

L'imagerie spectrale infrarouge comme une nouvelle approche pour la reconnaissance histopathologique automatique dans le diagnostic du cancer du colon

Jayakrupakar NALLALA, Marie-Danièle DIEBOLD, Cyril GOBINET,
Valérie UNTEREINER, Olivier BOUCHE, Michel MANFAIT,
Ganesh SOCKALINGUM, Olivier PIOT

cyril.gobinet@univ-reims.fr

olivier.piot@univ-reims.fr

ganesh.sockalingum@univ-reims.fr

michel.manfait@univ-reims.fr

Equipe MéDIAN, Biophotonique et Technologies pour la Santé

Unité MEDyC, FRE CNRS 3481

UFR de Pharmacie, Université de Reims Champagne-Ardenne

1) Le cancer colorectal

- Le cancer colorectal est l'un des principaux cancers en terme d'incidence et de mortalité dans le monde (GLOBOCAN 2008)
- **Les adénocarcinomes** représentent 90-95% des cancers colorectaux
- **Causes et facteurs de risque** - L'âge, les habitudes alimentaires, le tabagisme, les maladies inflammatoires, les facteurs génétiques, etc
- **Polypes adénomateux** - (lésions pré-cancéreuses) l'un des facteurs de risque les plus importants associé au développement du cancer colorectal

2) Méthodes actuelles de dépistage et de diagnostic

- **Méthodes de dépistage** : Recherche de sang occulte dans les selles (FOBT), coloscopie, and sigmoïdoscopie, etc



Confirmation du diagnostic: HISTOPATHOLOGIE

Besoin actuel : Nouvelles méthodes de classification, basées sur des informations biomoléculaires de cellules et tissus, et complémentaires à l'histopathologie



Approches automatisées pour le diagnostic rapide

2) Méthodes actuelles de dépistage et de diagnostic

- **Méthodes de dépistage** : Recherche de sang occulte dans les selles (FOBT), coloscopie, and sigmoïdoscopie, etc



Confirmation du diagnostic: HISTOPATHOLOGIE

Méthodes candidates : infrarouge et Raman (spectroscopies vibrationnelles)

- **Imagerie spectrale infrarouge**
 - Non-destructive, et basée sur l’empreinte biochimique de cellules et tissus
 - Objective et sans marquage (pas de coloration chimique)
 - Permet une analyse directe de tissus paraffinés (réduction des traitements chimiques)

3.1) Protocole : méthodologie d'imagerie spectrale infrarouge de matrices paraffinées de tissus coliques

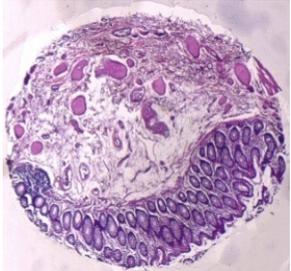
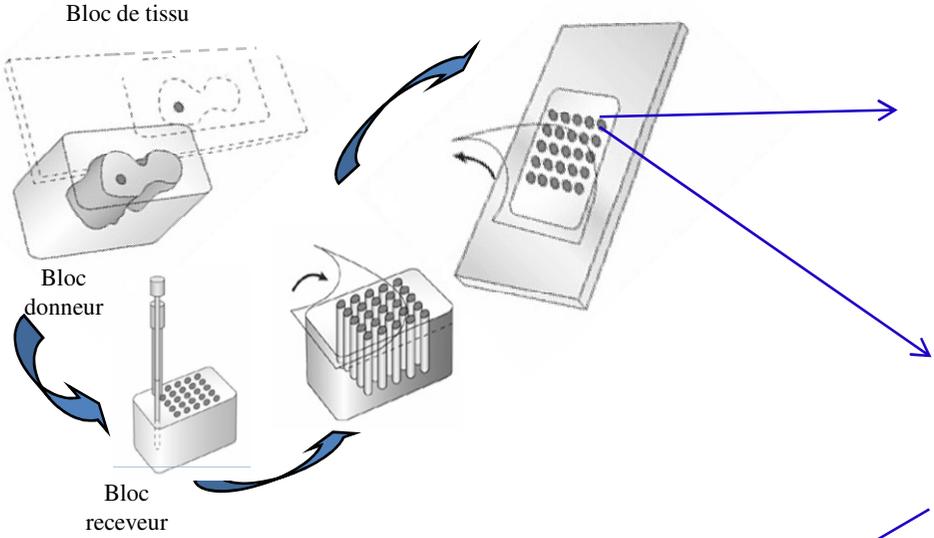
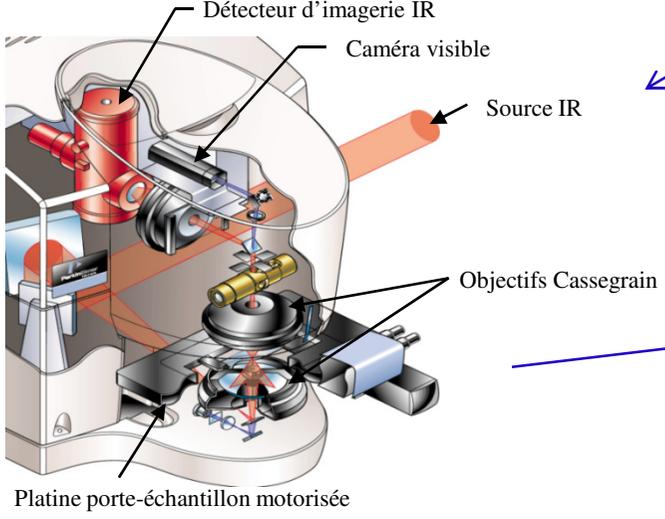


Image colorée par HPS



Image non colorée



Système d'imagerie IR

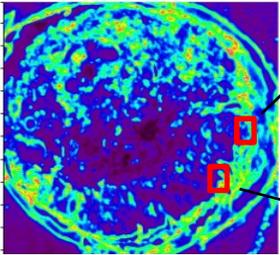
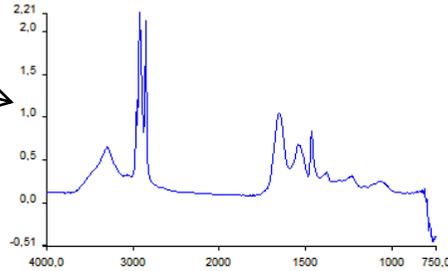
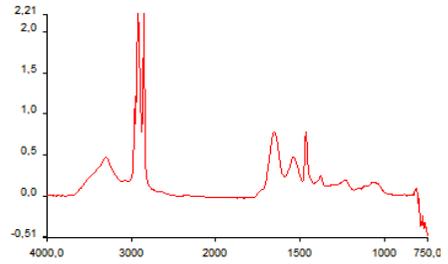


Image spectrale IR brute univariée



Chaque pixel (6.25x6.25 μm²) contient un spectre entier

3.2) Protocole : prétraitement par EMSC

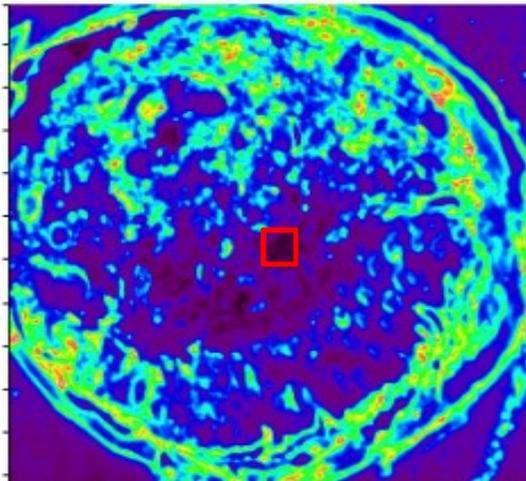


Image infrarouge brute
(absorbance totale)

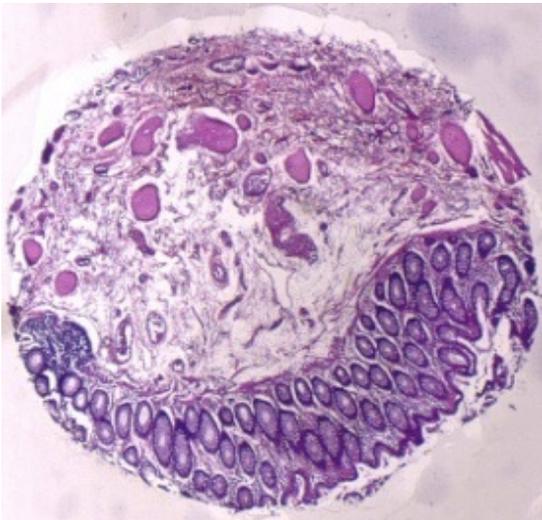
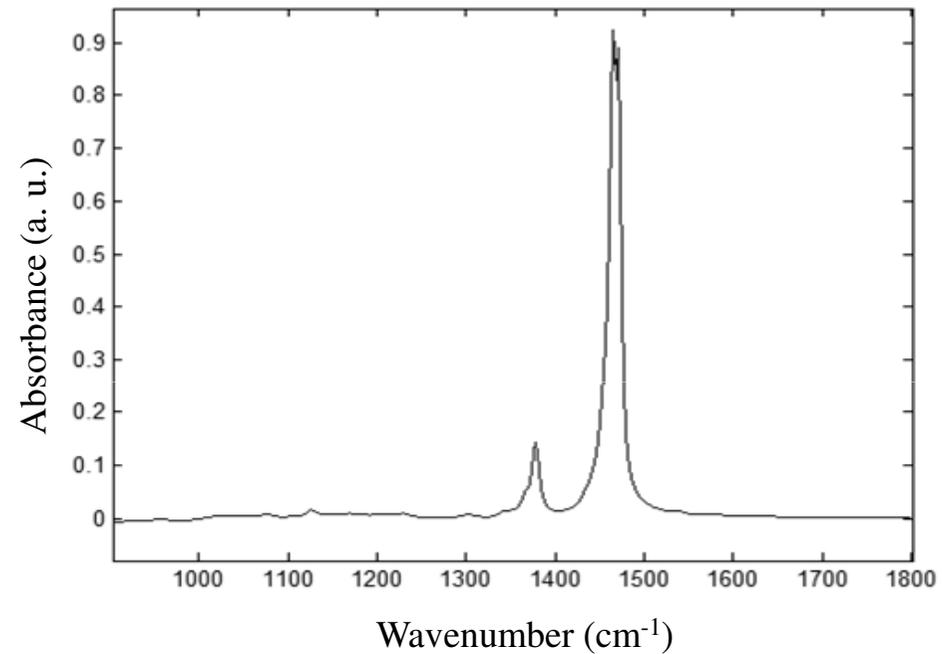


Image colorée par HPS
(5X magnification)

Paraffine



*Les spectres IR représentés sont les spectres cumulés moyens

3.2) Protocole : prétraitement par EMSC

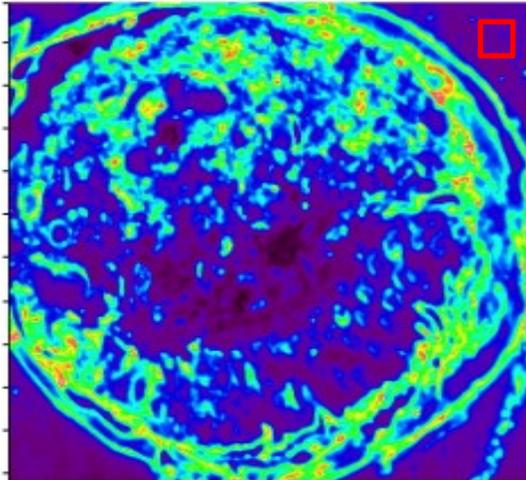


Image infrarouge brute
(absorbance totale)

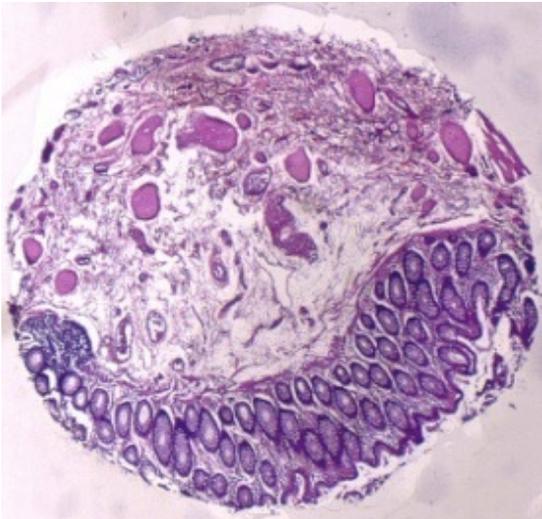
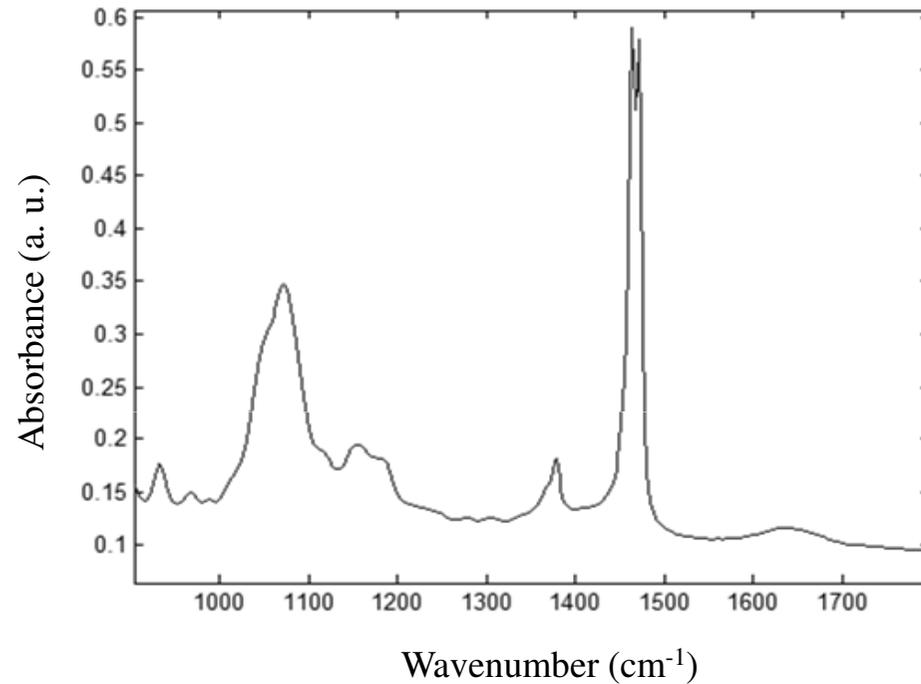


Image colorée par HPS
(5X magnification)

Agarose + Paraffine



*Les spectres IR représentés sont les spectres cumulés moyens

3.2) Protocole : prétraitement par EMSC

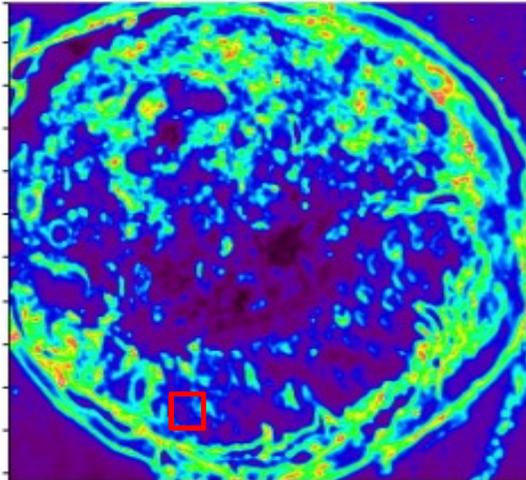


Image infrarouge brute
(absorbance totale)

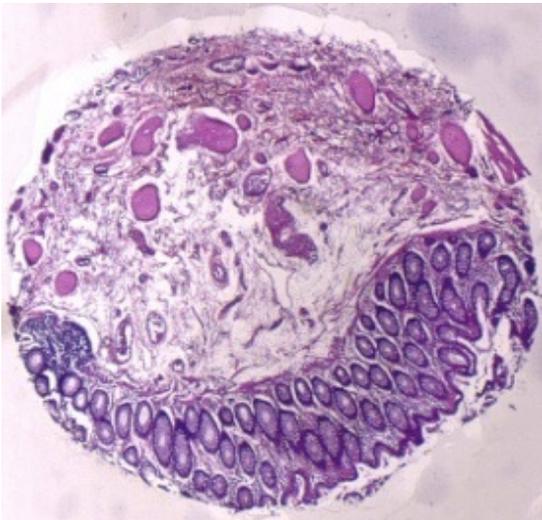
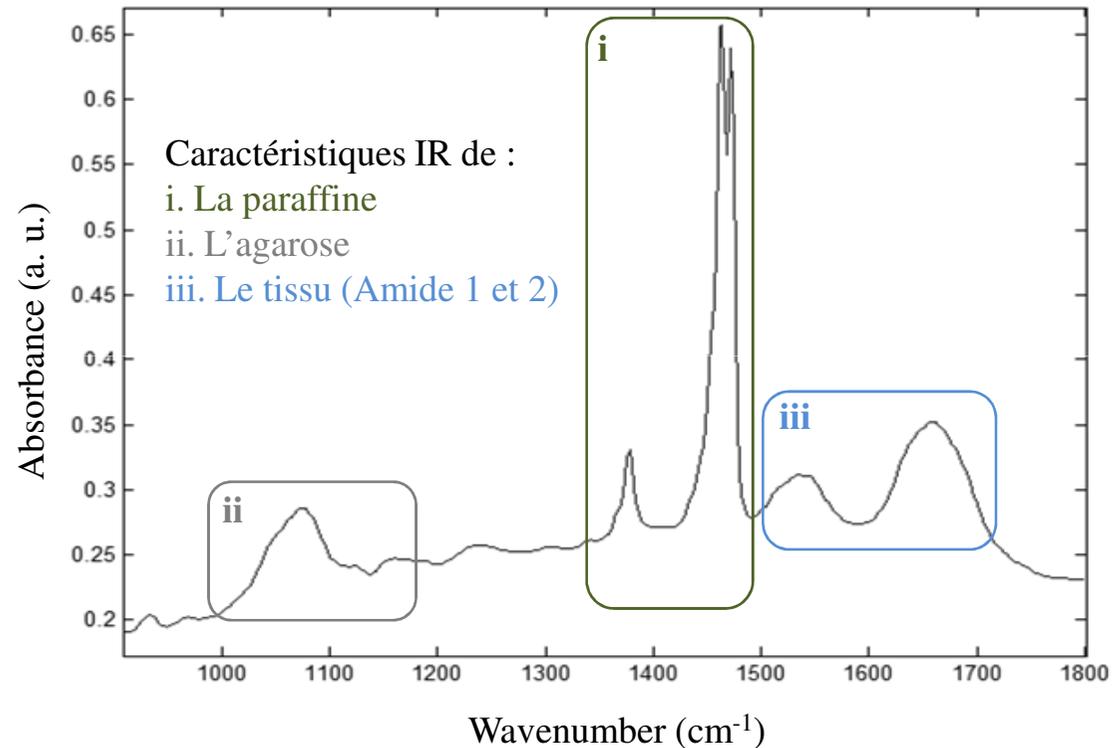


Image colorée par HPS
(5X magnification)

Tissue + Agarose + Paraffine

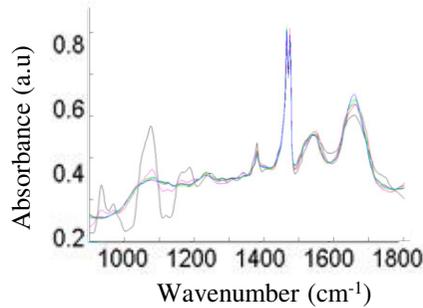


La variabilité des signaux de la paraffine et de l'agarose influence l'exploitation de l'information tissulaire

*Les spectres IR représentés sont les spectres cumulés moyens

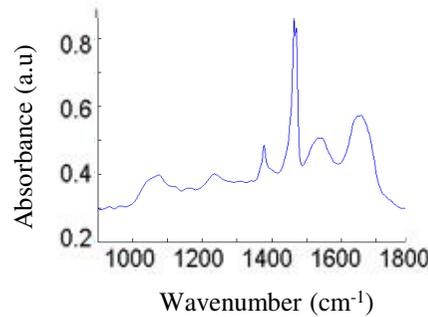
3.2) Protocole : prétraitement par EMSC

$$S_i = a_i \hat{S} + b_i I + c_i P + e_i$$



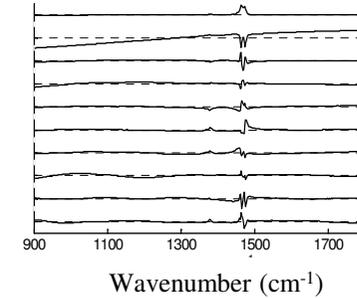
$$= a_i \hat{S}$$

S_i = spectre brut du jeu de données

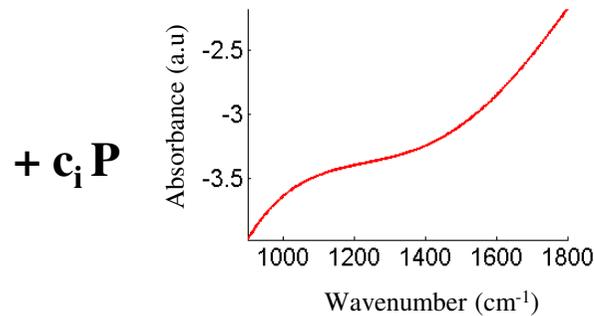


$$+ b_i I$$

\hat{S} = spectre moyen



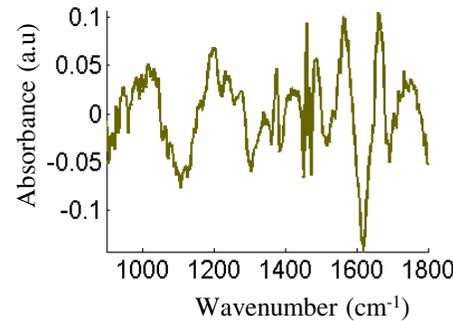
I = matrice d'interférence
(Paraffine / Agarose)



$$+ c_i P$$

$$+ e_i$$

P = polynôme pour l'effet de diffusion de la lumière

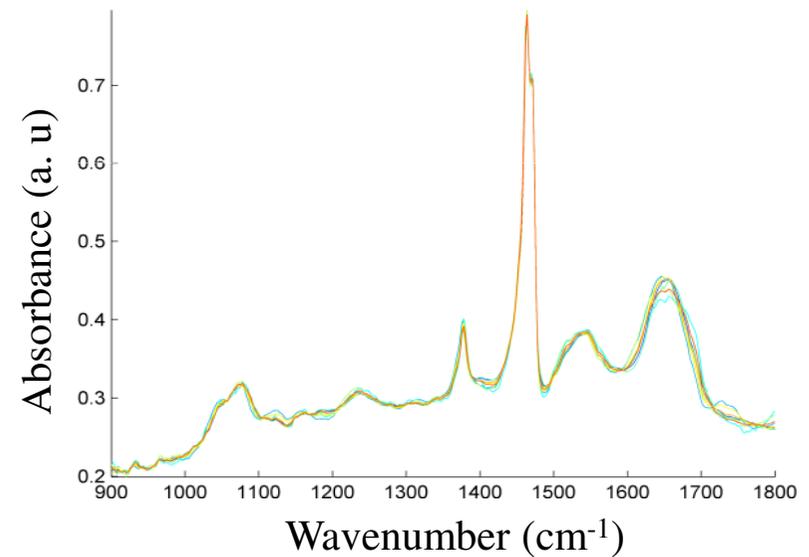
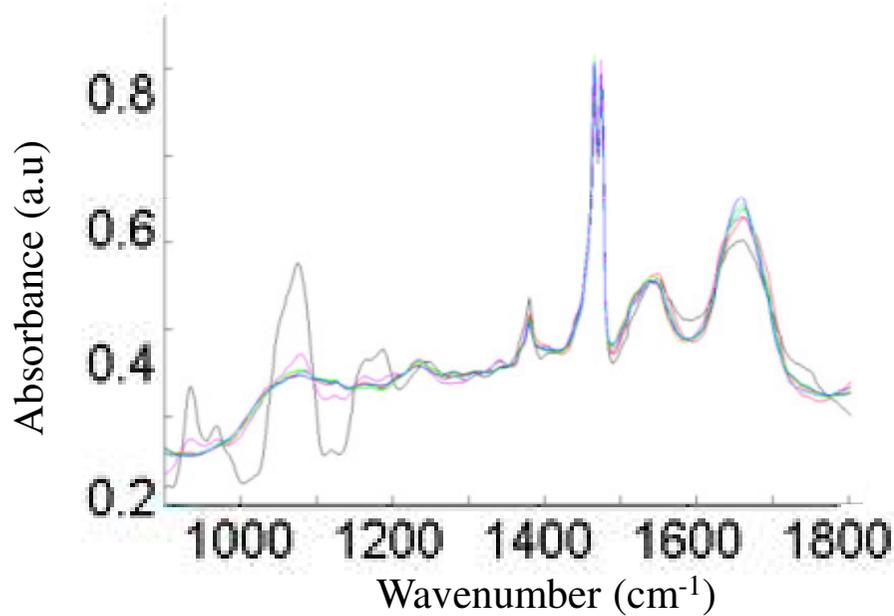


e_i = erreur de modélisation

$$S_{icorr} = \hat{S} + \frac{e_i}{a_i}$$

3.2) Protocole : prétraitement par EMSC

$$S_{icorr} = \hat{s} + \frac{e_i}{a_i}$$



Déparaffinage/neutralisation mathématique des contributions de la paraffine et de l'agarose

3.3) Protocole : classification non-supervisée par K-means

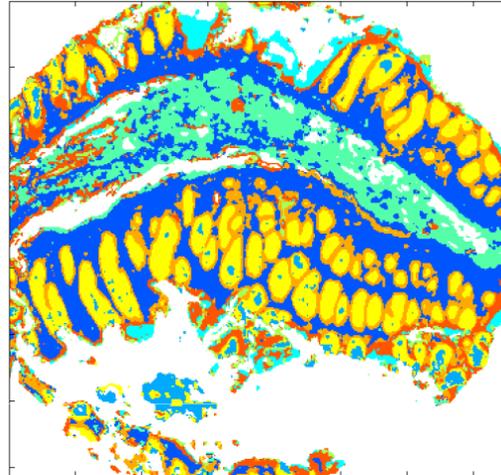
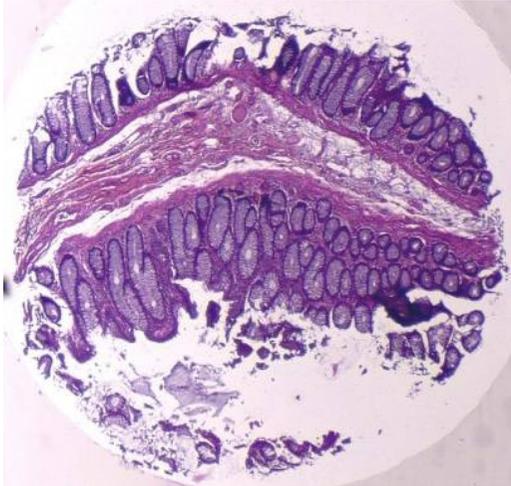
Pseudo-couleurs aléatoires

Tissu coloré par HPS
(grossissement 5X)

Image IR classée par K-means

Attribution histologique

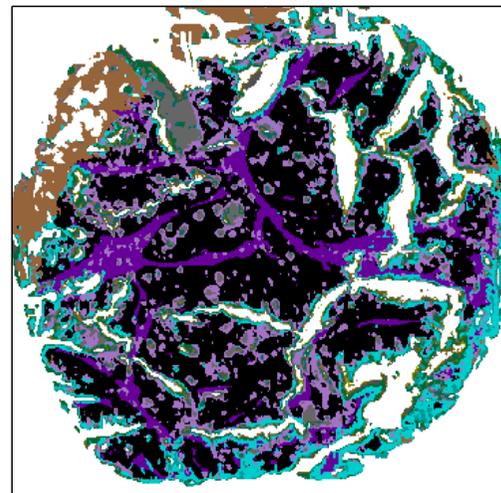
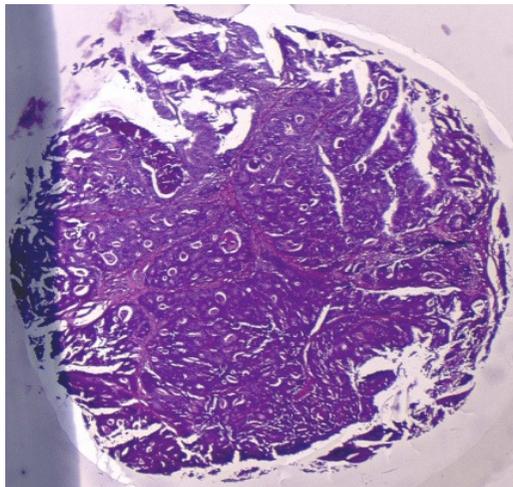
Tissu non-tumoral



Annotation par un pathologiste expert

- Cryptes (partie interne)
- Cryptes (partie externe)
- Lamina propria
- Sous-muqueuse

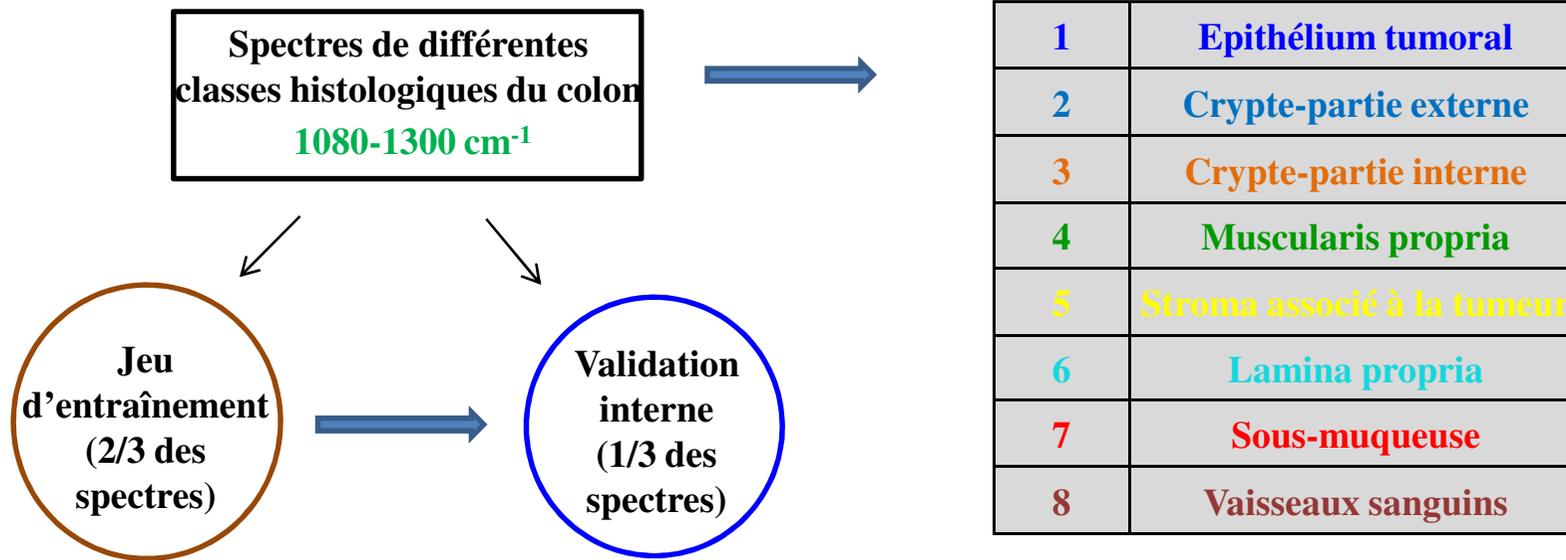
Tissu tumoral



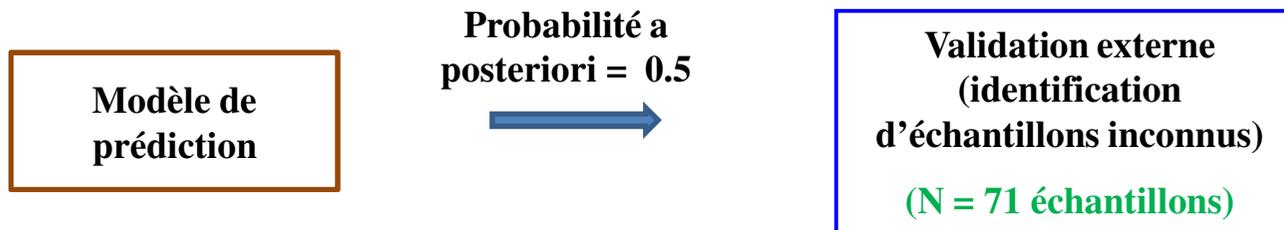
- Tumeur
- Stroma associé à la tumeur
- Autres - Non attribués

3.4) Protocole : construction d'un modèle de prédiction par analyse discriminante linéaire

(N = 9 échantillons; 86 802 spectres)



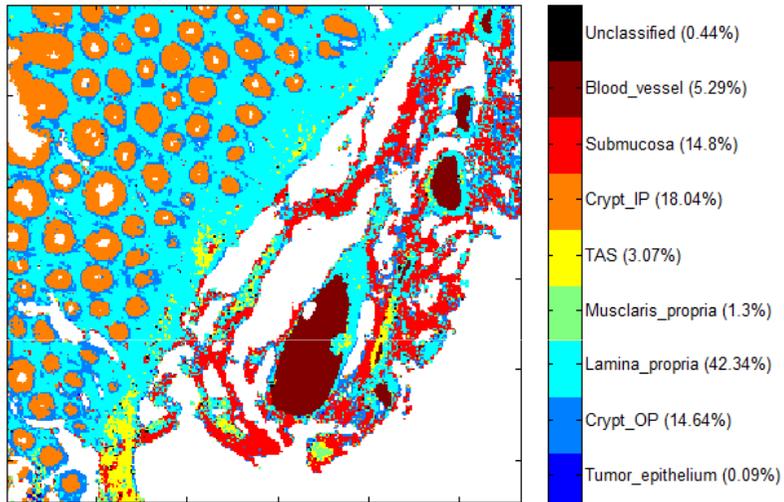
Sensibilité moyenne de 89.38 %



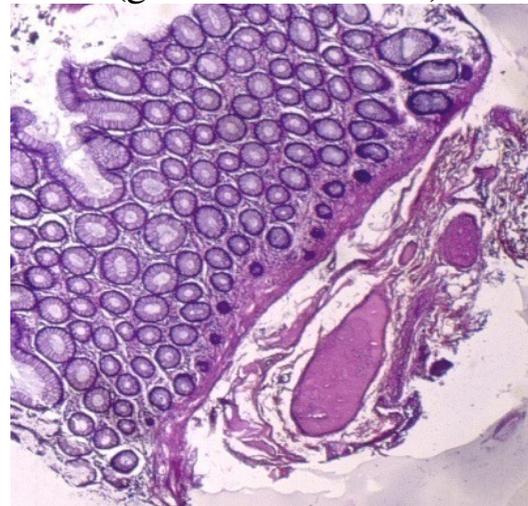
3.5) Protocole : validation externe : identification de tissus coliques inconnus par histopathologie spectrale

(Nb. de spectres testés: 3 620 287)

Images prédites par IR-LDA

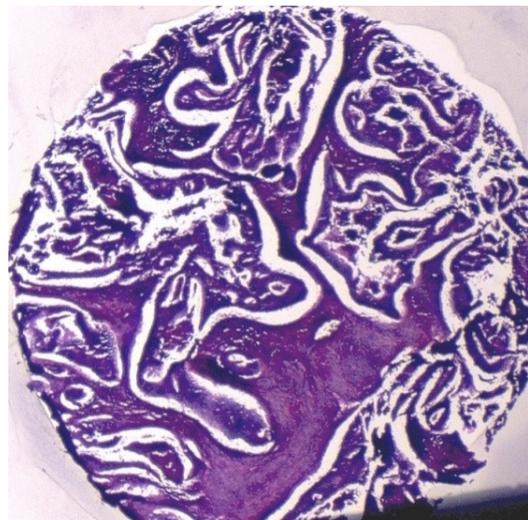
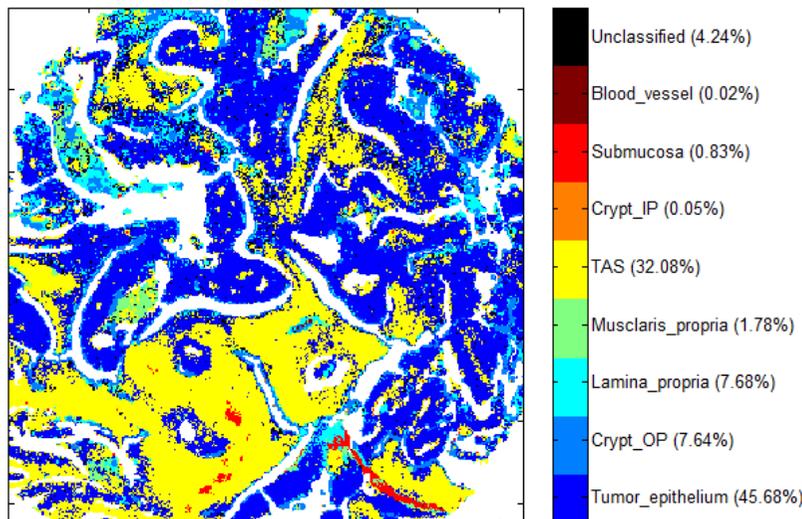


Tissus coliques colorés par HPS (grossissement 5X)



Diagnostic final

Non tumoral



Adénocarcinome

3.5) Protocole : validation externe : écueil du modèle de prédiction

Influence de l'inflammation

LDA_1080 - 1300 cm ⁻¹	
Sensibilité	100 %
Spécificité	58 %
Accuracy	83 %

12 échantillons faux-positifs

L'inflammation, qui a été identifiée comme de la tumeur, influence directement la spécificité du modèle

Validation externe

	Tumoral	Non-tumoral	Total
Test positif	42 (TP)	12 (FP)	54
Test négatif	00 (FN)	17 (TN)	17
Total	42	29	71

4) Identification du budding tumoral

Budding :

- Petits regroupements de cellules tumorales isolées dans le stroma – Tumeurs invasives

Tissu colique coloré par HPS
(grossissement 5X)

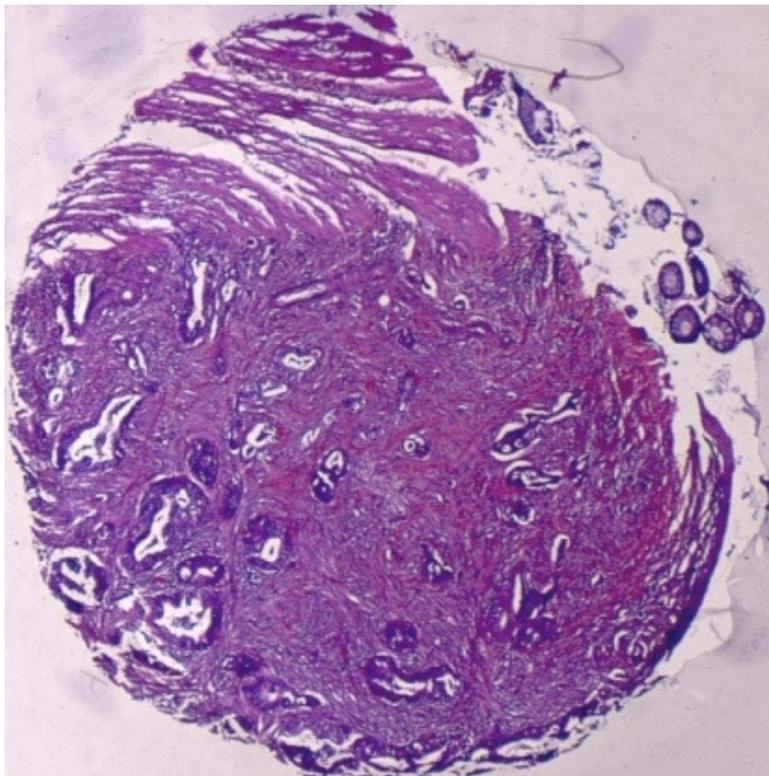
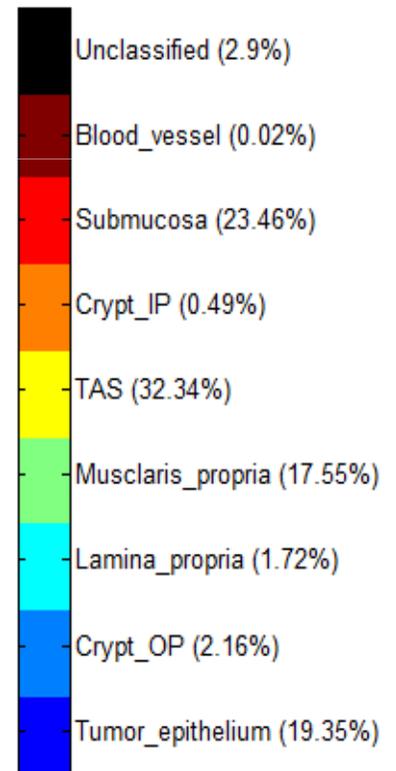
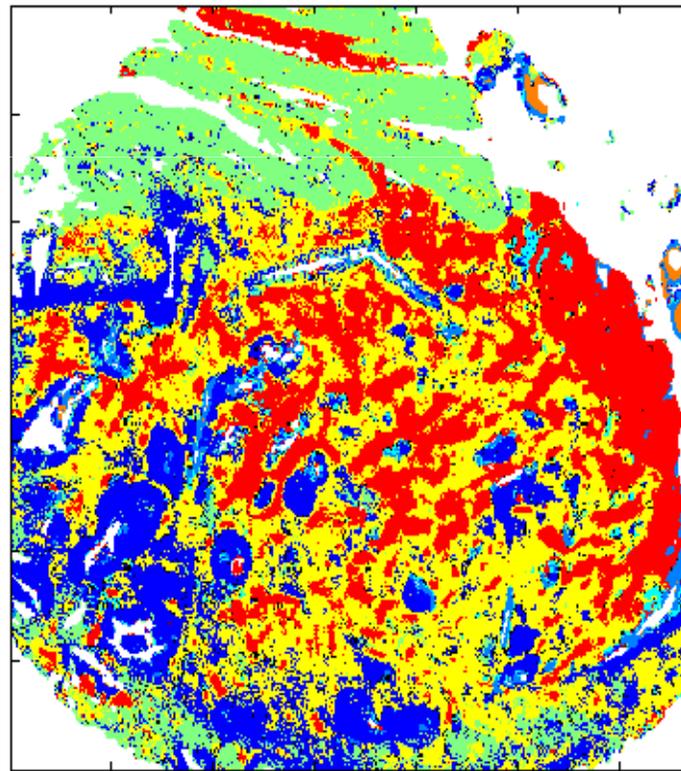


Image prédite par IR-LDA



4) Identification du budding tumoral

Tissu colique coloré par HPS
(grossissement 5X)

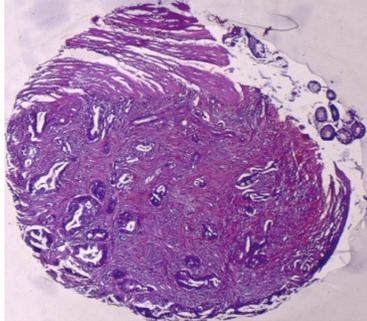


Image prédite par IR-LDA

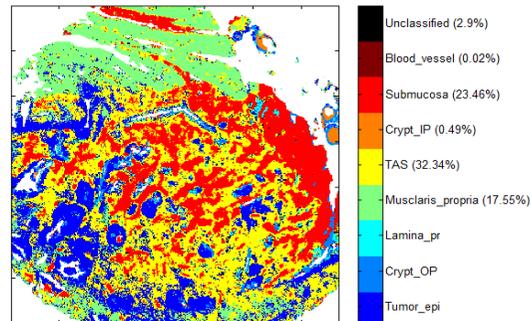
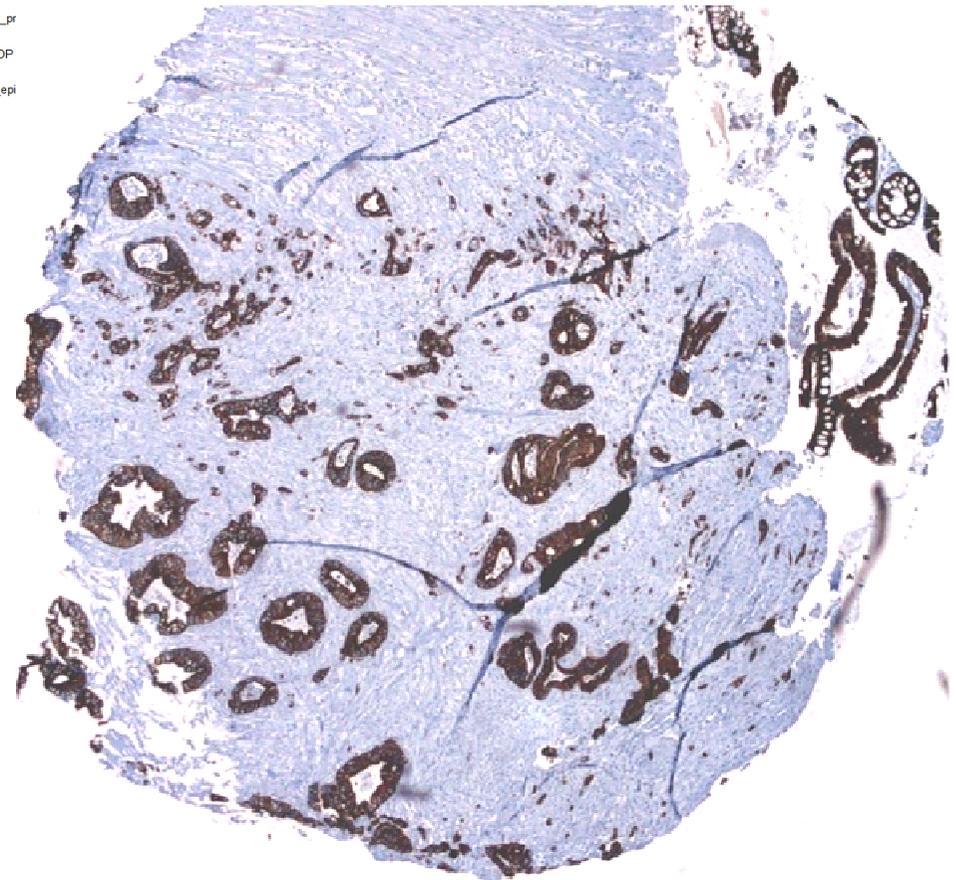


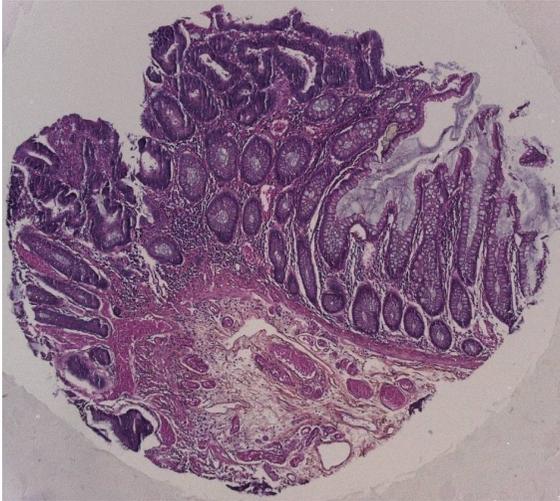
Image IHC (KL1)
(grossissement 5X)



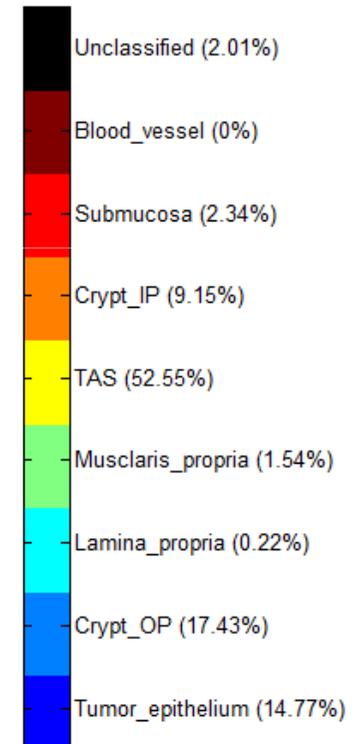
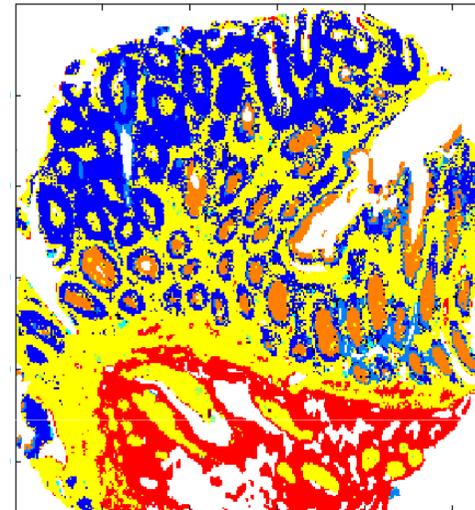
5) Identification de caractéristiques adénomateuses

Diagnostic précoce (analyse préliminaire)

Tissus colorés par HPS
(dysplasie de haut grade)



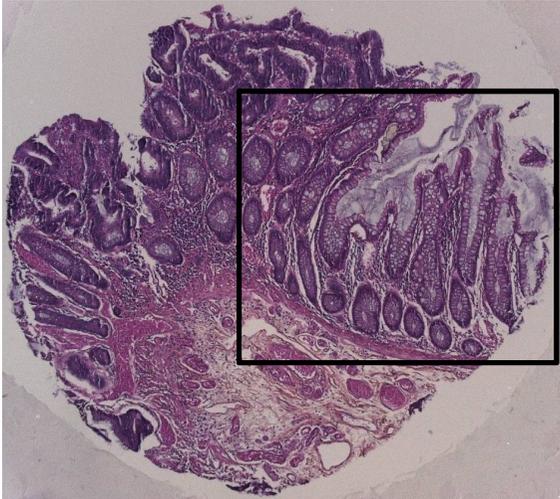
Images prédites par IR-LDA



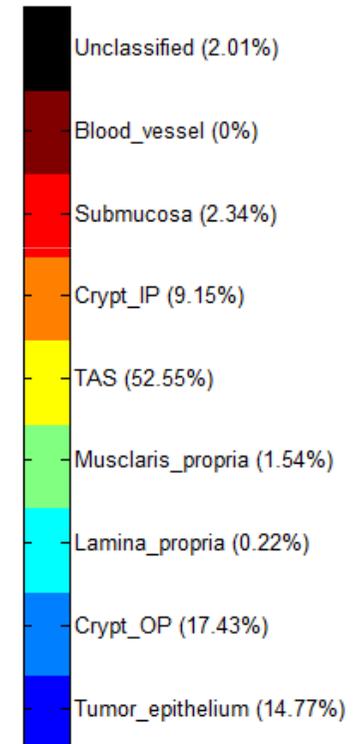
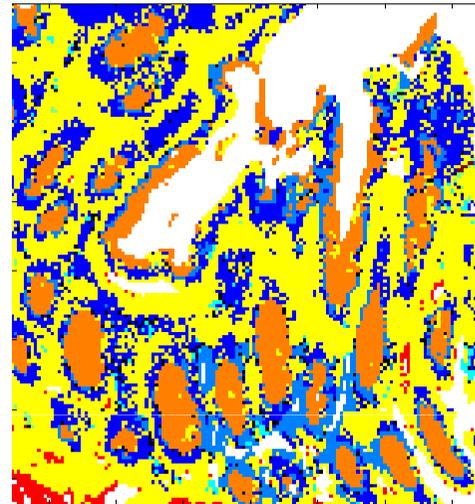
5) Identification de caractéristiques adénomateuses

Diagnostic précoce (analyse préliminaire)

Tissus colorés par HPS
(dysplasie de haut grade)



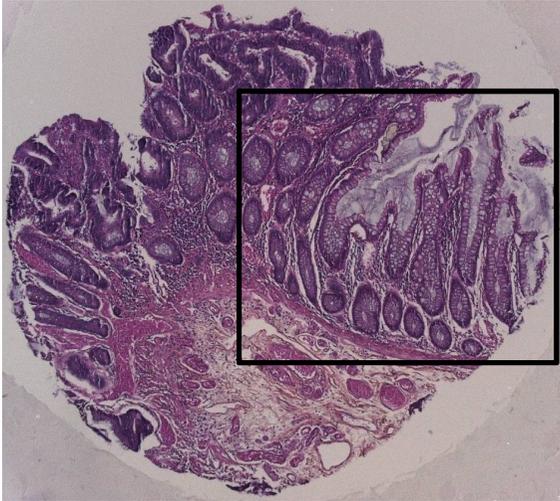
Images prédites par IR-LDA



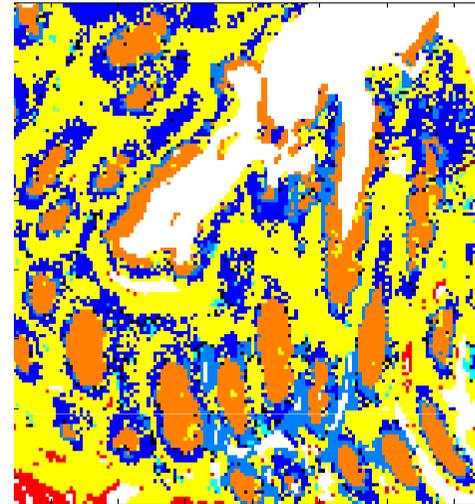
5) Identification de caractéristiques adénomateuses

Diagnostic précoce (analyse préliminaire)

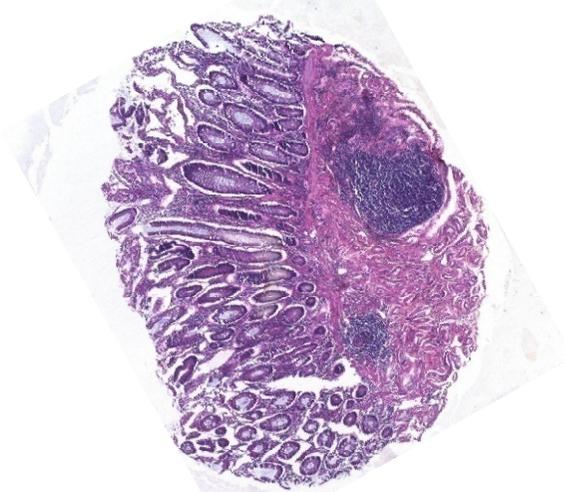
Tissus colorés par HPS
(dysplasie de haut grade)



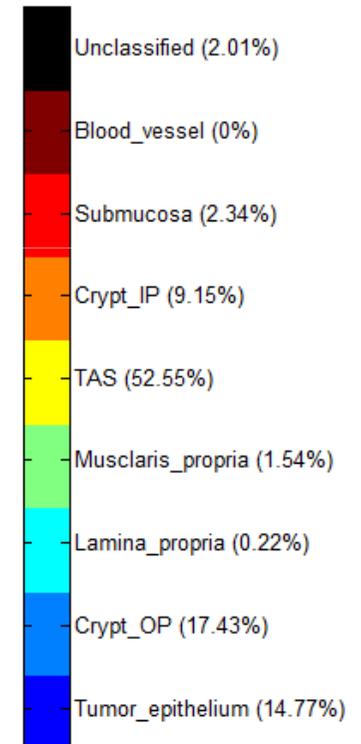
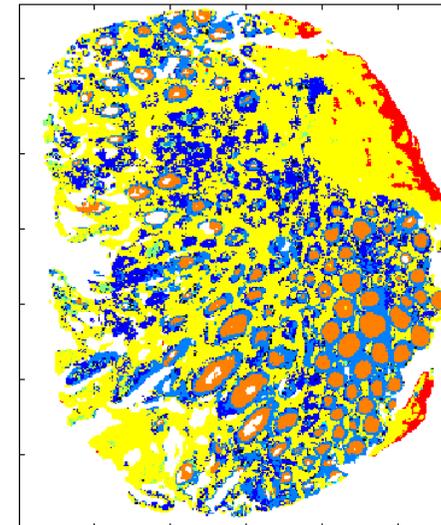
Images prédites par IR-LDA



(dysplasie de bas grade)



*grossissement 5X



Conclusions générales

- **Modèle de prédiction** : identification sans marquage de tissus inconnus (100 % de sensibilité); malignité et ses caractéristiques associés (budding tumoral, caractéristiques adénomateuses)
- Une nouvelle approche d'histopathologie spectrale infrarouge des tissus coliques, complémentaire à l'histologie conventionnelle : biomoléculaire, objective et diagnostic sans marquage
- **Application clinique potentielle** : diagnostic automatique de la malignité et de ses caractéristiques associées

Perspectives

- Amélioration de la spécificité du modèle de prédiction en ajoutant une classe de l'inflammation dans le modèle
- Développement de marqueurs pour d'autres cas rencontrés en clinique :
 - Métastases
 - Autres grades tumoraux
 - Hétérogénéité tumorale