

IA Robuste en robotique chirurgicale

Nicolas Thome

Professeur à Sorbonne Université (SU)

Chirurgie E-Novation

La robotique



Robotique chirurgicale: chirurgie du rachis

- Plate-forme de robotique chirurgicale à SU / ISIR
- Préparation de l'intervention
- « Co-manipulation » lors de l'intervention
 - Assistance du robot

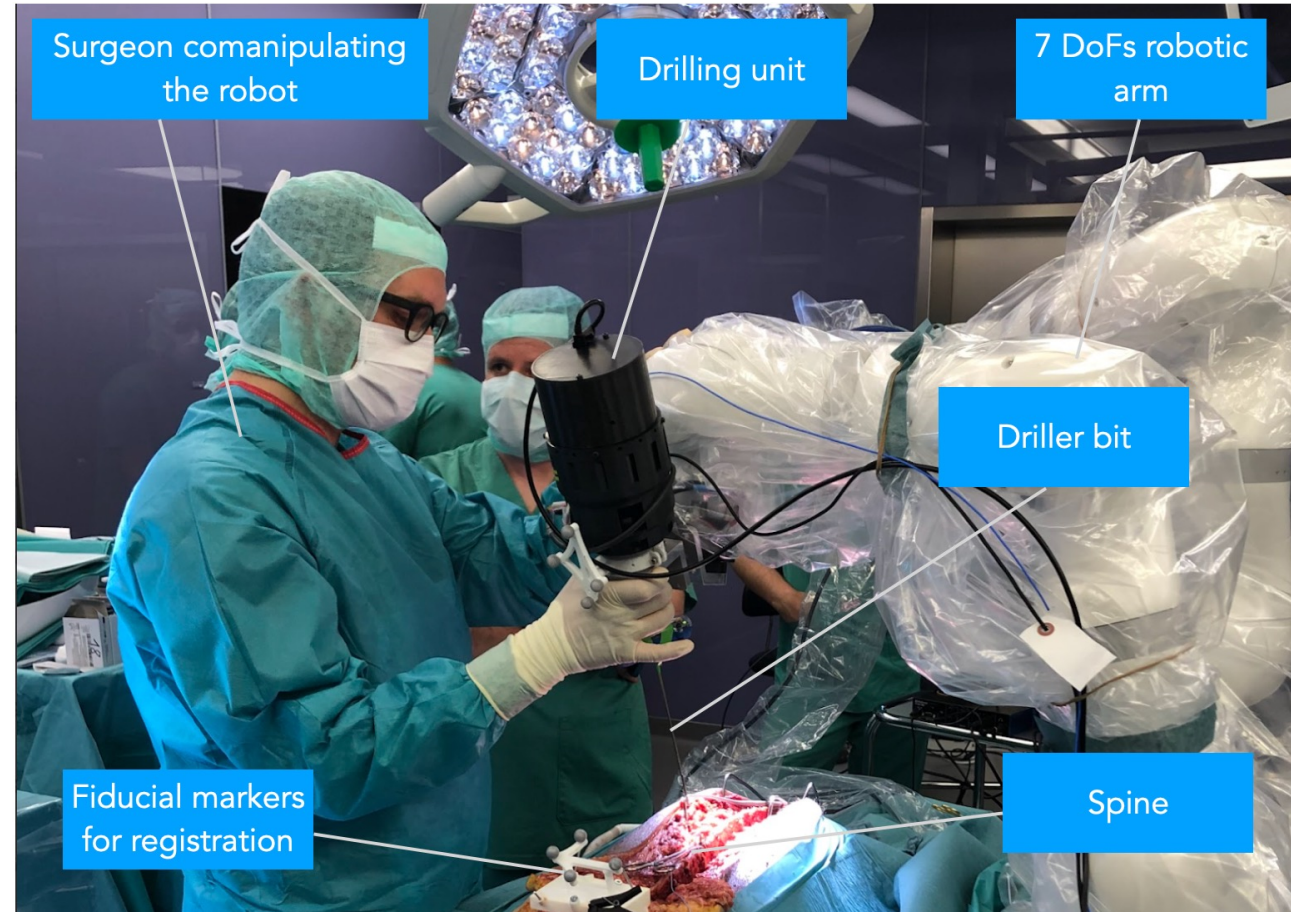
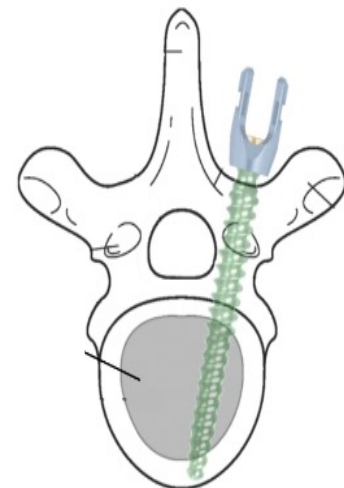
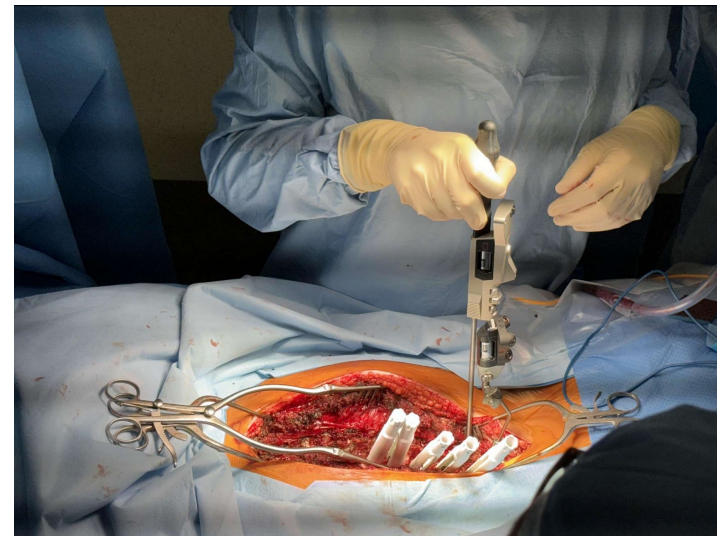


Plate-forme de chirurgie du rachis assistée par robotique

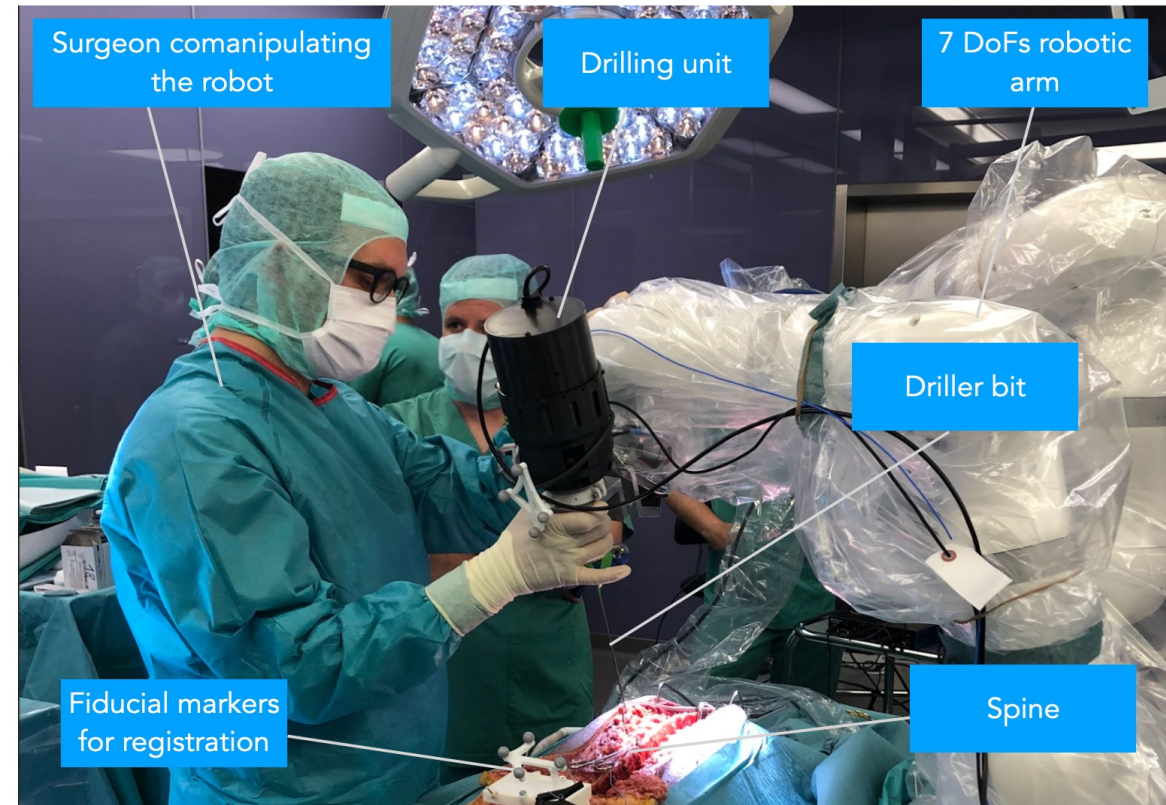
- Pré-opération : CT-scan 3D
 - Diagnostic, e.g., scoliose
 - Planification de la chirurgie , e.g., pose de vis pédiculaires (configuration précise)
- Opération
 - Patient sur le ventre
 - Seule partie arrière des vertèbres visible
 - Chirurgie de précision, sensible
 - Moelle épinière, aorte
 - ~80% des vis correctement placées (utile pour réduire les pathologies)



IA Robuste en robotique chirurgicale

Objectif : **Rendre la plateforme plus robuste**

- IA pour la navigation per-opératoire
 - ⇒ Perception visuelle, recalage pré-op
 - ⇒ Fluidifier les interactions par le langage
- Contrôleurs hybrides « analytiques + IA »
 - ⇒ Sécuriser la co-manipulation
- Quantifier l'incertitude & expliquer les étapes/ décisions automatiques
 - => acceptation patients / équipes médicales



Axe 1 : Planification multimodale soutenable

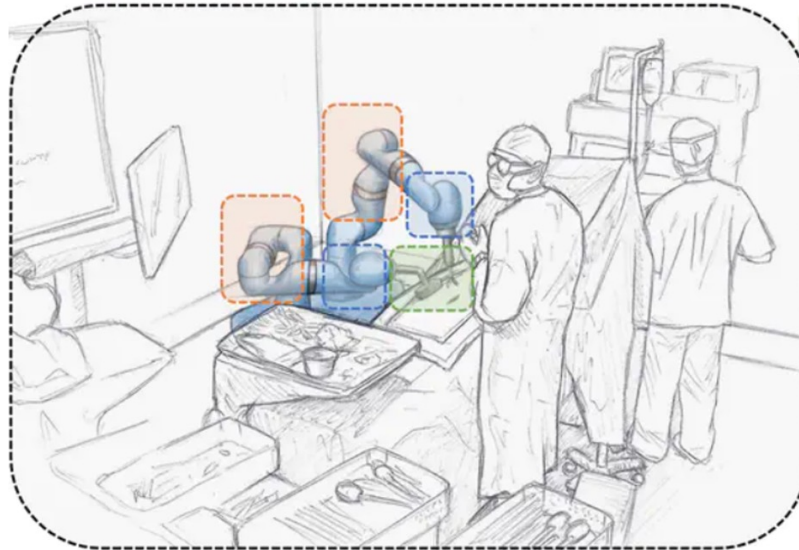
“aller à la vertèbre L5G”
=> **Comment faire ?**

Axe 2 : Contrôleurs Hybrides résiduels

“Aller face à la colonne vertébrale”

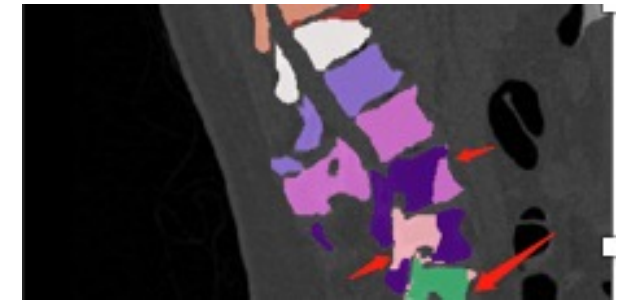
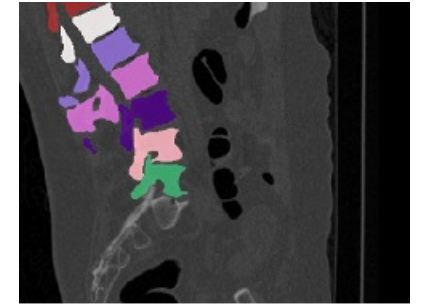
=> Contrôleurs avec physique partielle & résidu appris : frottements, contacts, vibrations

Axe 4 : Chirurgie du rachis guidée par robotique Perception visuelle



Axe 3 : Quantification d'incertitude (QI) & explicabilité

QI en segmentation :

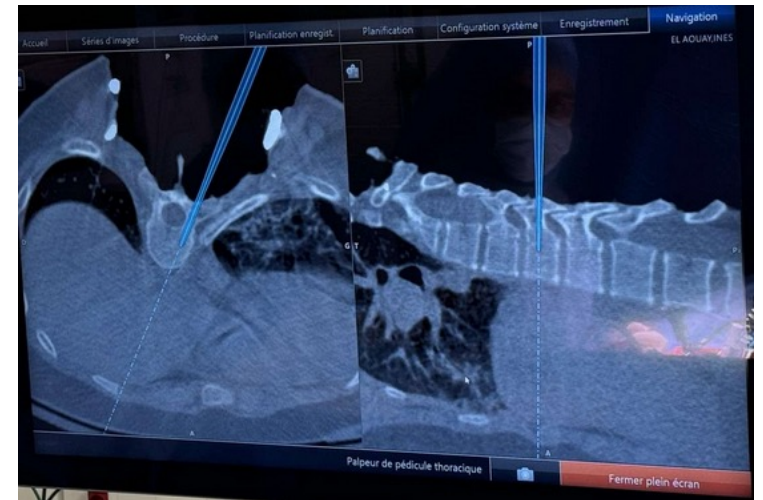
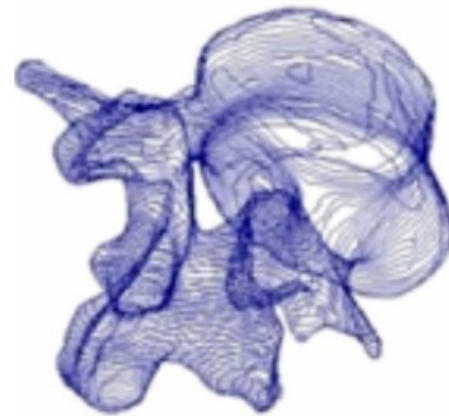
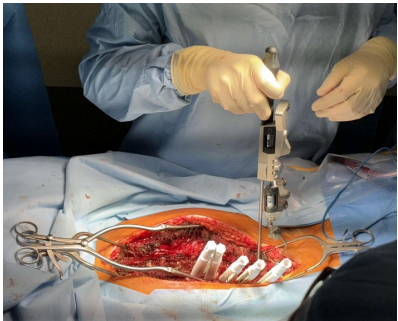


Segmentation échoue

- Pourquoi ?
- Que faire ?

Axe 4 : Perception visuelle

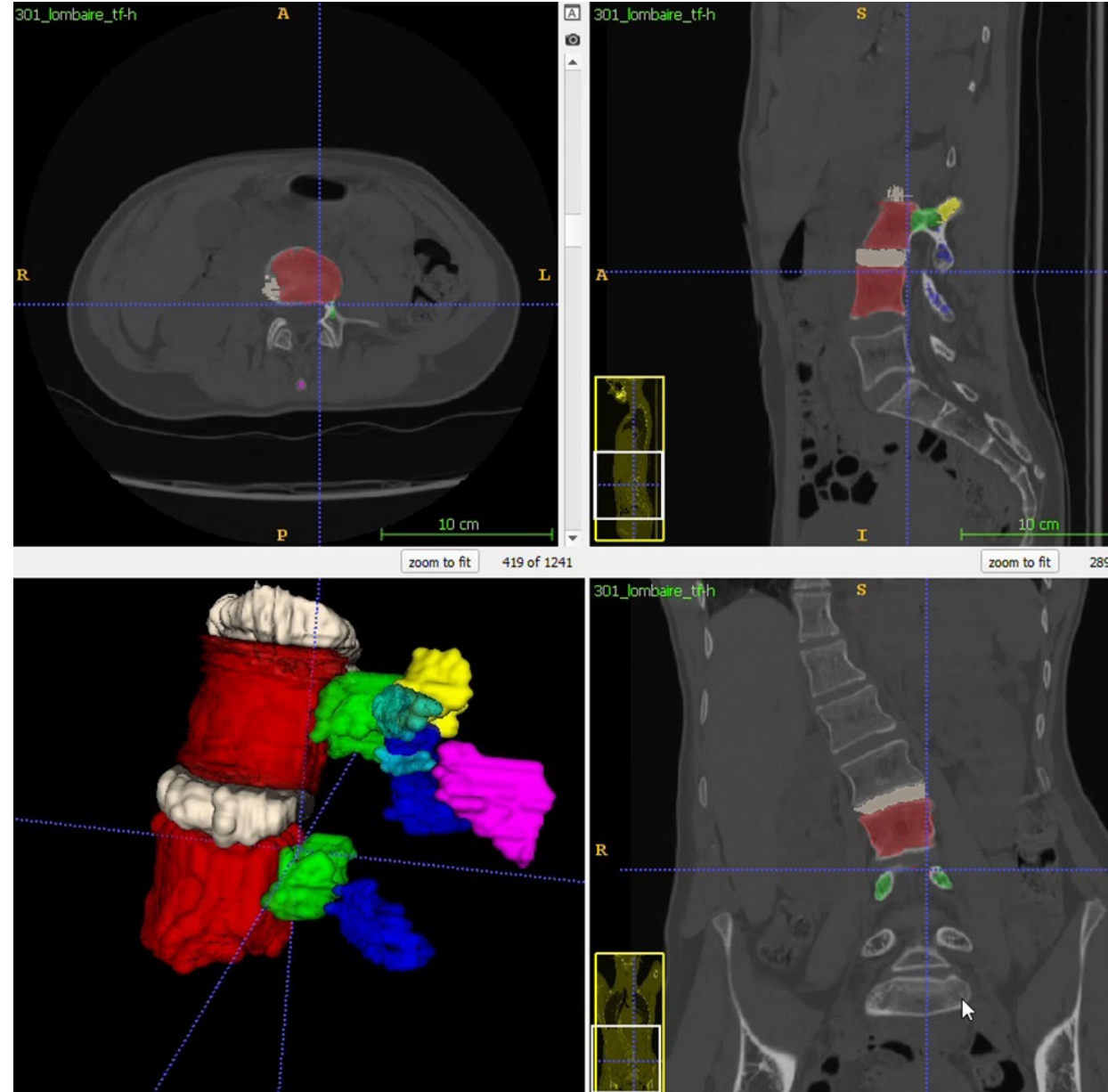
- Segmentations pré-op et per-op
 - Recalage avec données per/per-opératoire => navigation 3D
 - Segmentation des vertèbres individuelles pendant l'opération
 - Segmentation nécessaire pour l'axe 1 (planification guidée par le langage)



=> relâcher la charge mentale pendant l'opération

Axe 4 : Perception visuelle

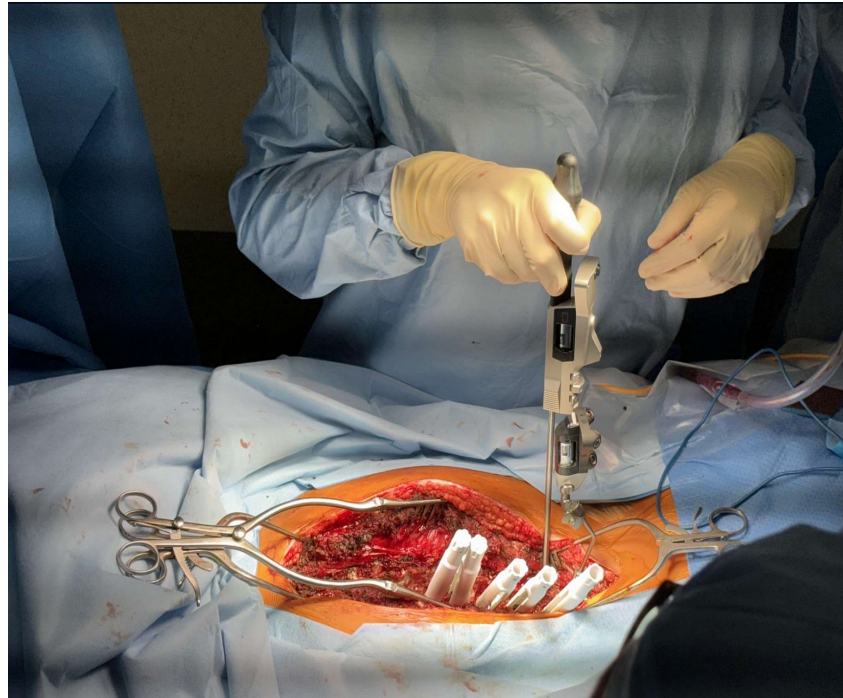
- Scanner 3D pré-opératoire : segmentation par IA
- Identification vertèbres individuelles + structures fines
 - **Corps vertébral**, disque, **pédicules**, **épineuse**, **lames vertébrales**, **apophyse transverse**, **articulaire** + canal rachidien



Axe 4 : Perception visuelle & recalage

- Collecte d'images per-opératoire : RGB-D, LIDAR, depth ?
 - Segmentation des structures per-op + recalage avec images pré-op
- **Recalage : sécuriser les opérations** (e.g., pose de vis pédiculaire)
 - Recalage des vertèbres individuelles par mise en correspondance des points des surfaces

- Assistance per-op : information 3D manquante
- Visualiser position instruments chirurgicaux / structures médicales 3D



Axe 1 : planification en robotique chirurgicale

- Planification guidée par le langage per-op

“Aller à la vertèbre L5G” => **étapes**

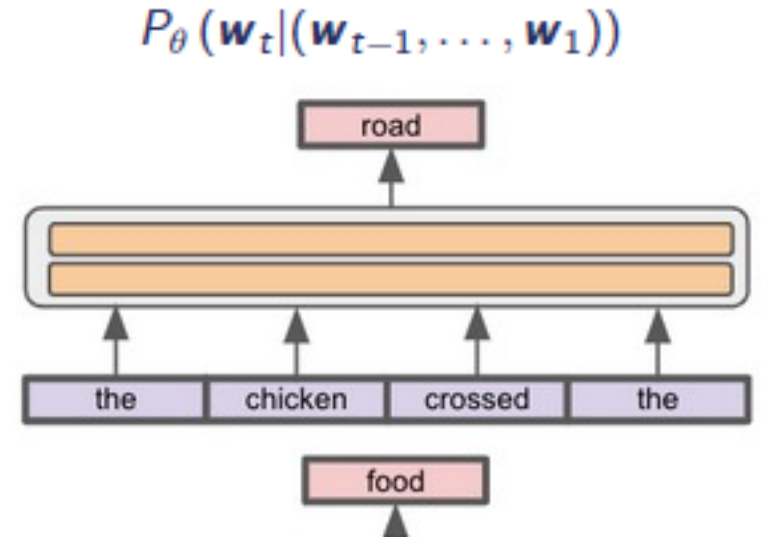
- ✓ *Aller face à la colonne vertébrale*
- ✓ *Localiser L5G dans l'image*
- ✓ *Aller en L5G*



- Fluidifier les interactions robot/humain
 - Interactions en langage naturel avec les équipes médicales

Large Language Models (LLMs) pour la planification

- LLMs, e.g., ChatGPT : entraînés à partir de corpus texte massifs
 - Prédire le mot suivant
- Capacité émergente : compréhension du langage, dialogue
 - Zero-shot, In-context learning (ICL)
- LLMs pour en planification ?
 - Pas de connaissances du monde réel
 - Raffiner («Fine-tuning») les LLMs par interaction avec l'environnement
 - Travaux récents SU [1,2]



“Aller à la vertèbre L5G”

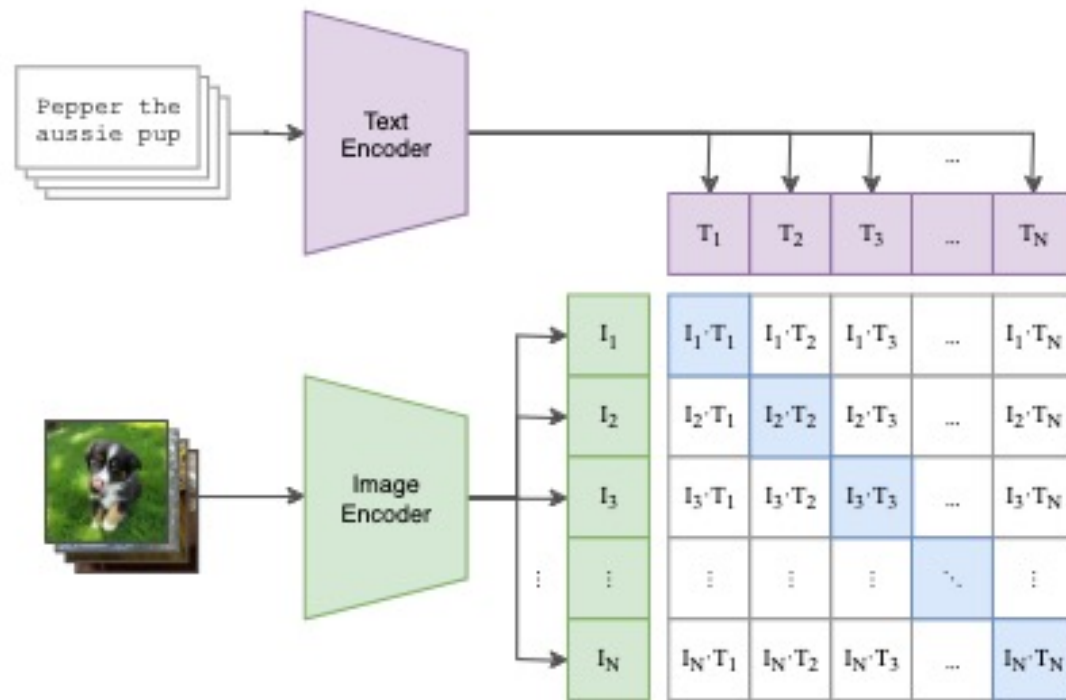
[1] T. Carta, C. Romac, T. Wolf, S. Lamprier, O. Sigaud, and P.Y. Oudeyer. Grounding large language models in interactive environments with online reinforcement learning. ICML 2023.

[2] S. Aissi, C. Romac, T. Carta, S. Lamprier, P.Y. Oudeyer, O. Sigaud, L. Soulier, N. Thome Reinforcement Learning for Aligning Large Language Models Agents with Interactive Environments: Quantifying and Mitigating Prompt Overfitting. Arxiv 2024.

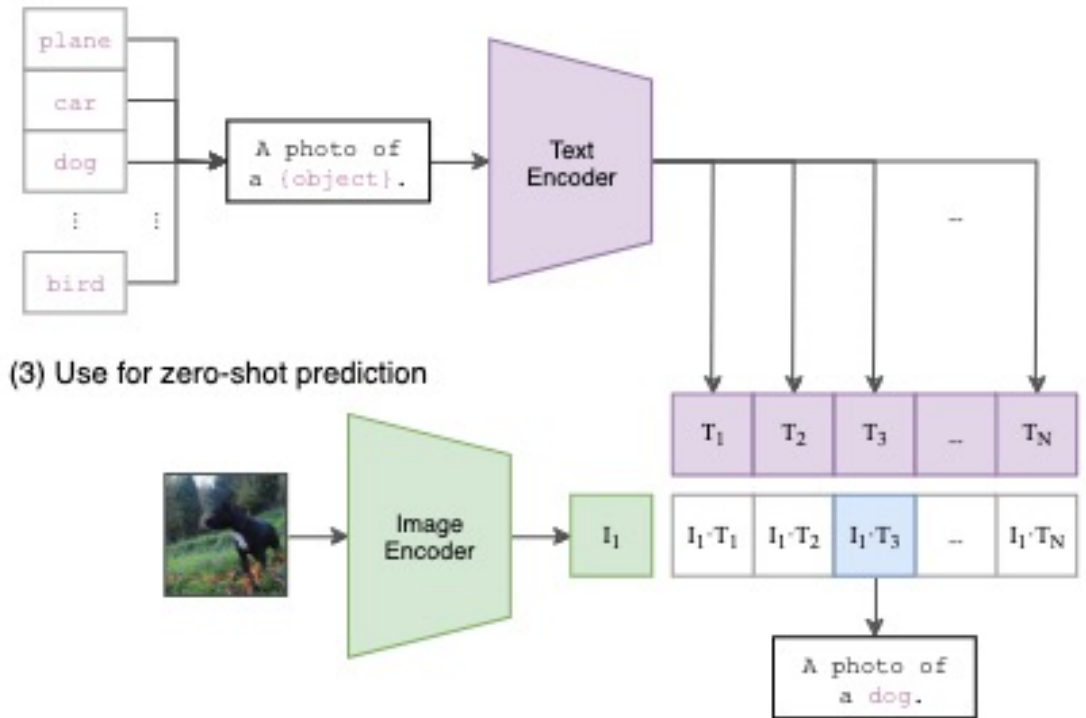
Vision-Language Models (VLMs) pour la planification

- Vision-Language Models: e.g., text—image (CLIP) [3]

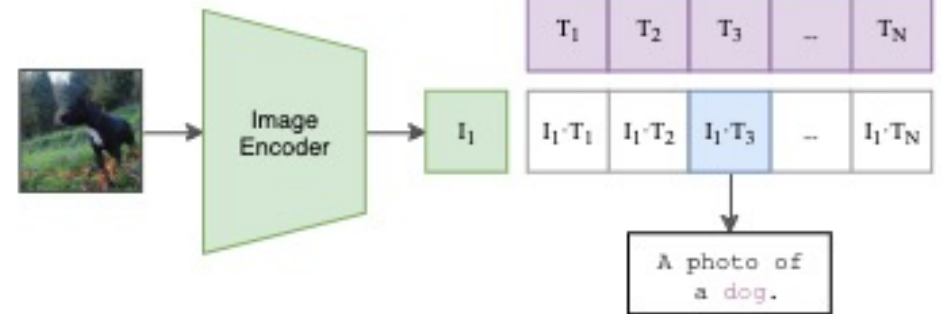
(1) Contrastive pre-training



(2) Create dataset classifier from label text

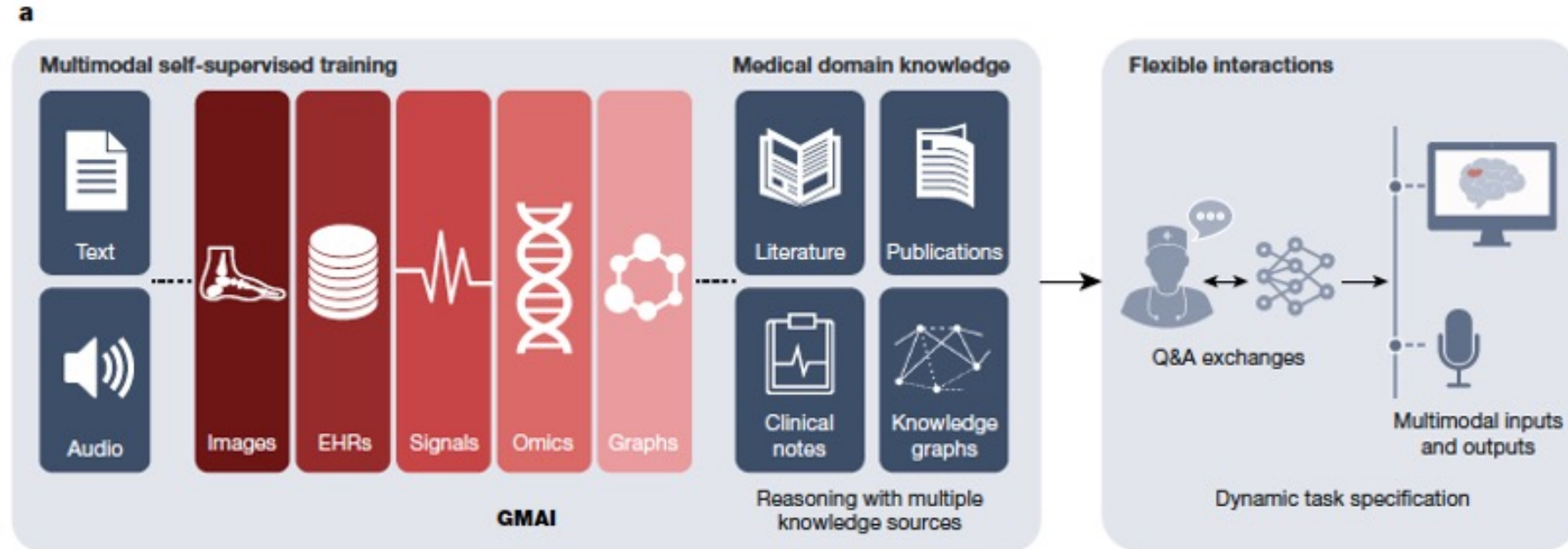


(3) Use for zero-shot prediction



(VLMs) pour la planification en robotique

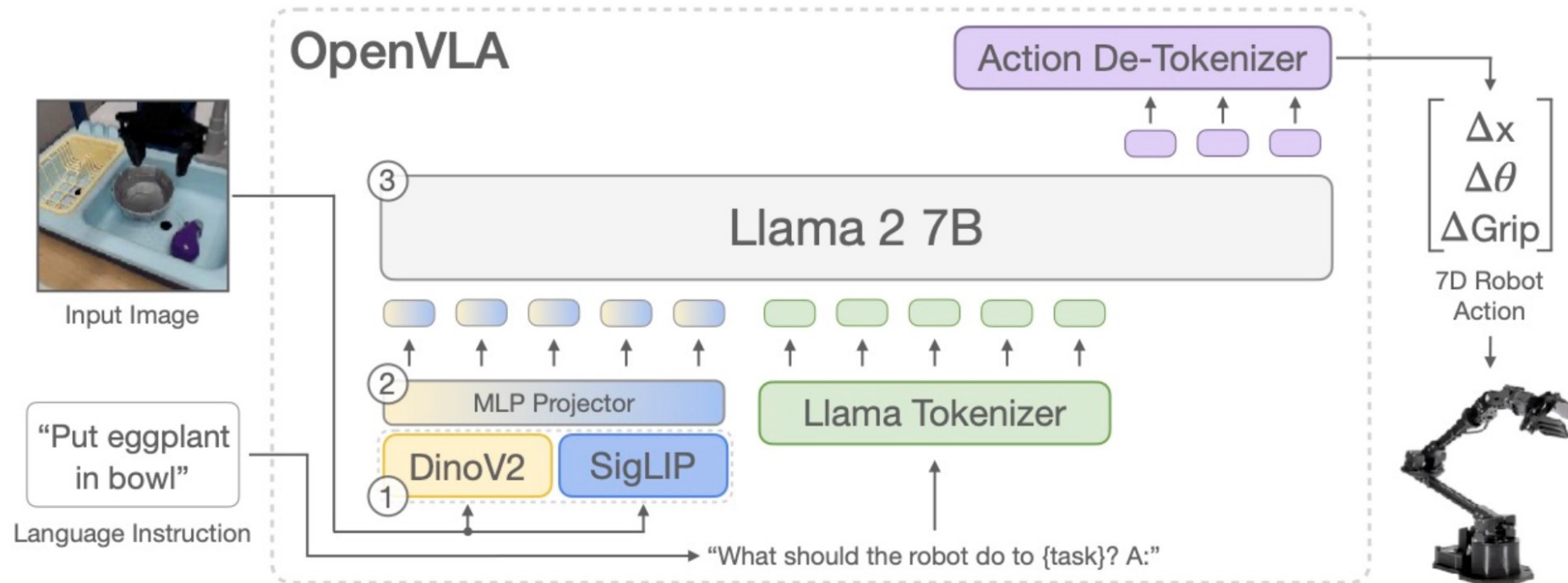
- CLIP: zero-shot prediction en santé : vers des modèles de fondation en santé[4]



For example, a clinician might say, “Check these chest X-rays for Omicron pneumonia. Compared to the Delta variant, consider infiltrates surrounding the bronchi and blood vessels as indicative signs”⁴⁰.

Modèles de fondation en robotique

- « Vision Language Action » (VLA), e.g., OpenVLA [5]
- Hot topic depuis 1 an, bcp d'acteurs sur le sujet
 - En France, Hugging Face « Le Robot »
 - Apprendre des actions robotiques génériques, dans la lignée des VLMs



Défis / enjeux scientifiques en IA

- Grounding textuel des modèles de langages
- Grounding dans la scène visuelle, lien avec la perception visuelle
 - Capacité de détecter des structures médicales précises, spécifiques au rachis
- Apprentissage joint de la planification et du contrôle
- Utiliser / adapter les modèles de fondation en robotique chirurgicale

“Aller à la vertèbre L5G” => étapes

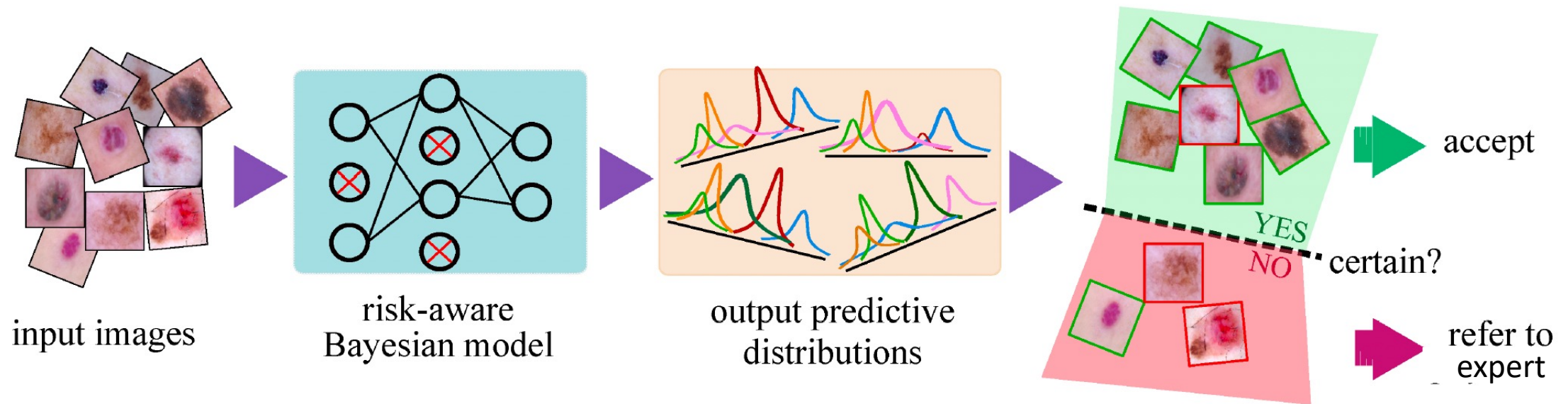
- ✓ ***Aller face à la colonne vertébrale***
- ✓ ***Localiser L5G dans l’image***
- ✓ ***Aller en L5G***



Axe 3 : Quantification de l'incertitude (QI)

QI : Robustesse & interactions

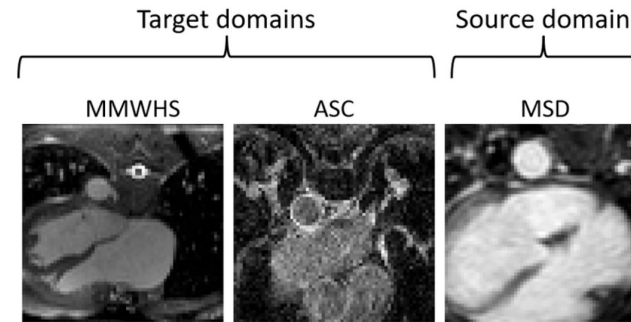
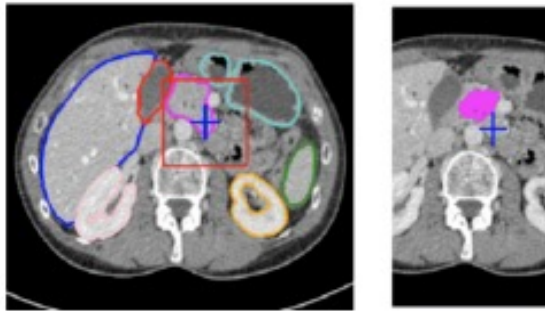
- S'abstenir de réaliser des predictions



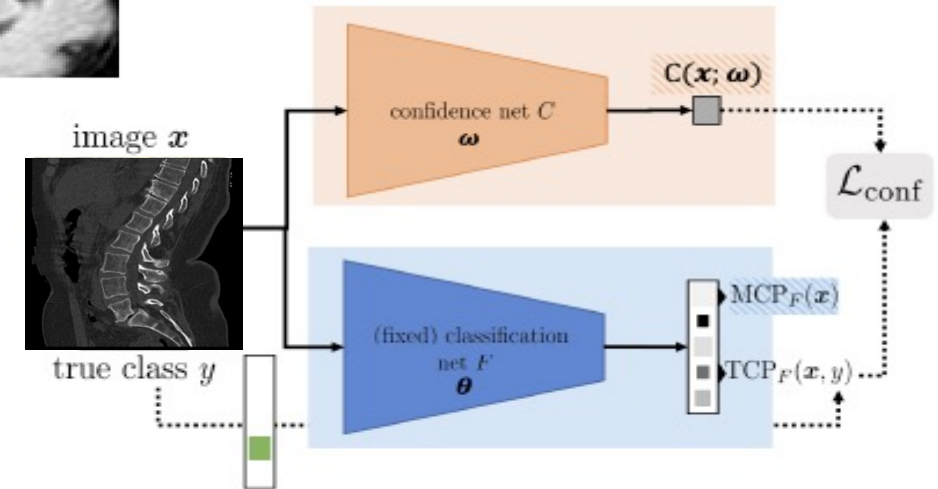
=> Informer le chirurgien, rendre la main, interagir

Axe 3 - QI : défis et objectifs

- Modèle d'IA / deep learning: difficile d'avoir une QI fiable
- Incertitude aléatoire vs épistémique : approches spécifiques



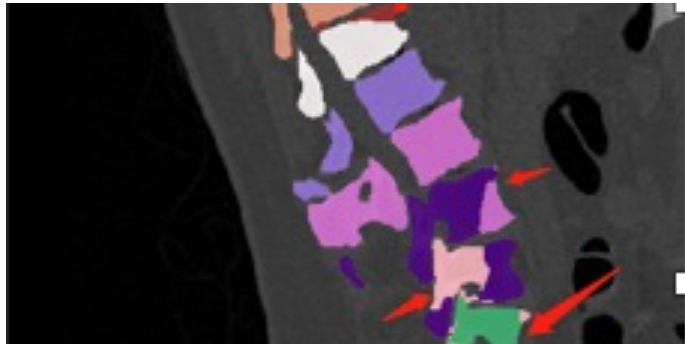
- Apprentissage de la QI [6]
- Objectifs :
 - QI unifiée
 - QI pour les données structurées



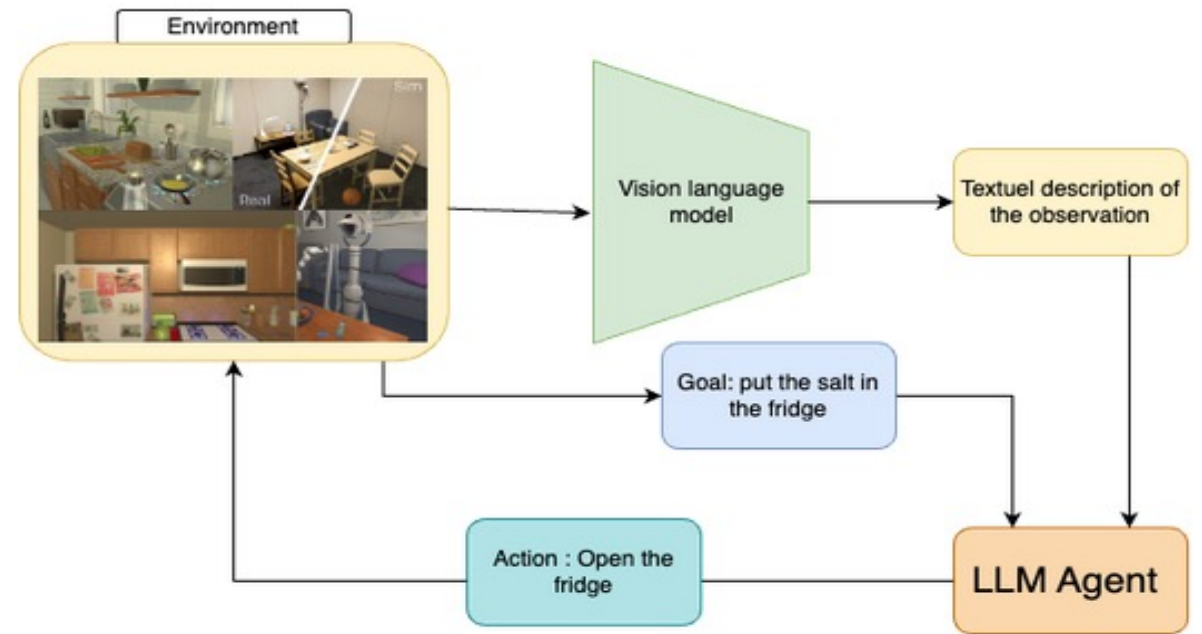
Axe 3 : QI et explicabilité

- QI : indicateurs numériques pas interprétables
- LLMs/VLMs => explication en langage des décisions et de l'incertitude associée

La segmentation a échoué



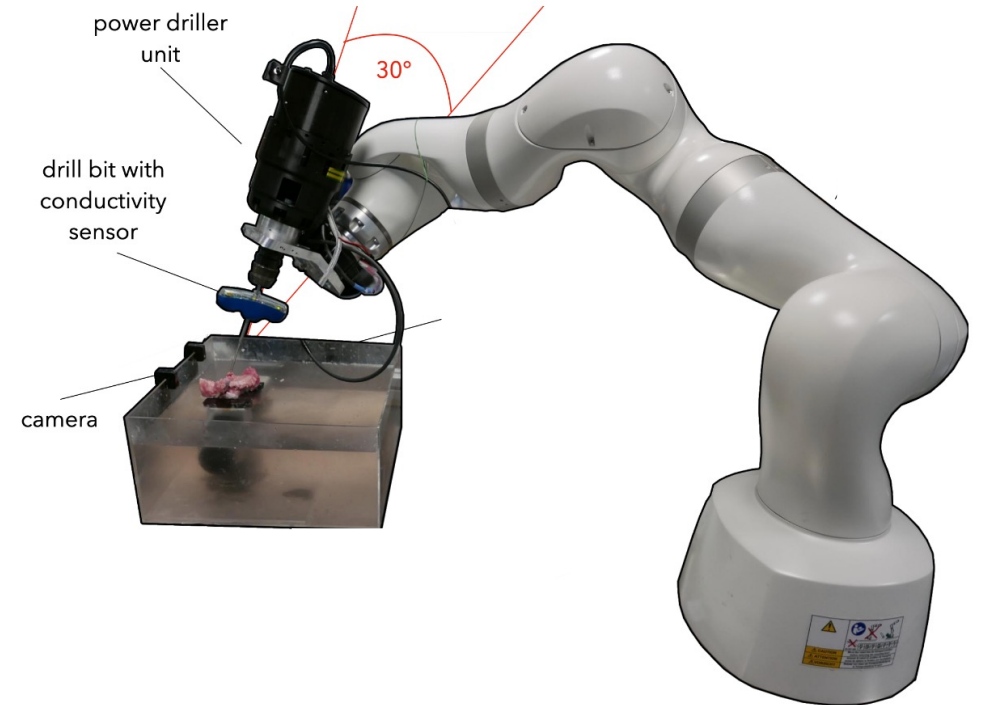
- Pourquoi ? => *car L5G est occultée*
- Que faire ? => *descendre*



- Interactions basées sur la pédagogie / pragmatisme [7]

Axe 2 : contrôleurs hybrides en co-manipulation

- Robot chirurgical 7dof : connaissances physique précises
- **MAIS** nombreuses forces impossibles à modéliser :
 - Interactions chirurgien, e.g., frottements des outils sur les structures
 - Compensation gravité, vibrations (cruciales en manipulation fine), inertie

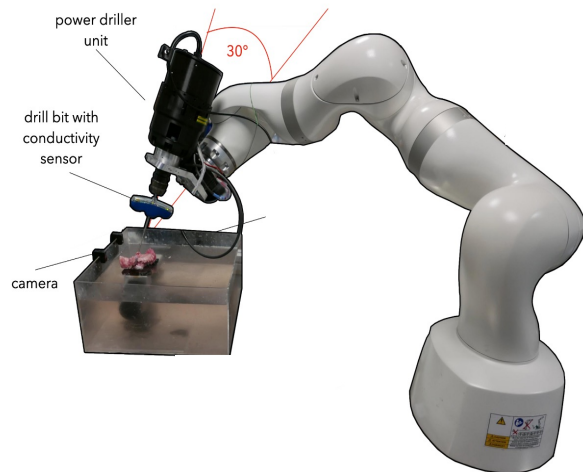


Enjeu : améliorer la précision des contrôleurs

- Toujours en co-manipulation, car tâches chirurgicales sensibles
- Mais limiter la charge mentale du chirurgien sur des tâches fastidieuses, se concentrer sur l'opération

Axe 2 : contrôleurs hybrides résiduels (CHR)

- **CHR [8]** : apprendre le complément des modèles analytiques
 - **Objectifs :**
 - Séparations de l'information physiques / apprise
 - apprendre directement sur le robot reel avec peu d'exemples
- ⇒ Méthodes Model-based RL pour apprendre le résidu
- ⇒ Les rendre plus efficaces



Merci de votre attention

- Questions ?

Axe 4 : Recalage automatique vs manuel

- **Recalage manuel**

- Pose de vis/marqueurs fiduciaires, puis scanner per-op
- Acquisition per-op capteurs caméra robot chirurgien



- **Limites & amélioration envisagées**

- Marqueurs fiduciaires : fastidieux à installer, encombre le champ opératoire
- Recalage dépend de l'image initiale : gérer les variations de géométrie ?
 - Changement de position du patient
 - Respiration, affaissement colonne vertébrale, déformations dues à la pose des vis