TransBTS	94.1	87.2	85.4	75.7	87.6	78.44	12.44	83.8	75.5	88.1	80.6
EnsemDiff	94.3	87.8	84.2	74.4	88.7	80.9	10.85	83.9	75.3	88.2	80.7
MTSeg	90.3	83.6	82.3	73.1	82.2	74.5	15.74	82.3	75.2	87.5	79.7
UltraUNet	91.5	82.8	83.1	73.8	84.5	76.3	14.03	84.5	76.2	89.0	81.8
FAT-Net	91.8	84.8	80.9	71.5	79.2	72.8	17.35	80.8	73.4	90.7	83.9
BAT	92.3	85.8	82.0	73.2	79.6	73.5	15.49	81.7	74.2	91.2	84.3
SegDiff	92.6	85.2	82.5	71.9	85.7	77.0	14.31	81.9	74.8	87.3	79.4
nnUNet	94.7	87.3	84.9	75.1	88.5	80.6	11.20	84.2	76.2	90.8	83.6
TransUNet	95.0	87.7	85.6	75.9	86.6	79.0	13.74	83.5	75.1	89.4	82.2
UNetr	94.9	87.5	83.2	73.3	87.3	80.6	12.81	81.7	73.5	89.7	82.8
Swin-UNetr	95.3	87.9	84.3	74.5	88.4	81.8	11.36	83.5	74.8	90.2	83.1
CoTr	95.8	87.1	85.9	75.3	87.7	81.2	12.91	82.6	74.5	90.6	83.9
TransUNet3D	-	-	-	-	88.9	81.7	11.08	-	-	-	-
Med-SA everything	97.1	89.0	85.9	76.3	88.9	81.9	10.97	85.8	77.2	92.2	83.9
Med-SA 1 point	97.4	89.5	86.8	78.8	89.1	82.0	10.38	86.3	78.7	92.6	84.1
Med-SA 3 points	97.9	89.8	87.1	79.0	89.8	82.3	10.11	86.7	79.4	93.4	84.7
Med-SA BBox 0.5	97.6	89.6	86.4	78.5	89.5	81.9	10.35	86.6	78.9	92.1	83.0
Med-SA BBox 0.75	98.3	90.1	87.5	79.9	90.5	83.0	9.50	88.4	80.4	93.0	84.2

FIGURE 2.1 : Comparaison des performances de Med-SA avec SAM et d'autres méthodes de segmentation de l'état de l'art sur différentes modalités médicales [3].

On observe que Med-SA obtient systématiquement des **scores Dice et IoU supérieurs** aux autres architectures, tout en nécessitant beaucoup moins de paramètres ajustés. Par exemple, sur la segmentation des tumeurs cérébrales, Med-SA atteint un Dice de 88.9% contre 85.7% pour SegDiff. De même, sur la segmentation des mélanomes, Med-SA surpassé TransUNet (92.2% contre 89.4%).

Ces résultats confirment l'intérêt d'adopter Med-SA dans notre contexte : il combine la puissance d'un modèle fondationnel avec une adaptation ciblée au domaine médical, garantissant ainsi une meilleure généralisation et des performances accrues.

2.4 Explication de l'architecture de Med-SA

La Figure 2.2 illustre différentes stratégies mises en place pour adapter le **Segment Anything Model (SAM)** au domaine médical. L'adaptation repose sur l'ajout de modules appelés **Adapters** (en orange), qui sont des composants légers et entraînables, insérés dans les blocs d'origine (en bleu,

figés). Ces ajustements permettent de conserver la puissance du modèle généraliste tout en introduisant des connaissances propres aux images médicales.

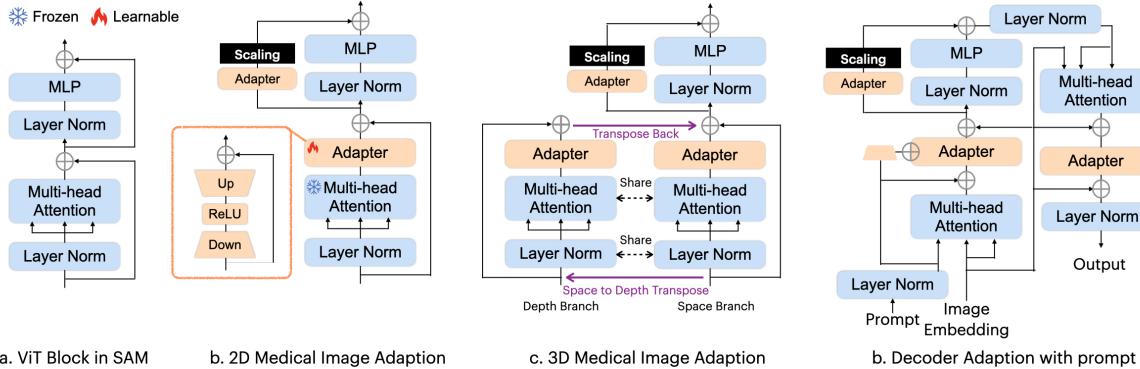


FIGURE 2.2 : Schéma illustrant les différentes adaptations de SAM pour le médical : (a) bloc ViT original dans SAM, (b) adaptation 2D pour l'imagerie médicale, (c) adaptation 3D pour données volumétriques, (d) adaptation du décodeur guidé par prompts.

(a) ViT Block in SAM

SAM repose sur un **Vision Transformer (ViT)** composé de trois éléments principaux :

- **Layer Norm** : normalisation des activations pour stabiliser l'entraînement,
- **Multi-Head Attention** : capture des relations à longue distance entre pixels,
- **MLP (Multi-Layer Perceptron)** : transformation non linéaire des représentations.

(b) 2D Medical Image Adaptation

Pour les images médicales **2D** (IRM en coupe, histologie), un **Adapter** est inséré dans le bloc ViT. Il suit une structure simple *Up* → *ReLU* → *Down* :

- **Up** : expansion de la dimension pour capter davantage de variations,
- **ReLU** : transforme toutes les valeurs négatives en 0 et gardent toutes les valeurs positives comme elles sont,
- **Down** : réduction de dimension pour revenir au format original.

Cela permet au modèle d'apprendre les spécificités médicales tout en conservant la base de SAM.

(c) 3D Medical Image Adaptation

Pour des données **volumétriques (IRM 3D, scanner)**, deux branches sont introduites :

- **Depth Branch** : traite l'information en profondeur (axe Z),
- **Space Branch** : capture les dépendances spatiales classiques (axes X-Y).

Des mécanismes **D-to-S Trans** et **S-to-D Trans** assurent l'échange d'informations entre les deux branches, permettant une meilleure cohérence inter-coups et une compréhension plus fine du volume complet.

(d) Decoder Adaptation with Prompt

La sortie de segmentation est produite par un **Decoder** guidé par des **prompts** (points, boîtes ou masques). Med-SA enrichit cette étape avec un **Hyper-Prompting Adapter**, qui améliore :

- l'alignement entre représentations d'images et instructions données par l'utilisateur,
- la précision des contours segmentés,
- la détection des structures fines et difficiles à repérer.

En résumé, Med-SA conserve la puissance du modèle fondationnel SAM tout en introduisant des modules ciblés, permettant une adaptation **paramètre-efficace** au domaine médical (et entraînant moins de 2% des paramètres). Il est plus performant grâce aux prompts de l'utilisateur

2.5 Génération des données d'entraînement

2.5.1 Images histopathologiques

Les données utilisées dans ce projet proviennent de **biopsies rénales**, réalisées au *Centre Hospitalier Universitaire (CHU) de Dijon* dans le cadre du partenariat avec le laboratoire LEAD. Ces images histopathologiques, obtenues à l'aide de microscopes numériques, représentent la structure interne du tissu rénal à une échelle microscopique.

Elles permettent d'observer en détail :

- les glomérules, responsables de la filtration du sang,
- les tubules rénaux, impliqués dans la réabsorption et la sécrétion,
- les vaisseaux sanguins,
- ainsi que certaines anomalies liées aux pathologies (inflammations, lésions, etc.).

La Figure 2.3 illustre un exemple d'image histopathologique brute utilisée dans notre jeu de données. Ces images constituent la base du travail de segmentation : leur complexité visuelle et la variabilité entre patients justifient l'utilisation d'algorithmes avancés de Deep Learning.

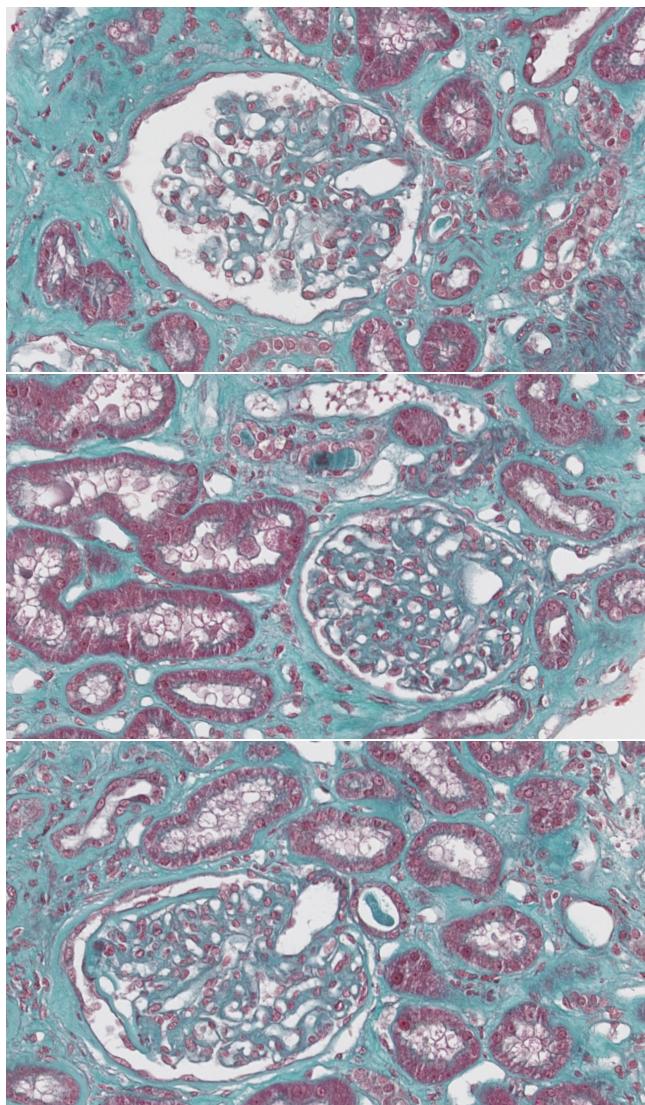


FIGURE 2.3 : Exemple d'image histopathologique issue d'une biopsie rénale.

2.5.1.1 Limites du modèle SAM en imagerie médicale

Plus récemment, un modèle fondationnel appelé **Segment Anything Model (SAM)** a été proposé par Meta AI [2]. SAM se distingue par son interface à base de *prompts* (points, boîtes, masques) et sa capacité à générer des segmentations dans un large éventail de contextes.

Cependant, plusieurs limites majeures ont été identifiées lorsqu'il est appliqué à l'imagerie médicale :

- **Absence de pré-entraînement spécifique au domaine médical** : SAM a été pré-entraîné sur des bases d'images générales (photos naturelles), qui diffèrent considérablement des images médicales en termes de contraste, de textures et de structures.
- **Données médicales limitées et spécifiques** : certaines pathologies rares ou certains tissus spécifiques (comme les biopsies rénales) disposent de peu ou pas de bases de données publiques annotées, rendant difficile l'adaptation directe du modèle.
- **Manque de robustesse face aux faibles contrastes et aux structures fines** : les frontières floues et la variabilité des échantillons histopathologiques posent un défi particulier que SAM ne parvient pas à surmonter sans adaptation.

Ces constats ont motivé le développement d'approches spécialisées telles que le **Medical SAM Adapter (Med-SA)**, visant à injecter des connaissances spécifiques au domaine médical tout en réduisant le coût d'adaptation du modèle.

2.5.2 Entrainement du Med-SA

Afin de rendre ces données exploitables par un modèle de segmentation, chaque image histopathologique a été associée à un **masque en niveaux de gris multi-classes** qui est la vérité terrain. Le masque est une image de même dimension que l'originale, où chaque pixel est attribué à une classe donnée :

- 0 : arrière-plan (fond de lames),
- 50 : tubules sains,
- 100 : tubules altérés,
- 150 : glomérule,
- 200 : vaisseaux,
- 250 : autres structures d'intérêt.

Ces masques ont été générés automatiquement à partir des annotations initiales au format xml et préparés de façon à ce qu'ils correspondent parfaitement aux images d'origine. La Figure 2.4 montre des exemples de masque en niveaux de gris obtenu pour la classe “**tubule sain**”.

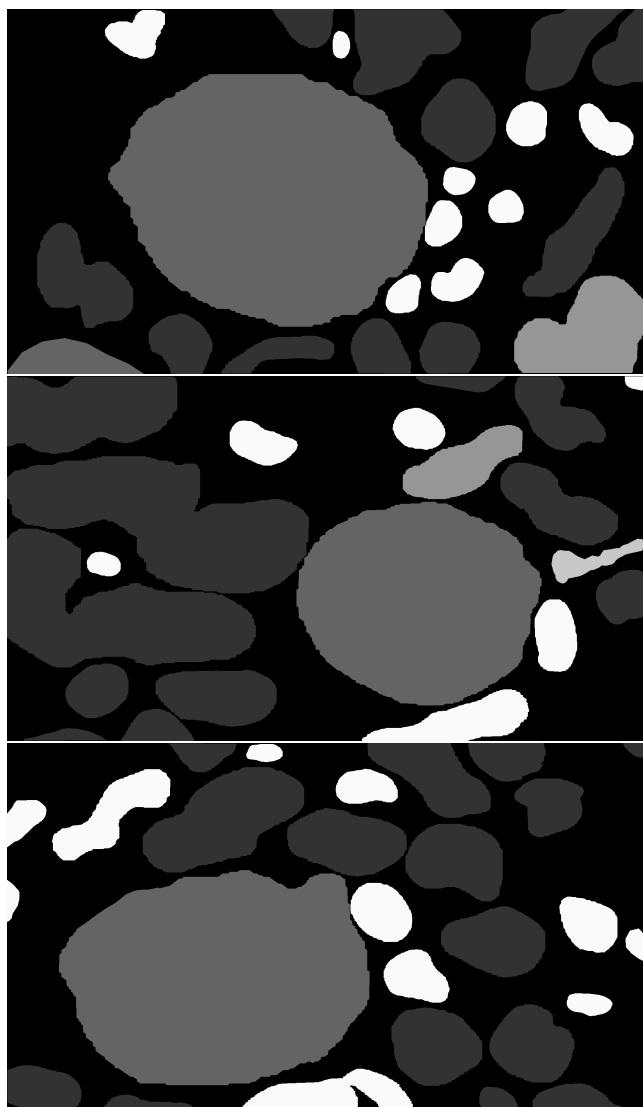


FIGURE 2.4 : Exemple de masque.

2.5.3 Prétraitement et organisation des données

Étant donné la résolution très élevée des images histopathologiques (allant parfois jusqu'à plusieurs centaines de millions de pixels), il n'est pas possible de les traiter directement par un réseau de neurones. Dans ce projet, chaque image a donc été redimensionnée et associée à son masque correspondant, afin d'obtenir un format homogène et directement exploitable pour l'entraînement des modèles (taille fixée à 256×256 pixels).

Un fichier CSV a été généré pour établir la correspondance exacte entre chaque image et son masque, permettant ainsi d'automatiser le chargement du jeu de données au sein de la classe Dataset de PyTorch.

Bilan du stage

Introduction

3.1 Résultats expérimentaux

L'entraînement du modèle **Med-SA** a été réalisé sur notre jeu de données d'images histopathologiques multiclasses. Les métriques principales utilisées pour évaluer les performances étaient le **Dice** et l'**IoU**.

- **Dice moyen** : stagnation autour de 0.18–0.20.
- **IoU moyen** : autour de 0.19–0.20.

Ces résultats traduisent une incapacité du modèle à apprendre efficacement la segmentation fine des structures rénales. Les visualisations qualitatives (Figure ??) montrent des segmentations incomplètes, avec notamment :

- Des zones non détectées ou mal délimitées.
- Une difficulté persistante à segmenter les petites structures pathologiques.

3.2 Analyse critique des performances

Plusieurs facteurs peuvent expliquer ces performances limitées :

1. **Domaine très spécifique** : Med-SA, bien que performant sur d'autres données médicales, n'a pas été pré-entraîné pour les images histopathologiques rénales complexes.
2. **Jeu de données limité et déséquilibré** : faible nombre d'images annotées et forte variabilité inter-échantillons (coloration, qualité d'image). Le jeu de donnée présentait 292 images aux totales.

3. **Complexité du problème multiclasses** : difficultés à distinguer des structures visuellement proches.
4. **Contraintes computationnelles** : limitations matérielles ayant imposé des compromis sur la taille des images et les hyperparamètres.

3.3 Axes d'amélioration

Pour améliorer les résultats, plusieurs pistes peuvent être envisagées :

3.3.1 Amélioration des données

- Collecter davantage d'images annotées en collaboration avec le service de néphrologie.
- Appliquer des techniques avancées d'augmentation de données (variation de teinte, flou, déformation géométrique).

3.3.2 Adaptation du modèle

- Réaliser un *fine-tuning* plus poussé en ré-entraînant certaines couches du ViT.
- Explorer des architectures spécialisées en histopathologie telles que HoVer-Net ou nnU-Net.
- Combiner Med-SA pour la localisation grossière avec un CNN/UNet pour la segmentation fine.

3.3.3 Optimisation des hyperparamètres

- Utiliser des fonctions de perte adaptées aux classes déséquilibrées (Dice Loss, Focal Loss).
- Ajuster dynamiquement le taux d'apprentissage (scheduler).

Bien que les performances obtenues soient limitées, ce travail met en lumière les défis spécifiques de la segmentation multiclasses en histopathologie rénale. Il constitue un point de départ pour des recherches futures orientées vers des modèles plus spécialisés, mieux adaptés aux contraintes du domaine médical.

Conclusion Générale

Ce projet s'inscrit dans la continuité des travaux menés par d'anciens étudiants de Polytech Dijon sur la segmentation automatique d'images histopathologiques rénales. L'objectif principal était d'explorer et d'implémenter une architecture de Deep Learning plus récente et plus performante que les approches précédemment utilisées, notamment le **Mask R-CNN**, afin d'améliorer la détection et la segmentation des structures rénales d'intérêt.

Pour répondre à cet objectif, nous avons étudié et implémenté le **Medical SAM Adapter (Med-SA)**, une adaptation du modèle fondationnel **Segment Anything Model (SAM)** spécifiquement conçue pour l'imagerie médicale. Cette approche repose sur un ajustement paramètre-efficace (*Parameter Efficient Fine-Tuning*) permettant de n'entraîner qu'une faible proportion des paramètres du modèle initial, tout en injectant des connaissances spécifiques au domaine médical.

Cependant, les expérimentations réalisées ont mis en évidence plusieurs difficultés. Les performances obtenues (métriques Dice et IoU stagnantes autour de 0.20) traduisent la complexité du problème : forte variabilité inter-échantillons, nombre limité d'images annotées, structures histologiques complexes, et surtout, limites intrinsèques de SAM, initialement conçu pour des images naturelles. De plus, l'adaptation à un cadre multi-classes, indispensable pour traiter simultanément plusieurs types de structures, n'a pas pu être pleinement aboutie faute de temps. Le calcul des métriques n'a ainsi pas pu être totalement adapté à ce nouveau cadre, limitant l'évaluation précise du modèle.

Malgré ces contraintes, ce travail a permis :

- de valider la faisabilité technique de l'adaptation d'un modèle fondationnel comme SAM à l'imagerie médicale,
- de mettre en évidence les obstacles spécifiques à l'histopathologie rénale,
- de poser les bases méthodologiques nécessaires pour des travaux futurs.

Plusieurs pistes d'amélioration ont été identifiées :

- enrichir et diversifier le jeu de données, éventuellement via des collaborations hospitalières,
- mettre en place des stratégies d'augmentation de données adaptées aux structures histologiques,
- adapter spécifiquement l'entraînement aux tâches multi-classes avec des fonctions de perte dédiées,
- explorer des architectures spécialisées en histologie (HoVer-Net, nnU-Net, ou hybrides ViT-CNN),

- intégrer des experts médicaux dans la boucle d'annotation et de validation.

En conclusion, ce projet illustre à la fois le potentiel et les limites des modèles de Deep Learning actuels lorsqu'ils sont appliqués à des tâches médicales complexes. Il ouvre la voie à des recherches futures visant à développer des solutions plus robustes, précises et cliniquement pertinentes. À terme, de tels outils pourraient considérablement améliorer l'assistance au diagnostic et optimiser la prise en charge des patients en néphrologie.

Bibliographie

- [1] O. Ronneberger, P. Fischer, and T. Brox, “U-net : Convolutional networks for biomedical image segmentation,” *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2015.
- [2] A. Kirillov, E. Mintun, N. Ravi, *et al.*, “Segment anything,” 2023.
- [3] X. Hu *et al.*, “Medical sam adapter : Adapting segment anything model for medical image segmentation,” 2023.

Table des matières

Remerciements	ii
Résumé	iii
	iii
Abstract	iv
	iv
Introduction	1
1 L'entreprise	4
Introduction	4
1.1 - Présentation de l'entreprise	4
1.2 Travaux antérieurs et motivations pour une nouvelle approche	4
2 Revue de Littérature	6
Introduction	6
2.1 Le contexte	6
2.2 Métriques d'évaluation des modèles de segmentation d'instances	6
2.2.1 Indice de Dice	6
2.2.2 Intersection-over-Union (IoU)	7
2.2.3 Lien entre Dice et IoU	7
2.3 Choix du réseau	7
2.3.1 Comparaison des performances dans la littérature	8
2.4 Explication de l'architecture de Med-SA	8
2.5 Génération des données d'entraînement	10
2.5.1 Images histopathologiques	10
2.5.1.1 Limites du modèle SAM en imagerie médicale	11
2.5.2 Entrainement du Med-SA	12
2.5.3 Prétraitement et organisation des données	13
3 Bilan du stage	14
Introduction	14
3.1 Résultats expérimentaux	14
3.2 Analyse critique des performances	14
3.3 Axes d'amélioration	15
3.3.1 Amélioration des données	15
3.3.2 Adaptation du modèle	15

3.3.3 Optimisation des hyperparamètres	15
Conclusion	16
Bibliographie	18
Table des matières	19

