

# 機械学習を用いた病理画像診断支援システム

埼玉医科大学  
保健医療学部 臨床工学科

2021年12月11日

小林直樹

# 医用画像解析における機械学習

---

## ■ 医用画像解析の研究

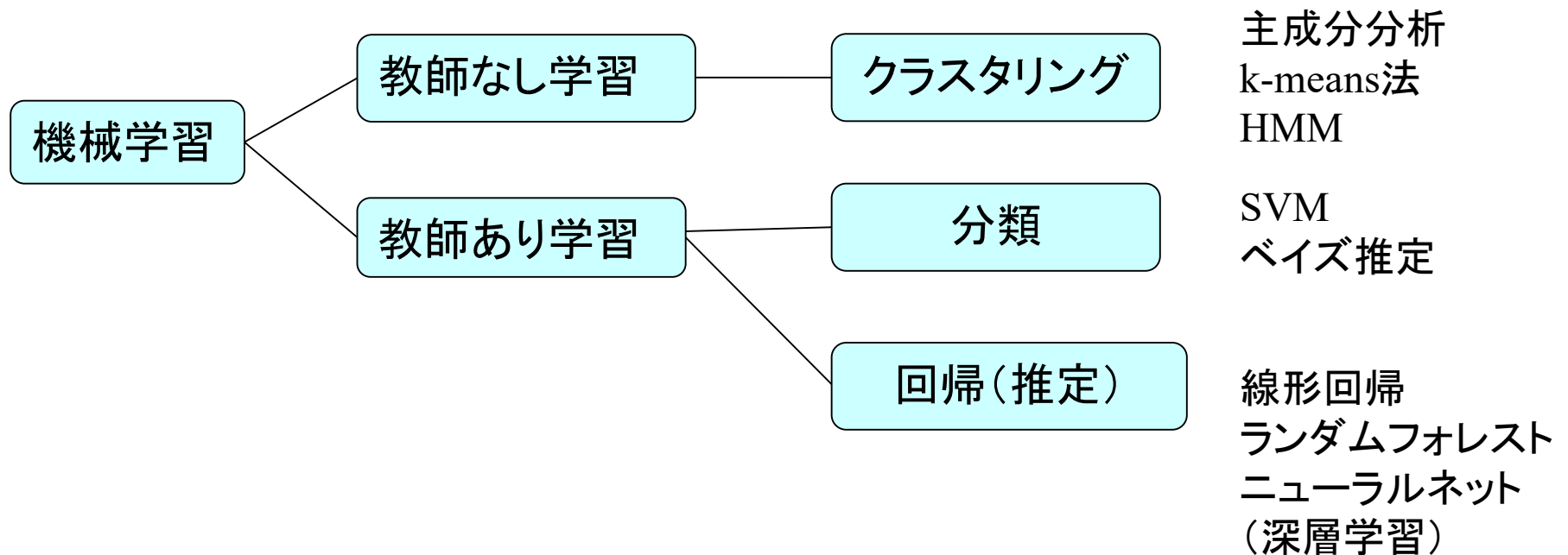
### 🌐 ハイパースペクトルを用いた定量的病理診断支援

- HE病理標本からのがん陽性核推定
- HE病理標本による線維組織の分類

# 機械学習

解析結果を求める ⇒ 数式モデルからアルゴリズム導いて解析

⇒ 得られているデータをもとにアルゴリズムを生成して解析。(機械学習)



# 機械学習の分類

- AI(人工知能)    ルールベース機械学習(決定木)  
                         統計的機械学習(SVM、ベイズ推定)  
                         ニューラルネット(パーセプトロン、深層学習)
- 古典的機械学習    ⇒ 特徴量を求め、特徴空間上で学習
- NN学習(深層学習) ⇒ 入力データで学習(特徴量は自動生成)

どの学習方法が適しているか？

# 医用画像解析における機械学習

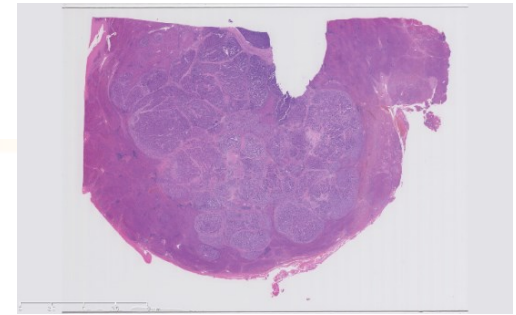
---

## ■ 医用画像解析の研究

### 🌐 ハイパースペクトルを用いた定量的病理診断支援

- HE病理標本からのがん陽性核推定
- HE病理標本による線維組織の分類

# 病理画像診断とは？



- 染色組織像を顕微鏡下で観察
- 病変の有無や種類、良悪性、進行度などを判断
- 確定診断、最終診断 → 治療方針の決定

## [病理診断のプロセス]

検体組織採取



組織標本  
作成



病理診断

パラフィン処理



薄切



染色



スライド化

### ■ 染色法

■ 一般染色：

HE (Hematoxylin & Eosin)

■ 特殊染色：

マツソントライクロームetc.

■ 免疫組織化学染色

Ki-67

- 生検組織診断 内視鏡組織切除、皮膚組織切除、針生検
- 手術での摘出臓器・組織の診断
- 術中迅速診断
- 凍結標本、病理解剖

# デジタル病理画像診断への期待

## 現在の病理診断

- 顕微鏡による形態学的目視観察が基本  
(要 経験的判断力)
- 半定量(N/C比、線維化組織の占有率、etc.)、  
手作業計測診断
- 他の臨床病理学的指標、予後・臨床経過との比較

- 定量性、再現性に欠ける
- 施設内の診断が中心  
専門外領域の診断が必要

## デジタル病理画像診断

### デジタル画像による高機能化

- 目視診断の迅速化  
(バーチャルスライド)
  - 画像処理による**定量的診断**
  - デジタル染色**
  - 病理画像転送
  - データベース化
  - 分子細胞学, PACS等と対応
- ▶病理診断の質向上・均てん化
  - ▶遠隔病理診断、専門別診断
  - ▶形態病理と分子細胞学・画像診断(MRI、CT等)を統合

# 医用画像解析における機械学習

---

## ■ 医用画像解析の研究

### 🌐 ハイパースペクトルを用いた定量的病理診断支援

- HE病理標本からのがん陽性核推定
- HE病理標本による線維組織の分類

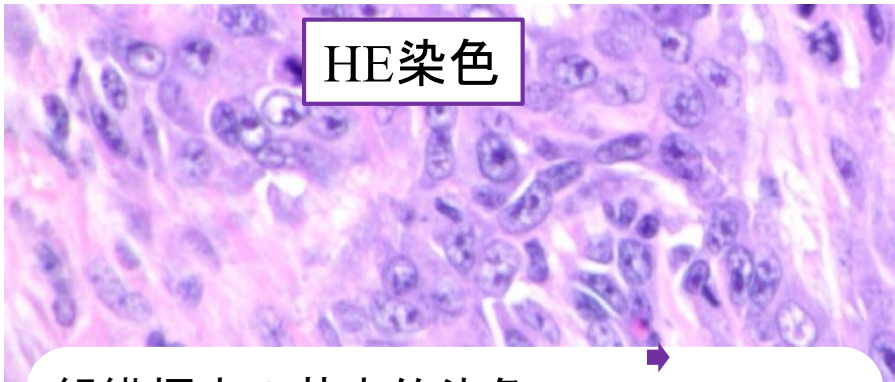


# 背景・目的

■ 病理診断： 手術切片、生検など組織 ⇒ がんの確定診断

- ・組織をHE染色し、顕微鏡で形態観察・診断
- ・必要に応じて特殊染色（Ki-67等）

## HE染色標本とKi-67標本



HE染色

- ・組織標本の基本的染色。
- ・組織構造を把握でき、観察がしやすい。
- ・ヘマトキシリン：核, リボゾーム 青紫色
- ・エオジン：細胞質, 線維 ピンク色



Ki-67

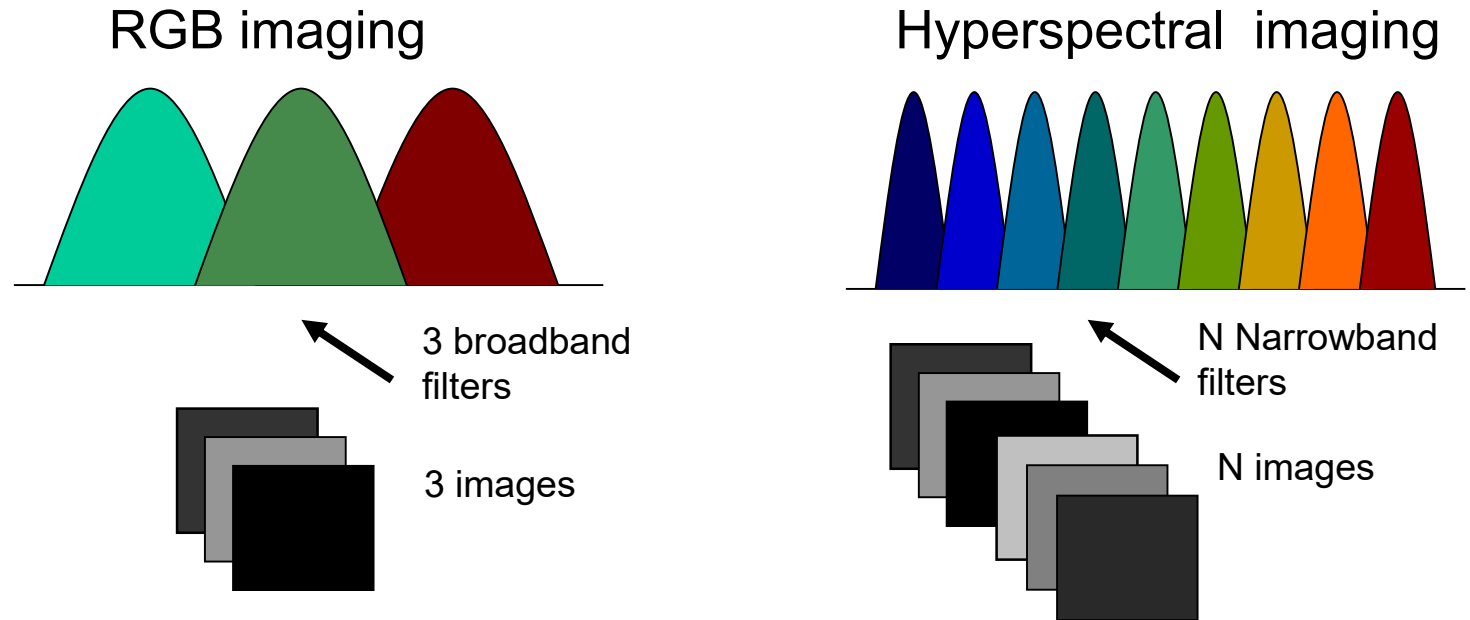
- ・細胞分裂するときが発生するタンパク質。
- ・細胞分裂が活発な細胞の核が茶褐色に染色される。

Ki-67は染色に時間を必要  
染色に加熱処理⇒標本への影響大きい。

ハイパースペクトルHE染色病理画像のみでがん陽性核推定

# ハイパースペクトルカメラ

- 多くの狭帯域フィルターを利用し、詳細な色情報を含む多バンド画像を撮影可能

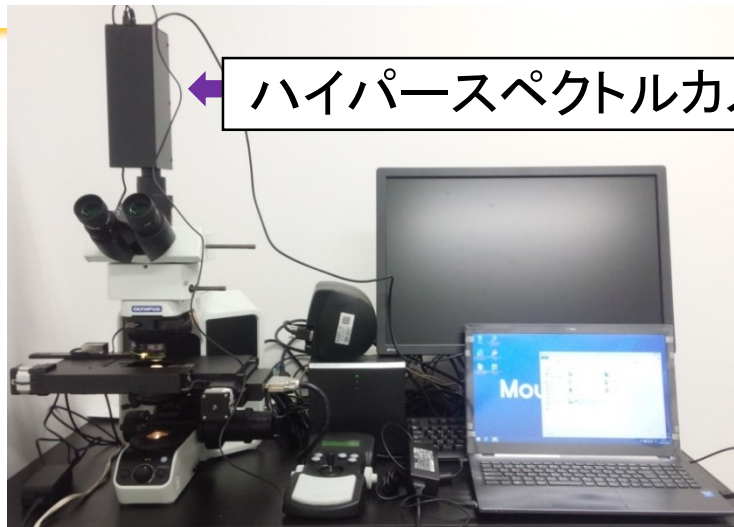


H&E画像の膠原線維と弾性線維にも、**微妙な色の差**があるのでは？



**ハイパースペクトル画像**で見分けられるか検討

# ハイパースペクトル画像 (HSI)

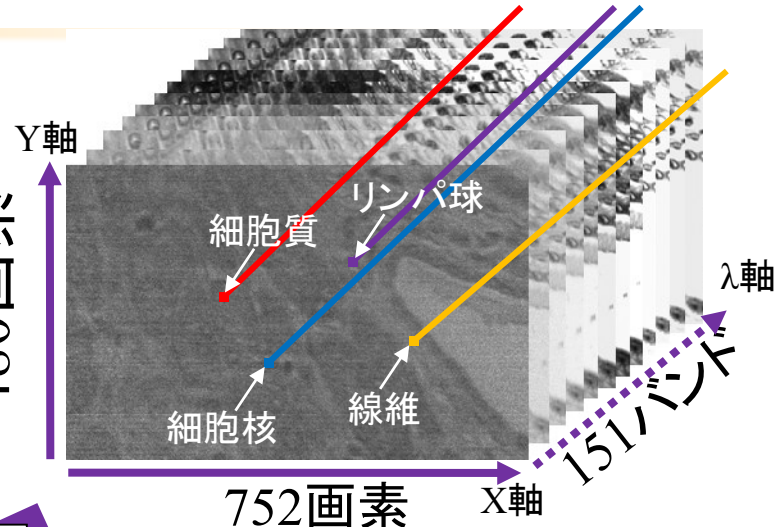


ハイパースペクトルカメラ

Eva-japan社製 NH-3

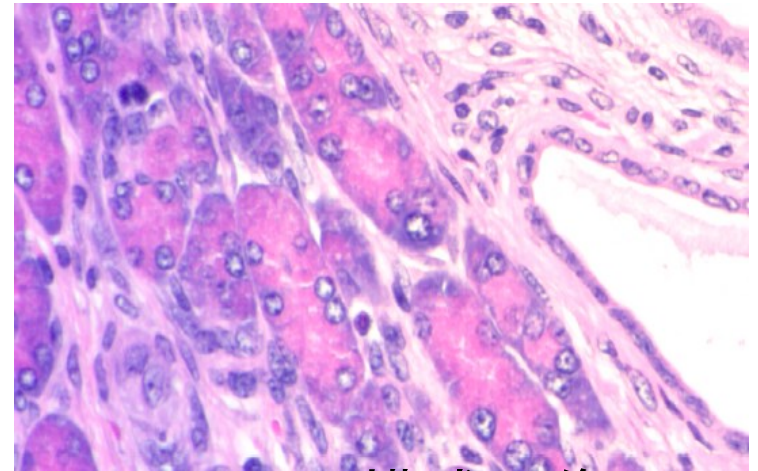
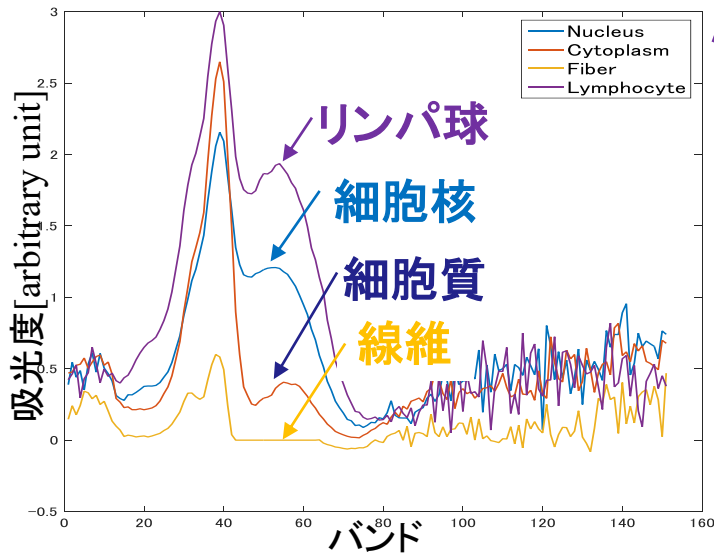


480画素



各組織の  
スペクトル

RGB再構成

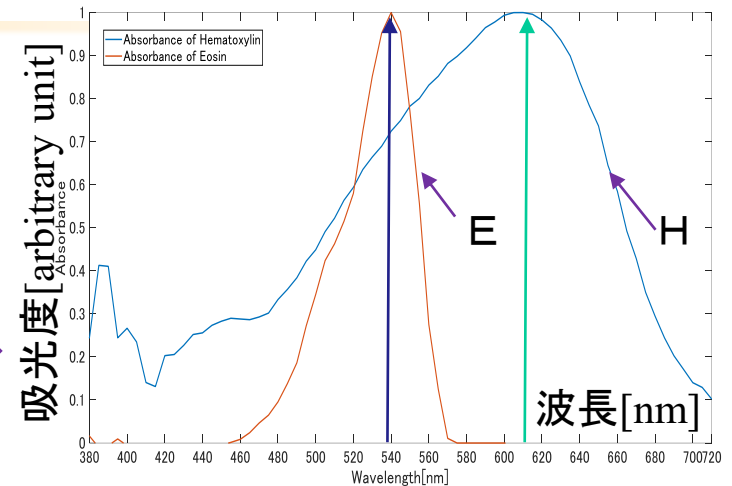
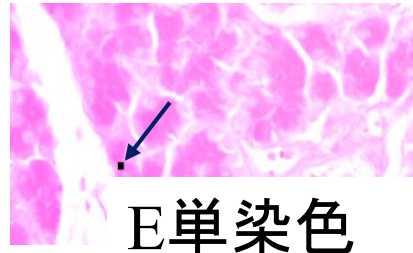
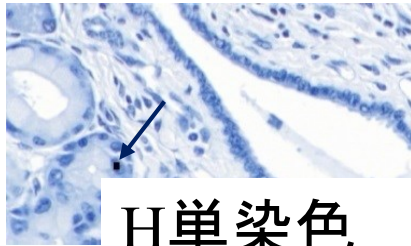


RGB再構成画像

RGB画像では捉えられない、膨大なスペクトルデータを一度の撮影で取得できる

# 特徴量1: HSI色素による量推定

HE染色標本は基本的にはHとEの2色で染色されており、その色の混ざり具合でスペクトルが決定すると仮定している



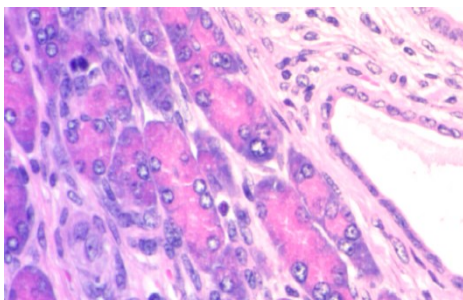
エオジンは540nm, ヘマトキシリンは610nmをピークに吸光帯がある

ランバートベールの法則から、HE染色病理標本は(4)式によって定義される

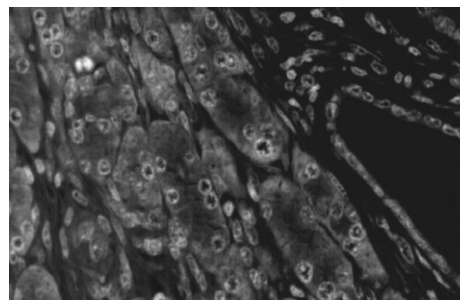
$$a = XC + e \quad (4) \quad a: \text{吸光度} \quad X: \text{分光吸収係数} \quad C: \text{色素量} \quad e: \text{その他の成分}$$

$$X^+ = (X^T X)^{-1} X^T e$$

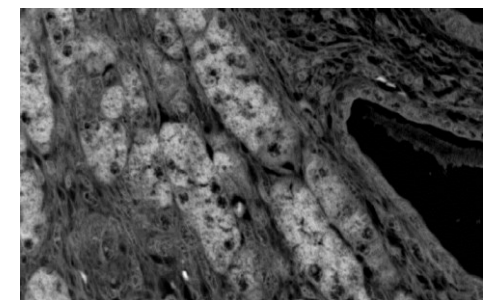
(5)式の推定行列 $X^+$ を吸光度ベクトル $a$ に、乗算することで推定色素量 $\hat{c}$ を得る



RGB再構成画像



ヘマトキシリン色素量



エオジン色素量



# 特徴量2: 色素量推定による差分情報

## 色素量推定

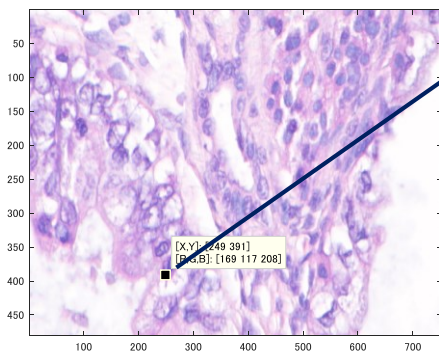
- HE染色標本はHとEの混ざり具合でスペクトル波形が決定すると仮定している

$$a = XC + e \quad (4) \quad X^+ = (X^T X)^{-1} X^T \quad (5) \quad \begin{array}{l} a: \text{吸光度} \quad X: \text{分光吸収係数} \\ C: \text{色素量} \quad e: \text{その他の成分} \end{array}$$
$$\hat{c} = X^+ a \quad (6) \quad \hat{c}: \text{推定色素量}$$

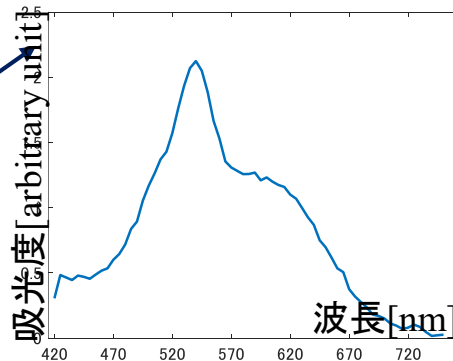
推定色素量から吸光度を再推定:  $a' = X\hat{c} \quad (7)$

$$e = a - (X\hat{c}) \quad (8)$$

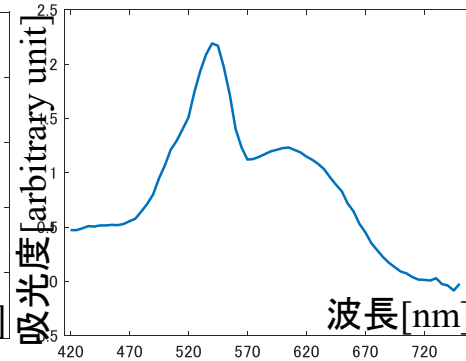
色素では表現されない部分に着目:  $e = a - a'$



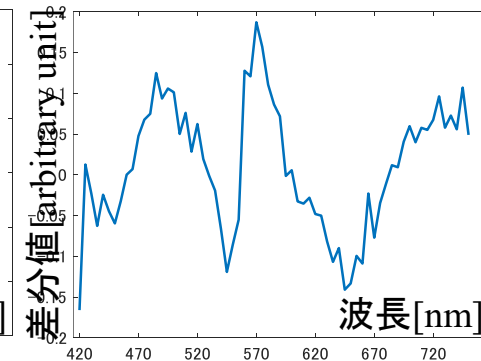
原画像



吸光度スペクトル



推定した色素量によって  
再構成したスペクトル

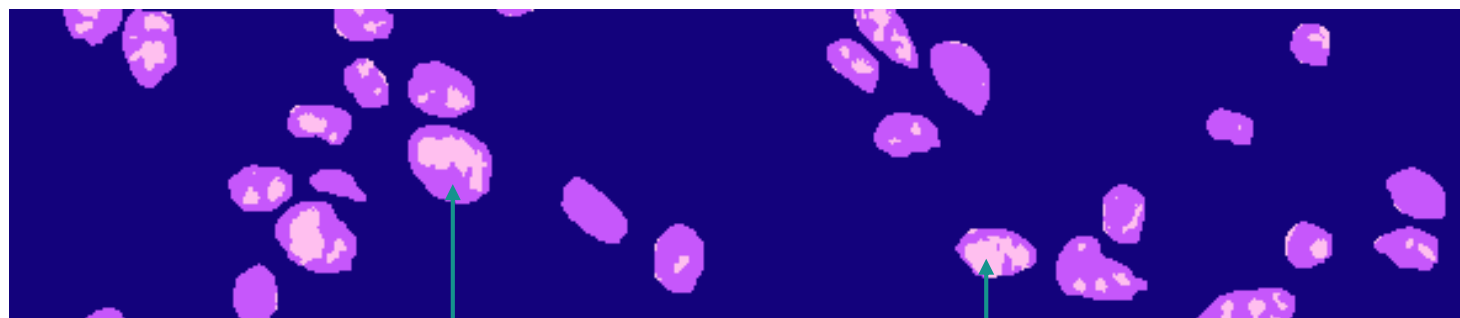
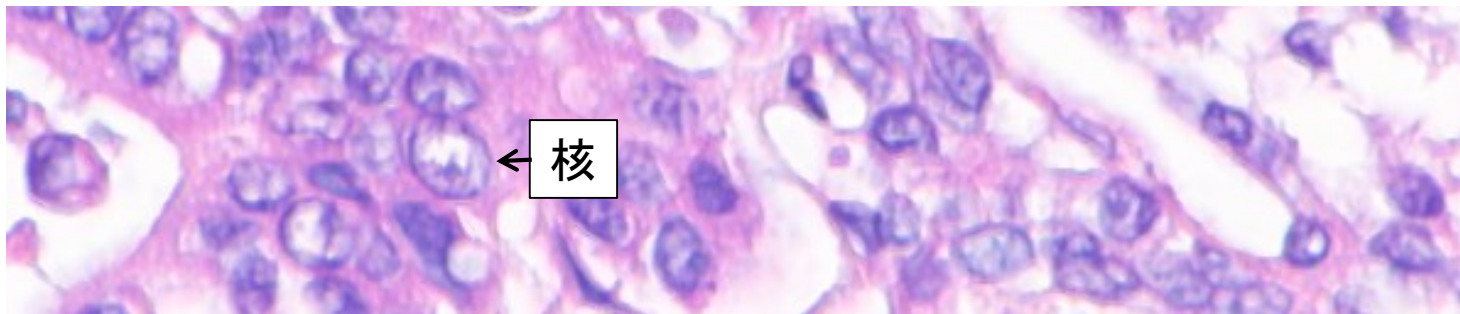


差分情報

差分情報を使用することで染色の影響も受けにくくなる

# 特徴量3 核の濃い,薄い の分類

- 細胞核には、濃い画素と薄い画素が存在する。



濃いと分類された画素

薄いと分類された画素

- そのまま機械学習に入力するとがん・正常に加えて染色度合いを含めた分類が必要となり複雑な計算が必要となる。

核の画素ごとに色の濃い画素と薄い画素に分類してから学習

# がん陽性核判別実験

## 学習条件

- ・学習画素数:がん、正常でランダムに各1000画素 合計2000画素
- ・学習:Random Forest 決定木:500 深さ:134band → 11, 67band → 8, RGB → 1
- ・12枚の画像を用いた被験者ごとの cross validationによる評価

## 学習のパターン

### 得られた波長情報に基づく分類

- RGB,HSI,EH比率が得られており、それらの評価とEH比率に算出した差分情報を加えたときに、結果がどうなるか実験を行った。
  - ・RGB・HSI・EH比率
  - ・EH比率+差分情報
- 4パターン ➤ Random Forest

### 核を濃い画素、薄い画素に分類してから識別

- 核の濃い場所、薄い場所で分けてから識別を行った方が、結果が良くなるか確認するために、実験を行った。
  - ・E 透過率による閾値:0.2
  - ・C-HSI・C-EH比率
  - ・HAPSS
- 3パターン
- 濃い画素 ➤ Random Forest  
薄い画素 ➤ Random Forest

# 実験環境

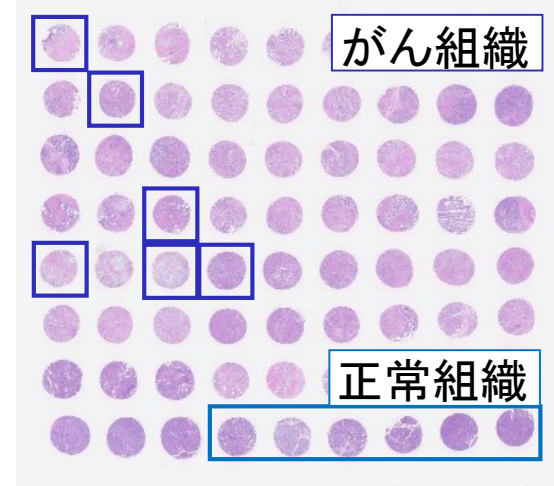
## 標本撮影

- ・Tissue Micro Array(TMA) 標本 US Biomex社のPancreatic cancer (PA721)
- ・正常組織6例とランダムに選択したがん組織6例 (Duct adenocarcinomaのグレード2)

## Tissue Micro Array(TMA)

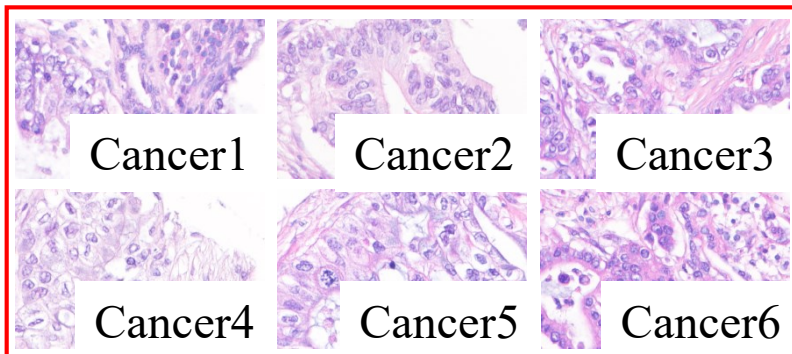
- ・異なる患者の組織切片を1枚のスライドガラスに整列したもの
- ・病理医によりがん・正常の診断結果とがんのグレードが付与されている

## ■ Tissue Micro Array

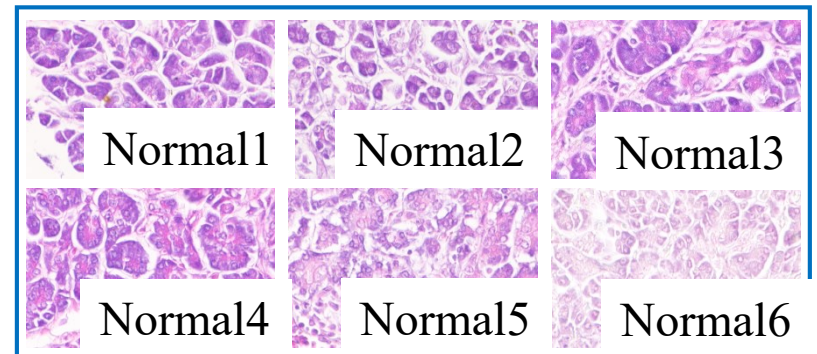


## 使用画像

### ■ がん組織



### ■ 正常組織

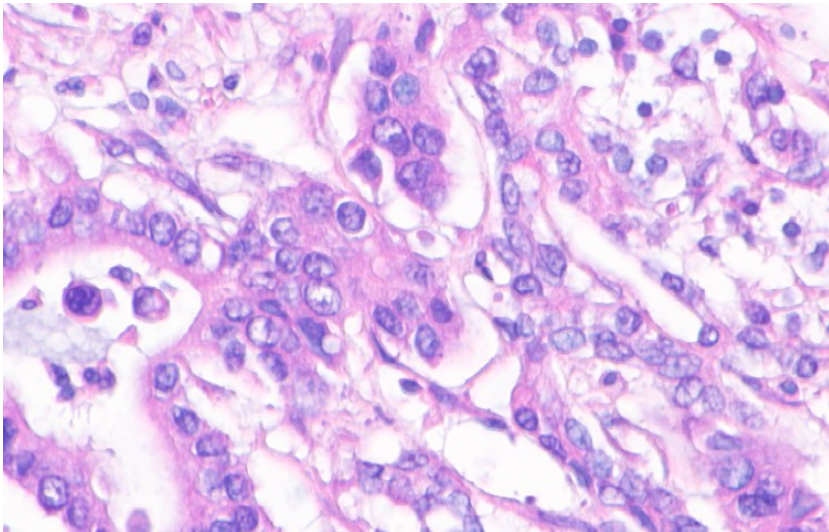




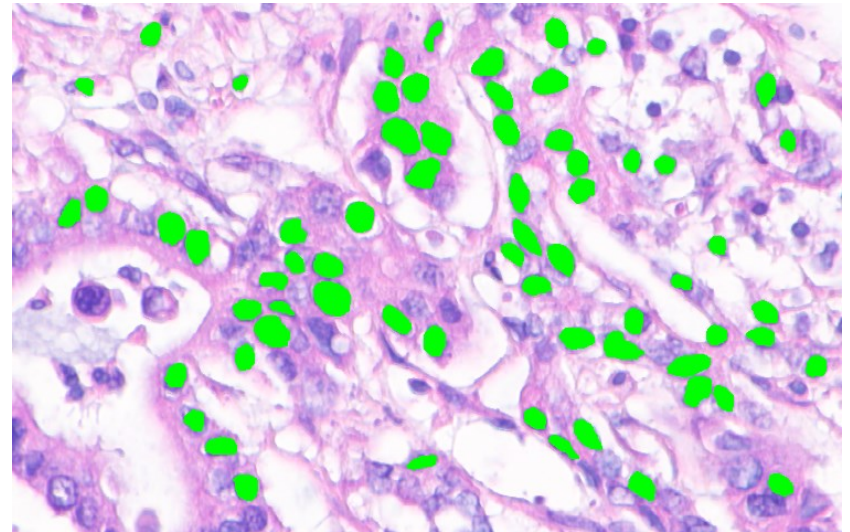
# 手動での細胞核抽出（正解データ作成）

- がんになると組織の増殖能が増すため、細胞核に変化が起こる。
  - ➡ がんと正常の細胞核の違いに着目し、細胞核に限定した解析を行う。
- 細胞核は手動で抽出した。

## ■ RGB画像



## ■ 核を抽出した画像（緑色の画素）



# 分類結果(平均精度)

## 得られた波長情報に基づく分類

- ・感度: がんをがんと分類した確率
- ・特異度: 正常を正常と分類した確率

	RGB(%)	HSI(%)	EH比率(%)	EH比率+差分(%)
精度				
感度	82.8	88.4	87.8	90.2
特異度	74.2	74.0	88.3	89.2

## 核の濃い画素、薄い画素の分類あり(E透過率による閾値:0.2)

	C-HSI(%)			C-EH比率(%)			HAPSS(%)		
	濃	薄	全	濃	薄	全	濃	薄	全
精度	91.0	85.2	87.6	90.2	86.8	88.2	92.1	90.0	90.9
感度	90.7	90.6	90.6	89.4	91.3	90.7	91.7	91.2	91.4
特異度	91.3	69.8	82.1	91.0	74.0	87.4	92.5	86.4	89.9

- ・濃: 濃い画素の結果
- ・薄: 薄い画素の結果
- ・全: 濃, 薄の結果を足したもの

- ・RGBの精度79.8%に比べ、HSI 83.3%, EH比率 88%, EH比率に差分情報を加えたものが89.9%とより精度が改善した。
- ・核の分類の有無だと、分類をした方がしていないものより精度が改善した。
- ・HAPSSの精度が最も高く、90.9%となった。
- ・濃い所だけで識別した場合はHAPSSでは92.1%と最も高い精度が得られた。

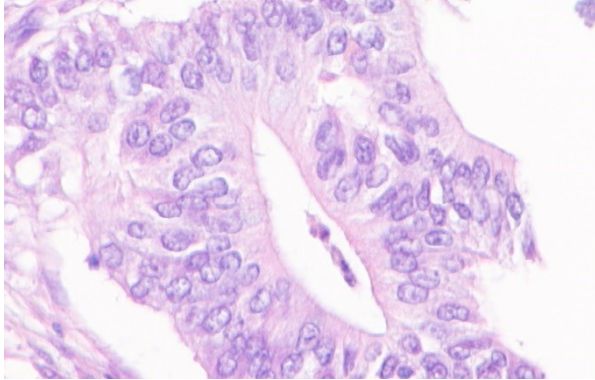
# 分類結果

・核の画素が紫:濃いと分類

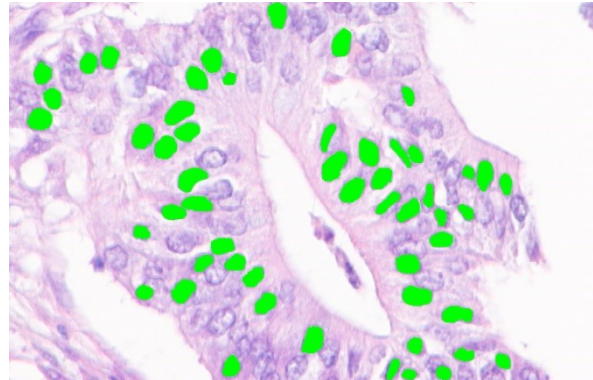
・核の画素がピンク:薄いと分類

・画素が橙(濃), 黄(薄):がんと分類

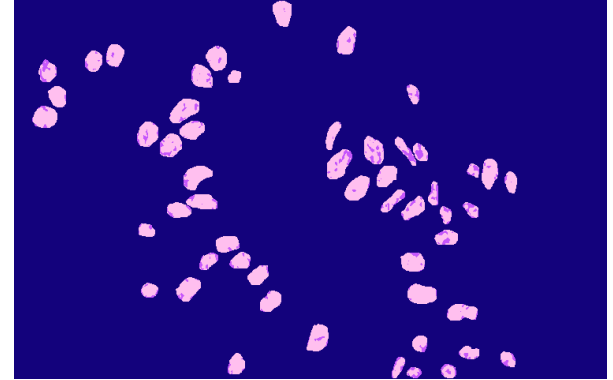
・画素が青(濃), 水(薄):正常と分類



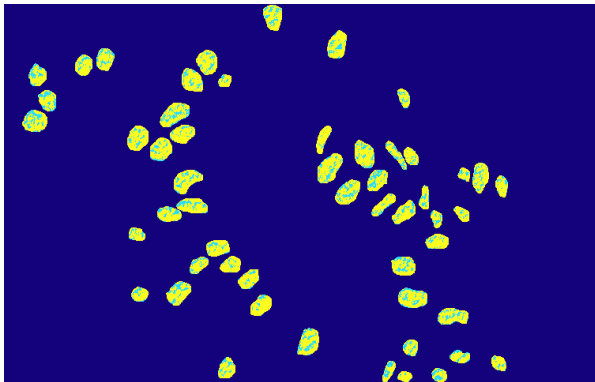
RGB画像 Cancer\_2



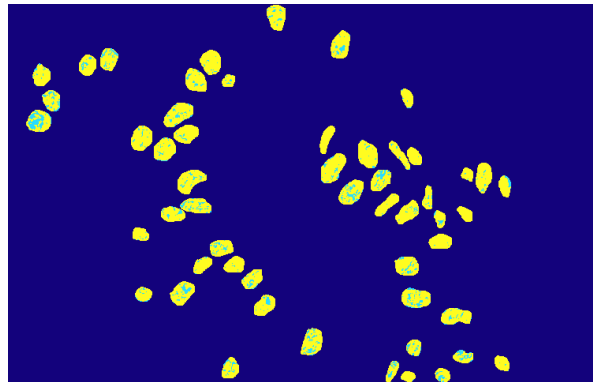
アノテーション画像



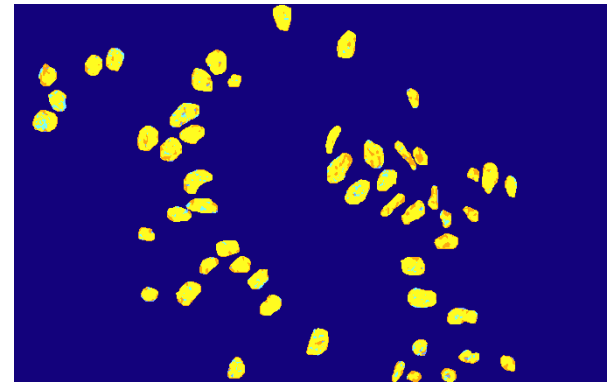
核分類(E透過率閾値:0.2)



RGBの分類結果(82.5%)



HSIの分類結果(91.5%)



HAPSSの分類結果(94.5%)

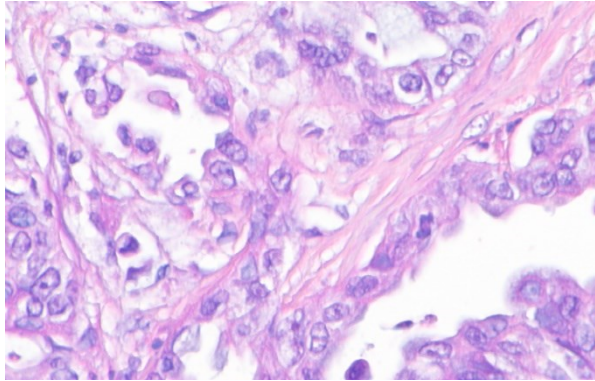
# 分類結果

・核の画素が紫:濃いと分類

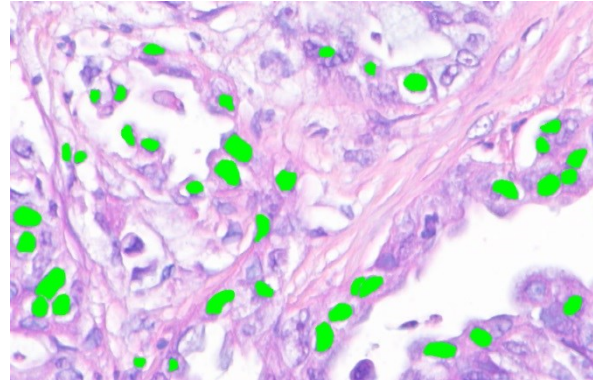
・核の画素がピンク:薄いと分類

・画素が橙(濃), 黄(薄):がんと分類

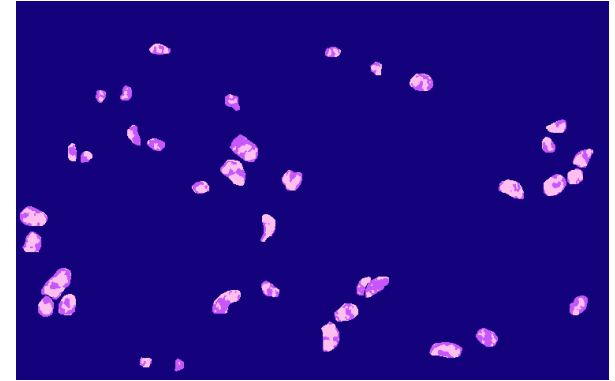
・画素が青(濃), 水(薄):正常と分類



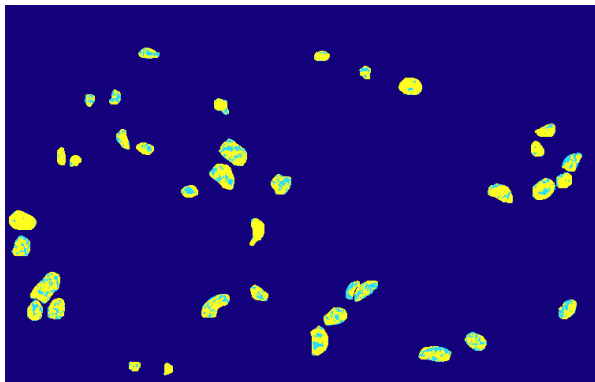
RGB画像 Cancer\_3



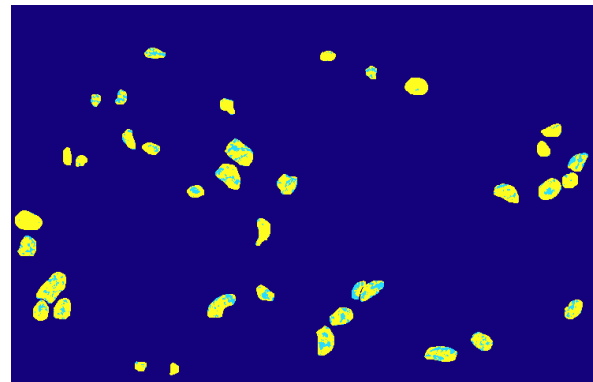
アノテーション画像



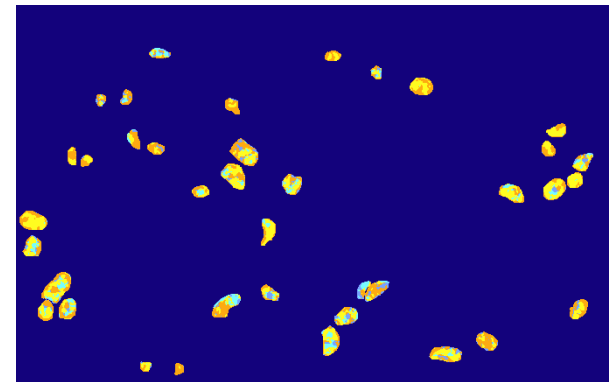
核分類(E透過率閾値:0.2)



RGBの分類結果(71.8%)



HSIの分類結果(78.4%)



HAPSSの分類結果(79.4%)



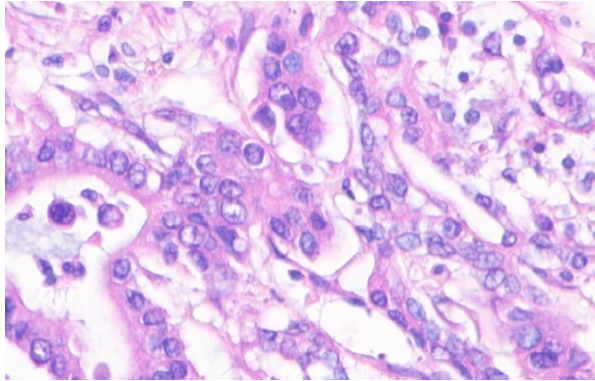
# 分類結果

・核の画素が紫:濃いと分類

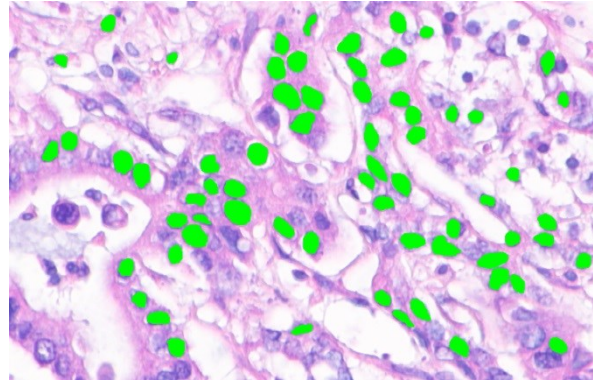
・核の画素がピンク:薄いと分類

・画素が橙(濃), 黄(薄):がんと分類

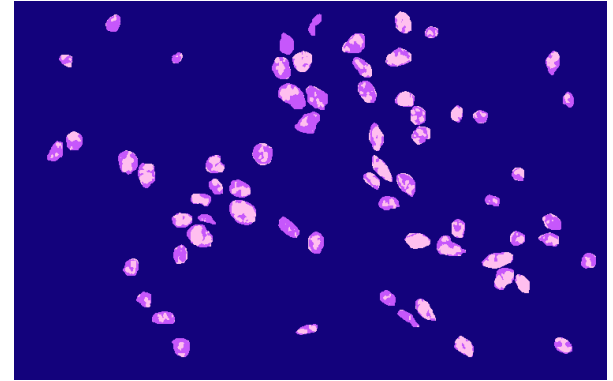
・画素が青(濃), 水(薄):正常と分類



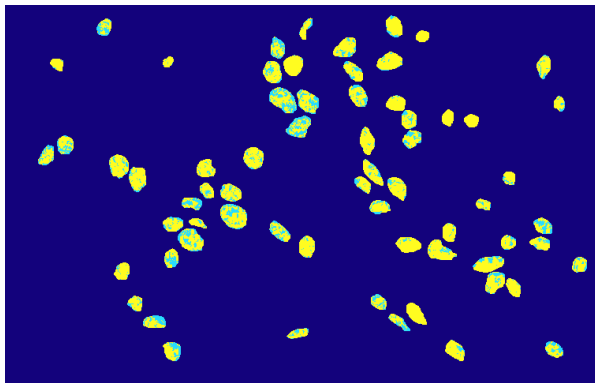
RGB画像 Cancer\_6



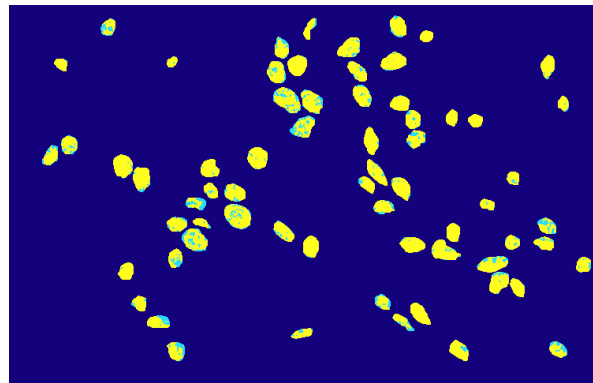
アノテーション画像



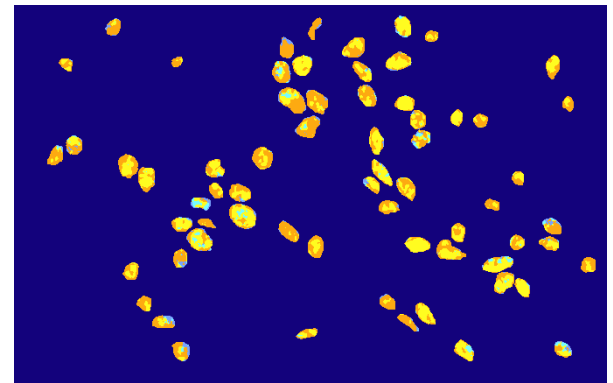
核分類(E透過率閾値:0.2)



RGBの分類結果(79.5%)



HSIの分類結果(81.8%)



HAPSSの分類結果(90.6%)

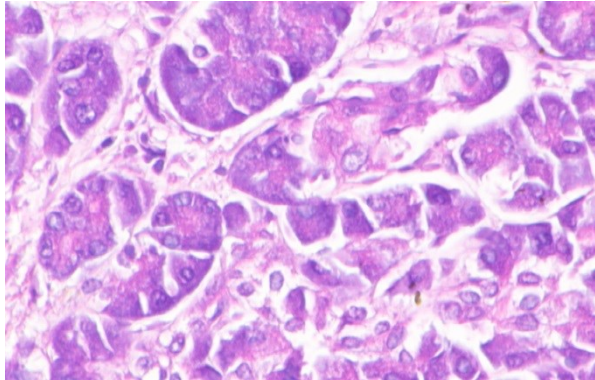
# 分類結果

・核の画素が紫:濃いと分類

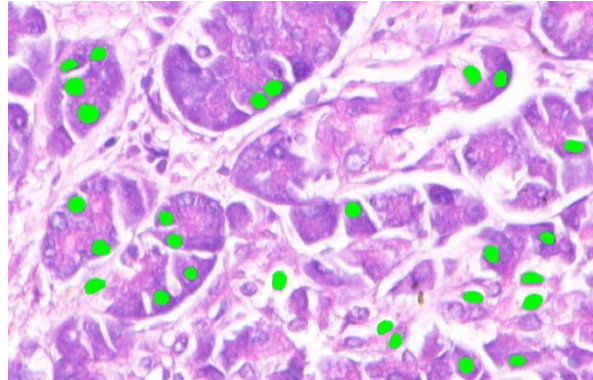
・核の画素がピンク:薄いと分類

・画素が橙(濃), 黄(薄):がんと分類

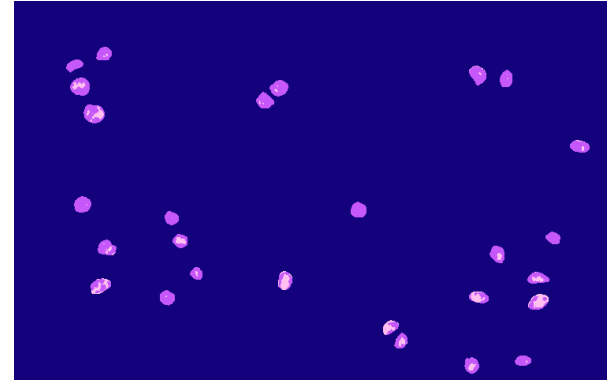
・画素が青(濃), 水(薄):正常と分類



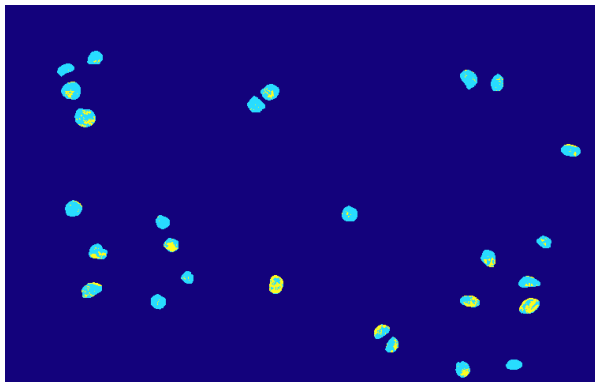
RGB画像 Normal\_3



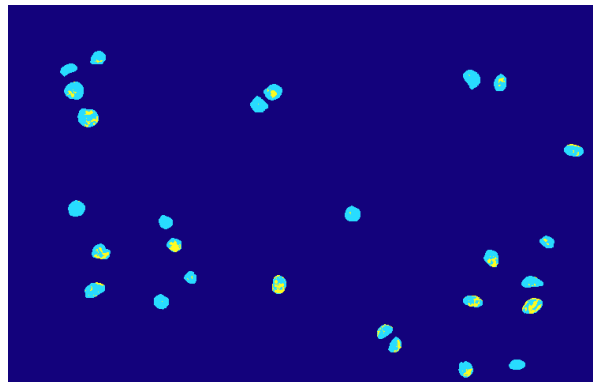
アノテーション画像



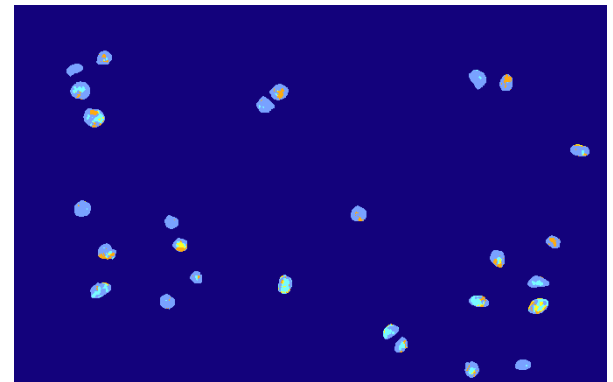
核分類(E透過率閾値:0.2)



RGBの分類結果(79.1%)



HSIの分類結果(85.9%)



HAPSSの分類結果(86.7%)

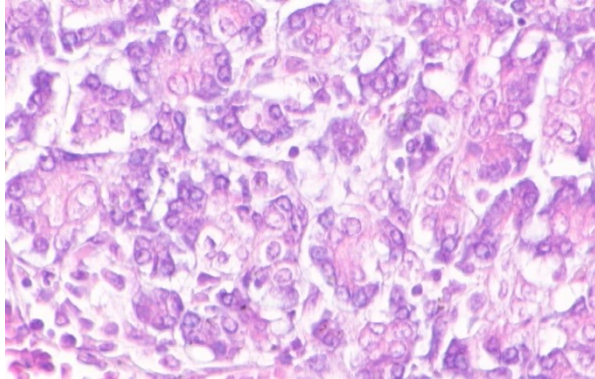
# 分類結果

・核の画素が紫:濃いと分類

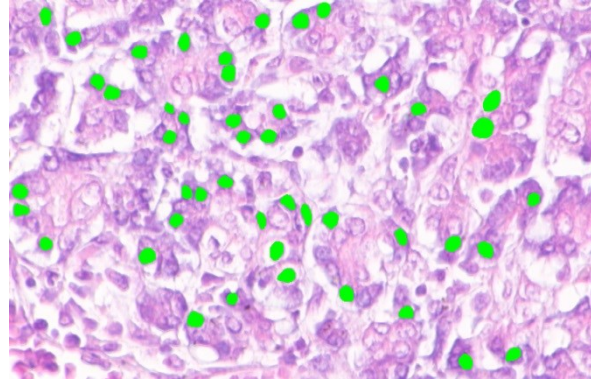
・核の画素がピンク:薄いと分類

・画素が橙(濃), 黄(薄):がんと分類

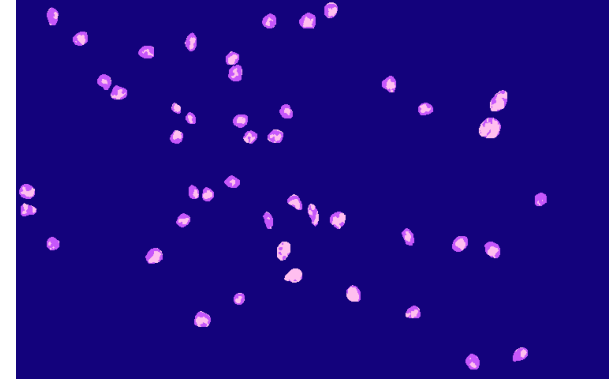
・画素が青(濃), 水(薄):正常と分類



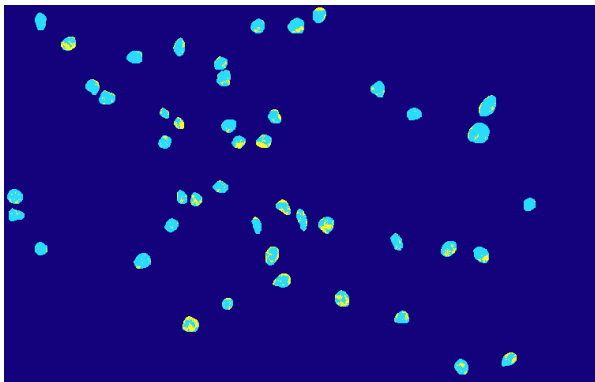
RGB画像 Normal\_5



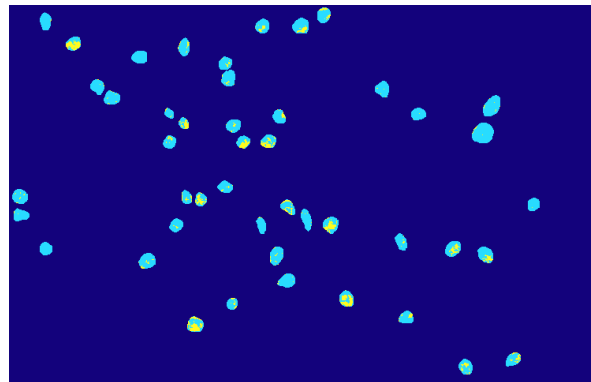
アノテーション画像



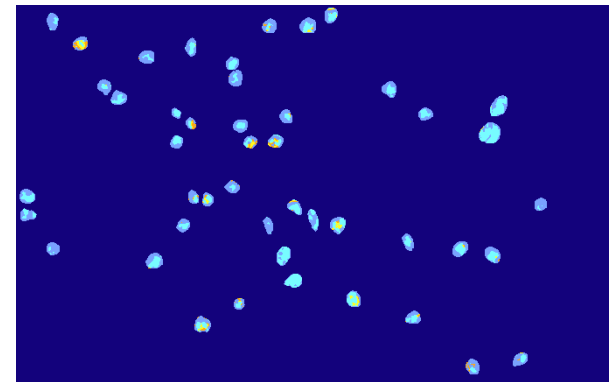
核分類(E透過率閾値:0.2)



RGBの分類結果(81.2%)



HSIの分類結果(85.6%)



HAPSSの分類結果(92.9%)

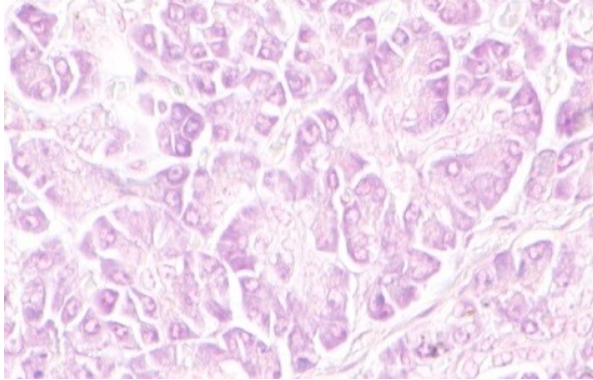
# 分類結果

・核の画素が紫:濃いと分類

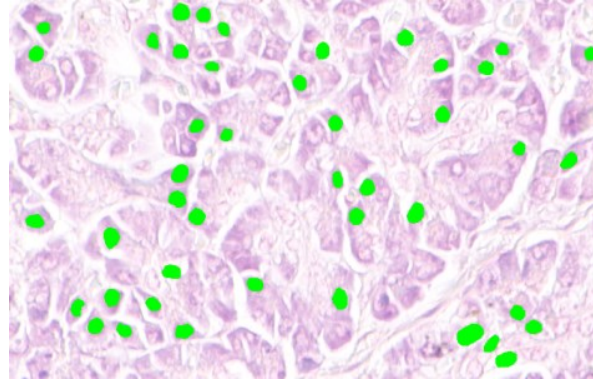
・核の画素がピンク:薄いと分類

・画素が橙(濃), 黄(薄):がんと分類

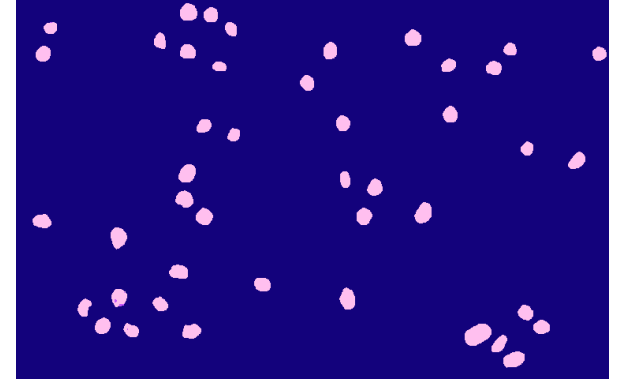
・画素が青(濃), 水(薄):正常と分類



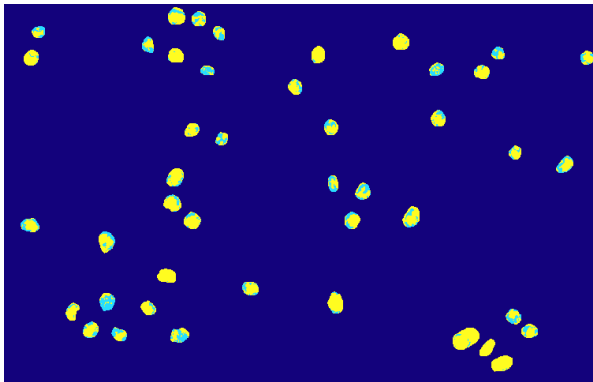
RGB画像 Normal\_6



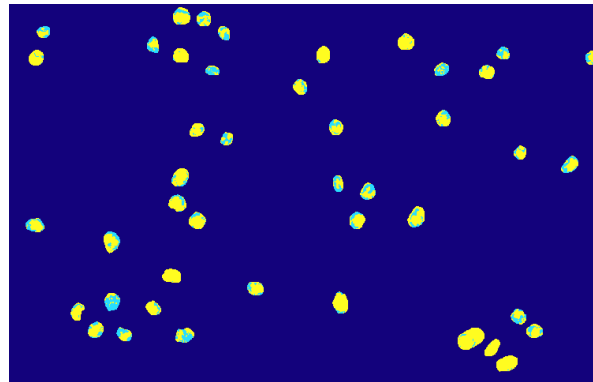
アノテーション画像



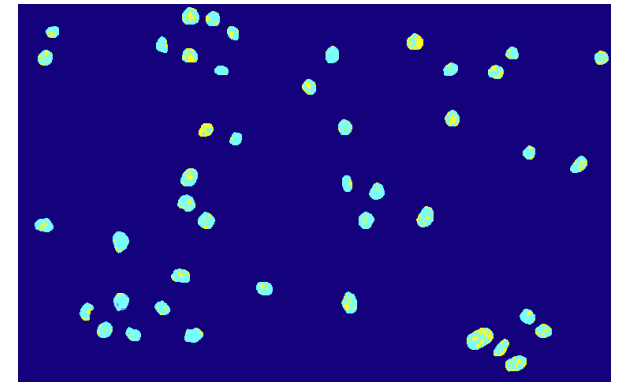
核分類(E透過率閾値:0.2)



RGBの分類結果(37.1%)



HSIの分類結果(20.3%)



HAPSSの分類結果(82.8%)

RGBやHSIの結果に比べてHAPSSの結果では水色の画素が多くなっている



# 医用画像解析における機械学習

---

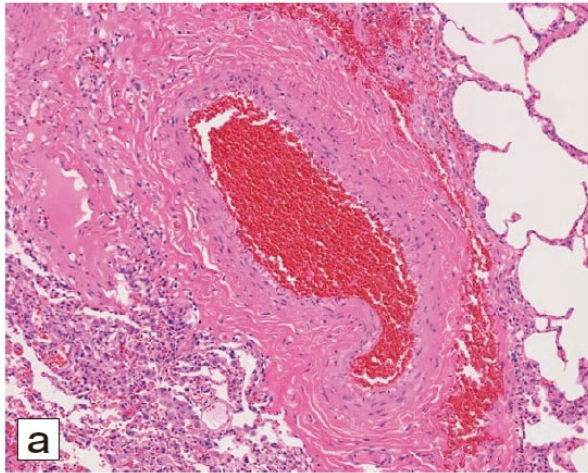
## ■ 医用画像解析の研究

### 🌐 ハイパースペクトルを用いた定量的病理診断支援

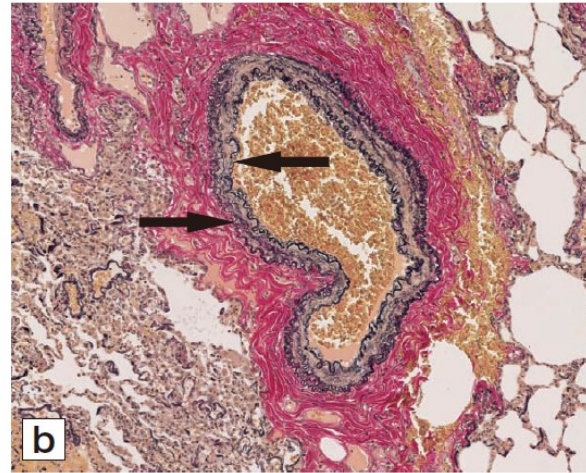
- HE病理標本からのがん陽性核推定
- HE病理標本による線維組織の分類

# 膠原線維と弾性線維

- 膠原線維（コラーゲン）と弾性線維（エラスティン）の分布は、病変部の状態を知るために重要な要素
- H&E染色では膠原線維と弾性線維の存在領域を見分けることが難しいが、Elastica van Gieson（EVG）染色等の特殊染色により見分けることが可能



H&E染色



EVG染色

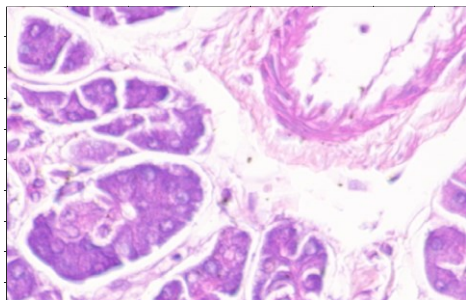
膠原線維＝赤色  
弾性線維＝紫黒色

<http://www.labo.city.hiroshima.med.or.jp/wp-01/wp-content/uploads/2017/01/center201704-03.pdf>

ただし特殊染色は、H&E染色に比べてコストや時間がかかる

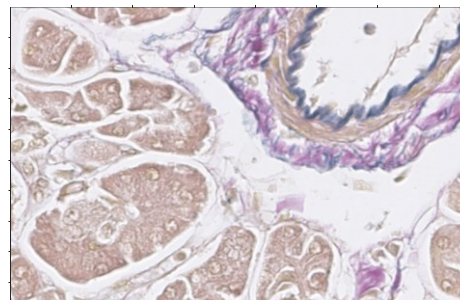
# 本研究の目的

ハイパースペクトル撮影されたH&S画像から**スペクトル情報のみ**で**膠原線維**と**弾性線維**を識別できるかを調査



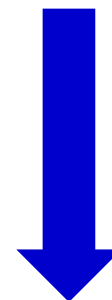
一般染色であるH&E染色画像では**膠原線維**と**弾性線維**の見分けがつかない

従来



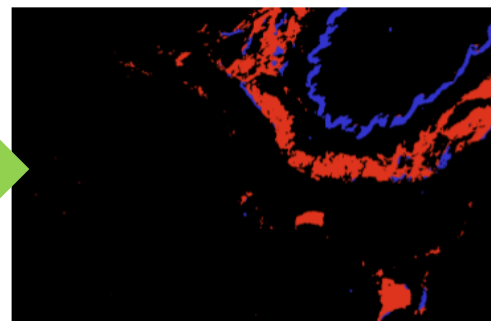
膠原線維と弾性線維を別の色で染め分ける**特殊染色**を利用

時間・コストがかかる



ハイパースペクトルカメラで撮影

提案



H&E染色画像からでも**膠原線維**と**弾性線維**の識別できるか？

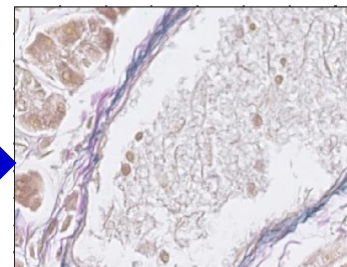
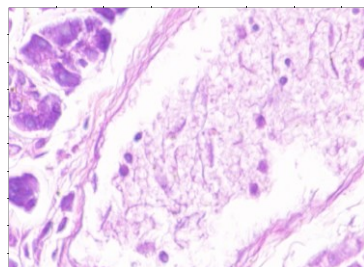
時間・コストを低減

# 方法(method)

## 肝臓の病理標本

H&E染色標本

EVG染色標本



ハイパースペクトル画像 (撮影)  
RGB画像 (再構成)

RGB画像 (Whole Slide スキャナ)

位置合わせ

膠原線維、弾性線維領域抽出

学習によって  
識別関数を作成

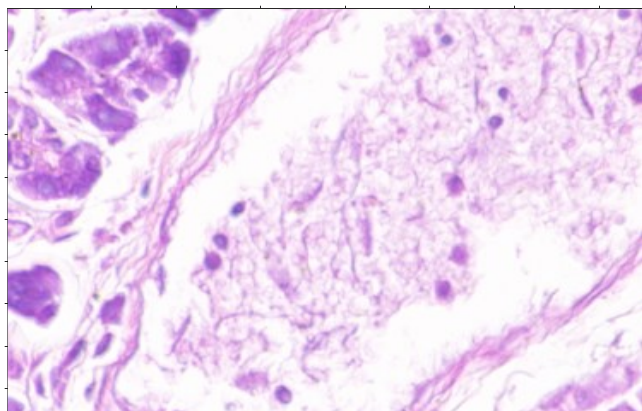
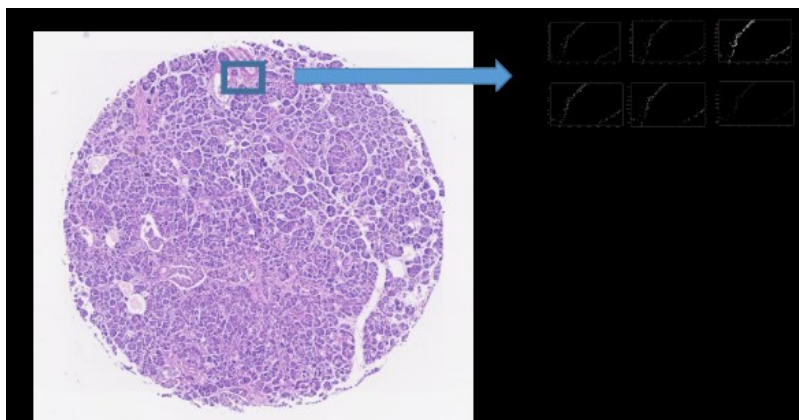
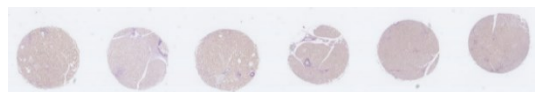
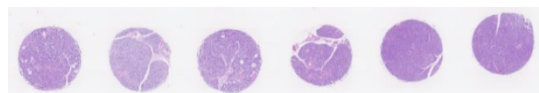
線維領域識別画像

正解画像

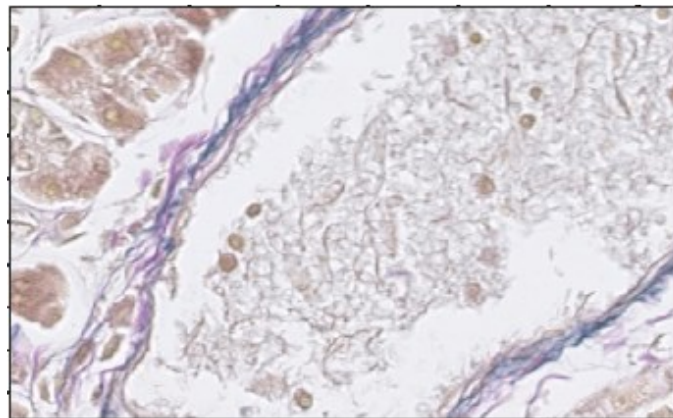


# 顕微鏡画像の撮影

- 大きな画像から特徴的な部位を取り出して利用



H&E染色画像



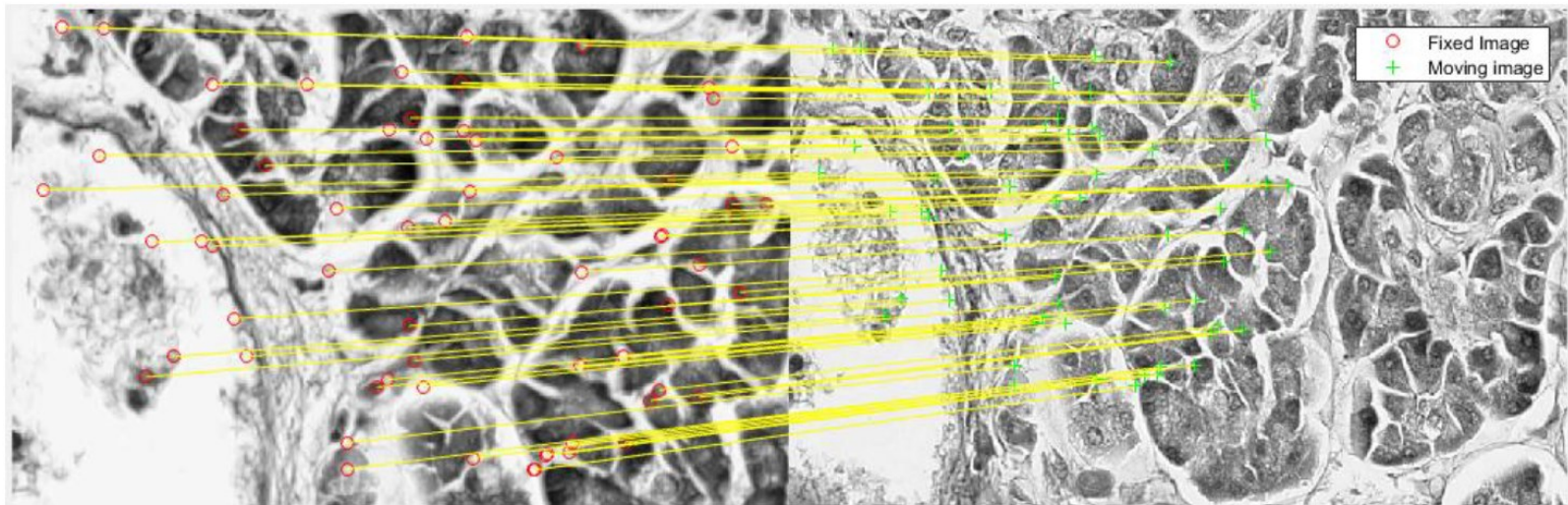
EVG染色画像

# H&E画像とEVG画像のレジストレーション

H&E染色画像とEVG染色画像は、異なる装置で撮影しているため、位置、解像度等が異なる



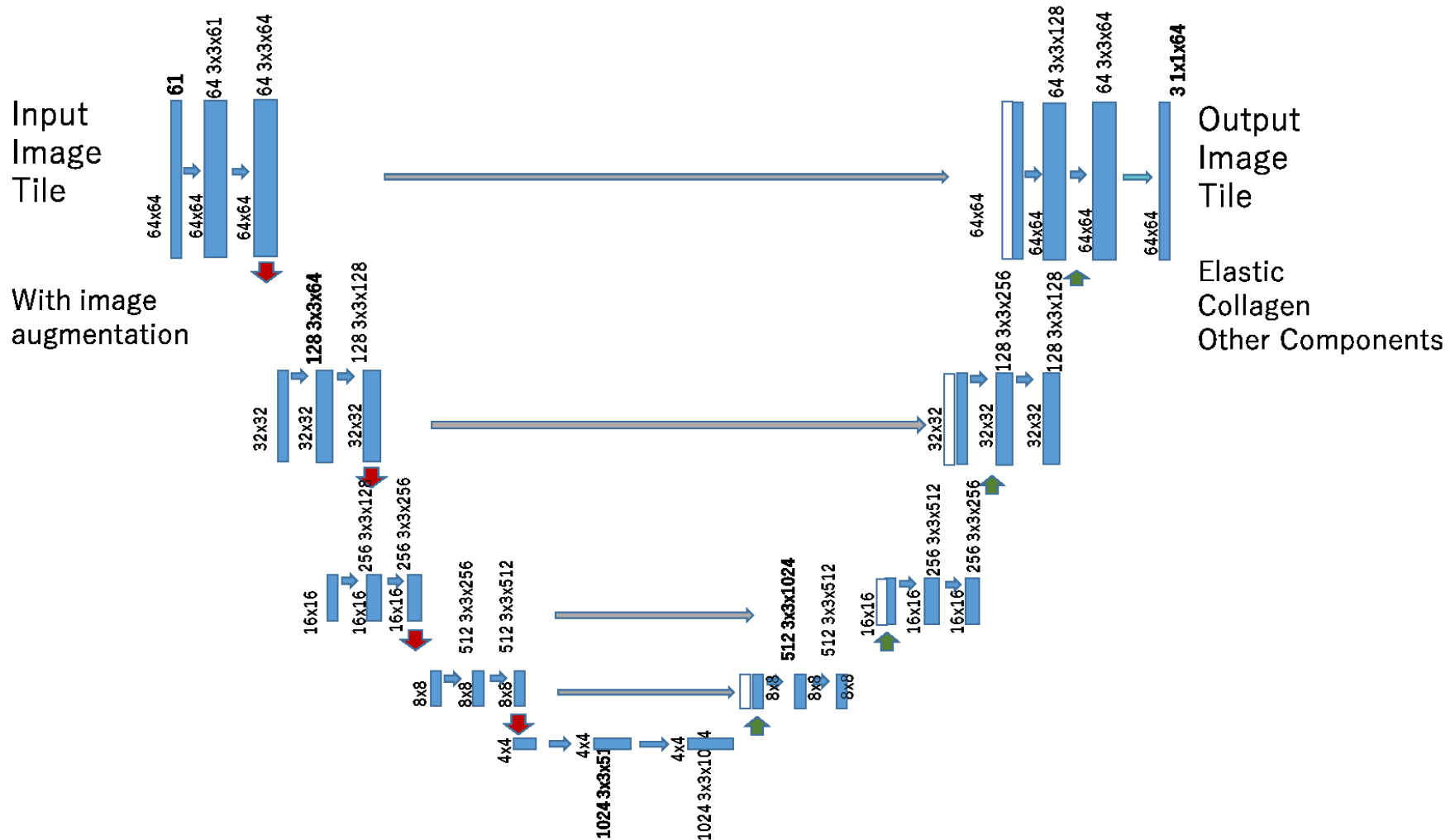
2つの画像の位置・解像度調整（レジストレーション）が必要



**SURF特徴量に基づく  
レジストレーション手法を適用**

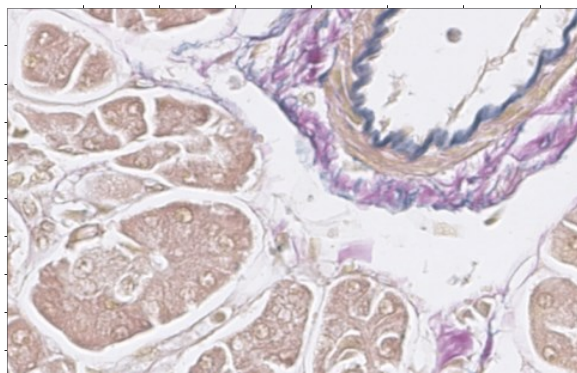
# U-NET

U-net:セグメンテーションに特化したCNN型深層学習の一種。  
画像サイズを変えながら学習することで対象の大小の影響を抑えながら学習可能となる。

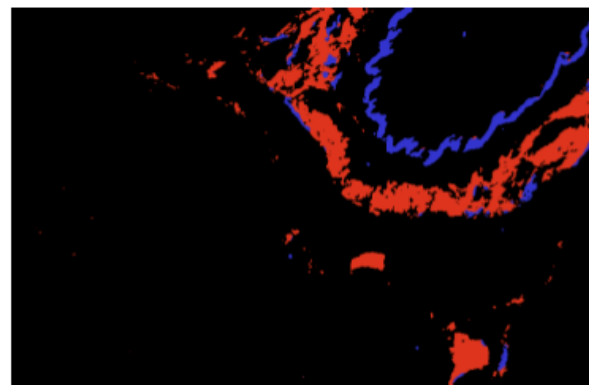


# 正解となる識別画像の作成

- EVG画像（RGB）から各要素を分類する手法を適用  
Abe et.al., Pathology International, 2013.
- ベイズ推定に基づく統計的分類により、H&E染色の5つの要素（**膠原繊維**, **弾性線維**, 核, 細胞質、グラス）を分類可能



EVG染色画像  
(RGB)

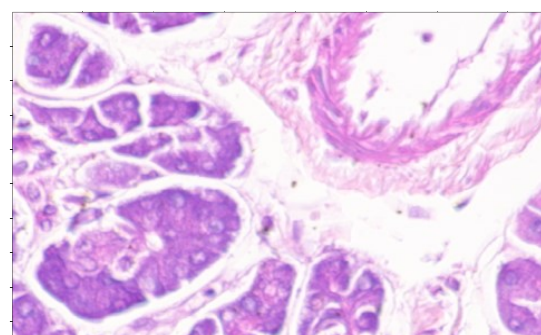
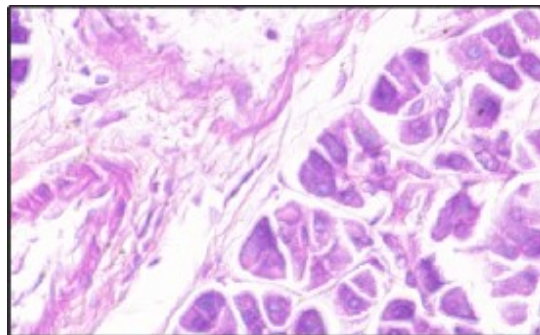
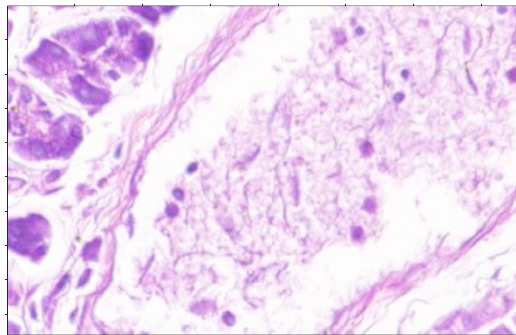


線維領域の識別結果画像  
(これを識別結果の正解とする)

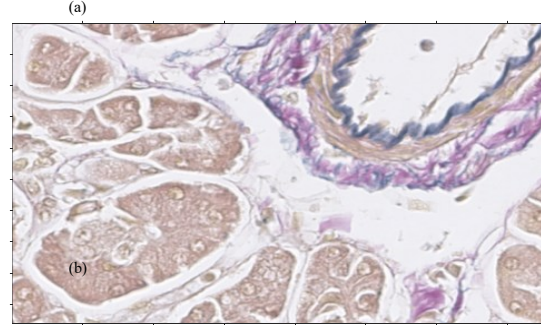
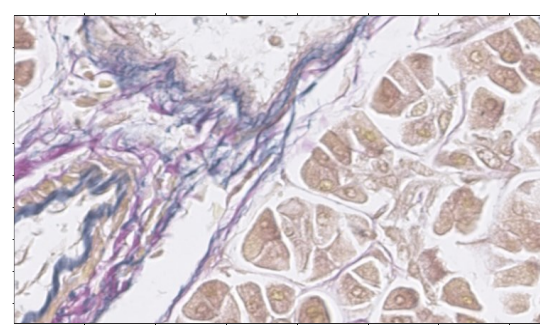
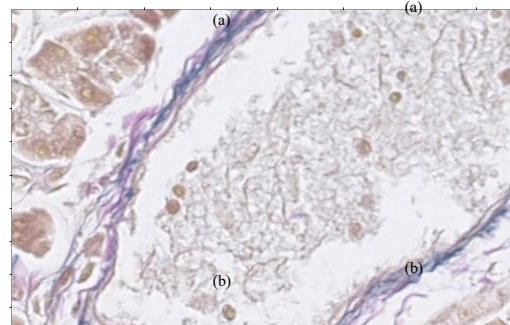


# 実験データ 3画像

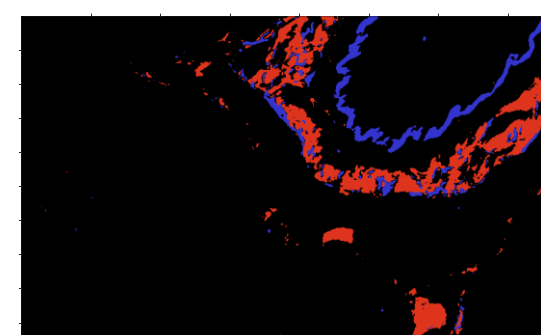
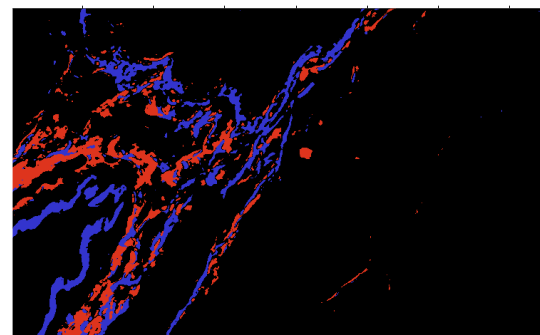
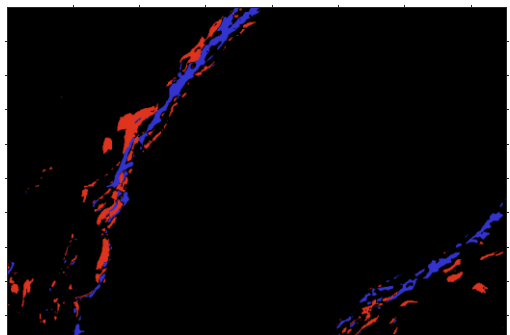
HE



EVG



正解



画像1

画像2

画像3

# サブジェクトベース交差検証

交差検証1 学習データ 画像1、画像2

検証データ 画像3

パッチ数	Hyperspectral	RGB
100	83.9411	80.2338
500	86.3672	85.0735
1000	87.1848	86.6857

交差検証2 学習データ 画像2、画像3

検証データ 画像1

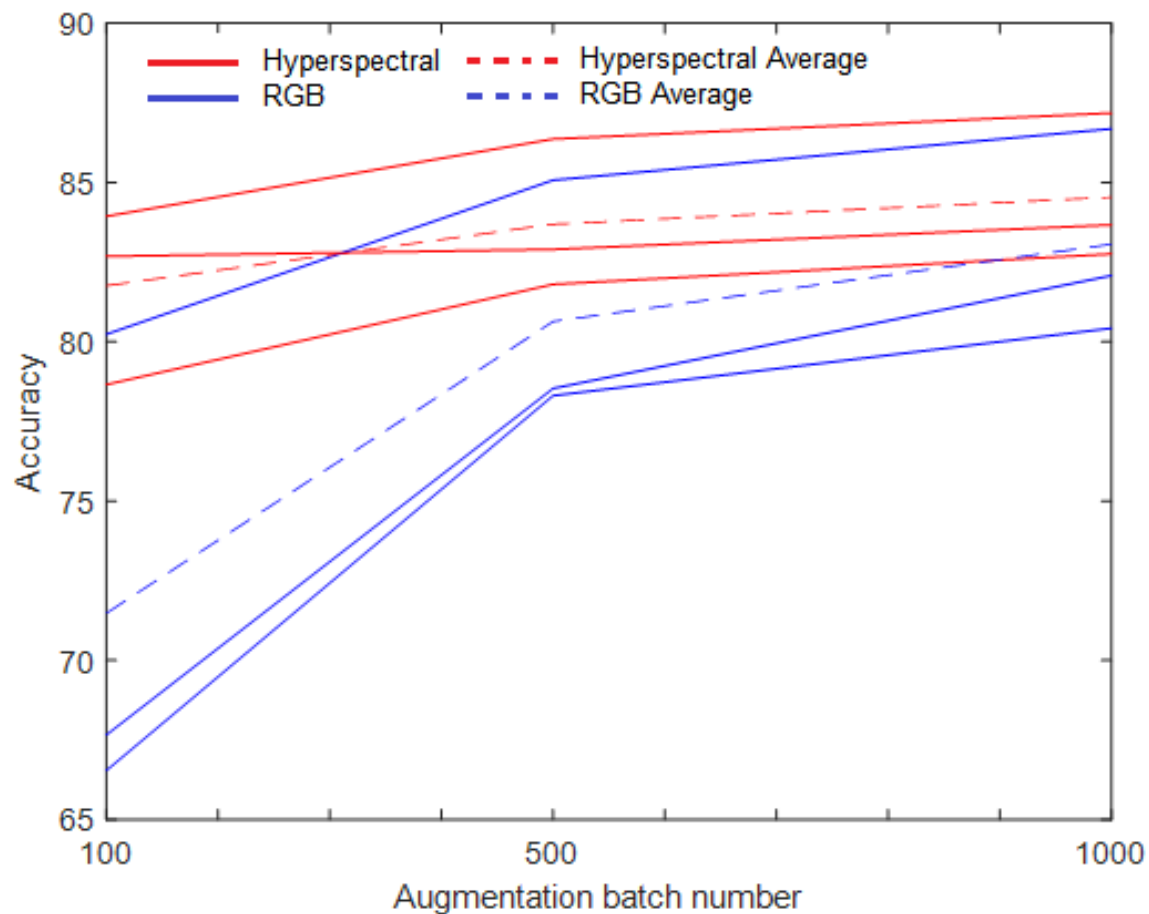
パッチ数	Hyperspectral	RGB
100	82.6852	67.6588
500	82.9025	78.5357
1000	83.6701	82.0769

交差検証3 学習データ 画像1、画像3

検証データ 画像2

パッチ数	Hyperspectral	RGB
100	78.6587	66.5400
500	81.8063	78.3075
1000	82.7576	80.4277

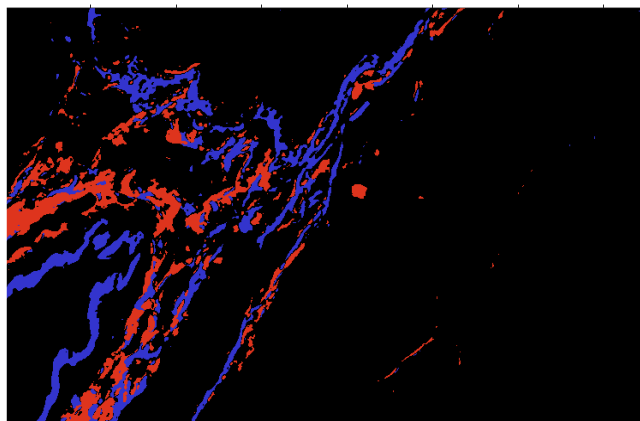
# 結果： 平均精度(パッチ数)



パッチ数	Hyperspectral	RGB
100	81.76167	71.47753
500	83.692	80.6389
1000	84.5375	83.06343

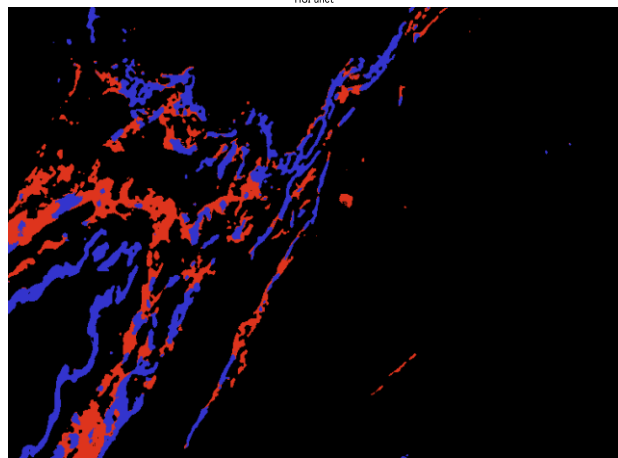
# 交差検証 3 (学習: 画像1、画像3, 検証: 画像2)

正解データ



パッチ数100

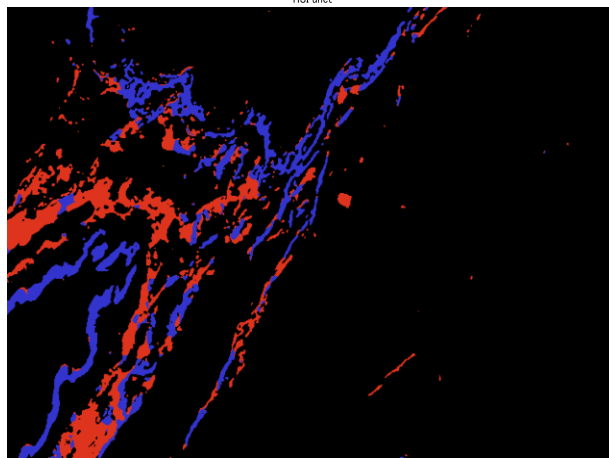
HSI unet



78.6%

パッチ数500

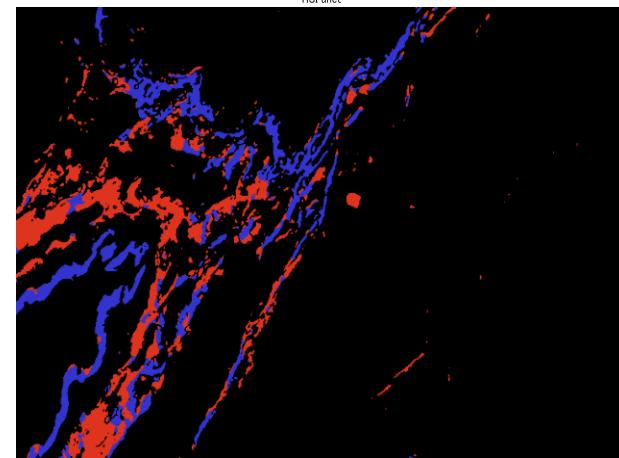
HSI unet



81.8%

パッチ数1000

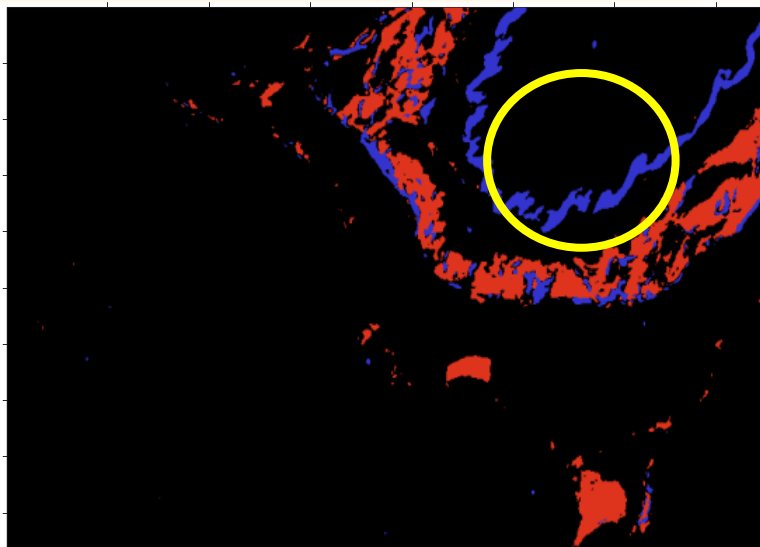
HSI unet



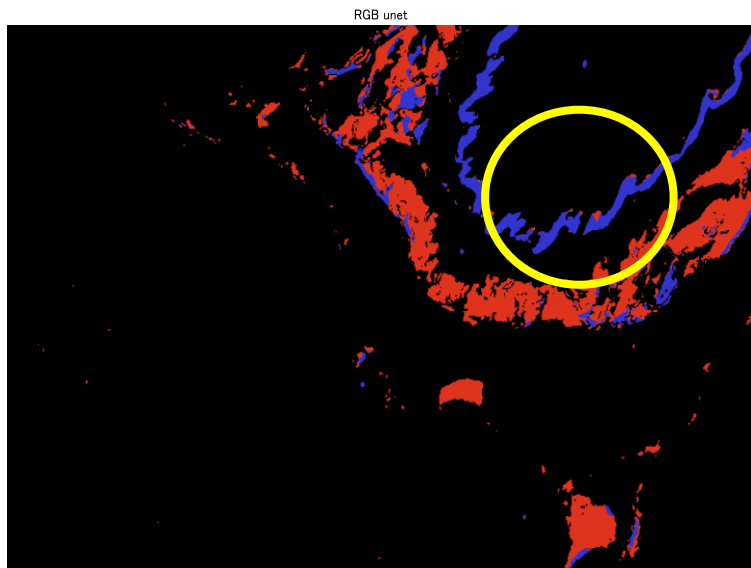
82.8%

# 交差検証1: 学習データ: Case 1 Case 2 検証データ: Case 3

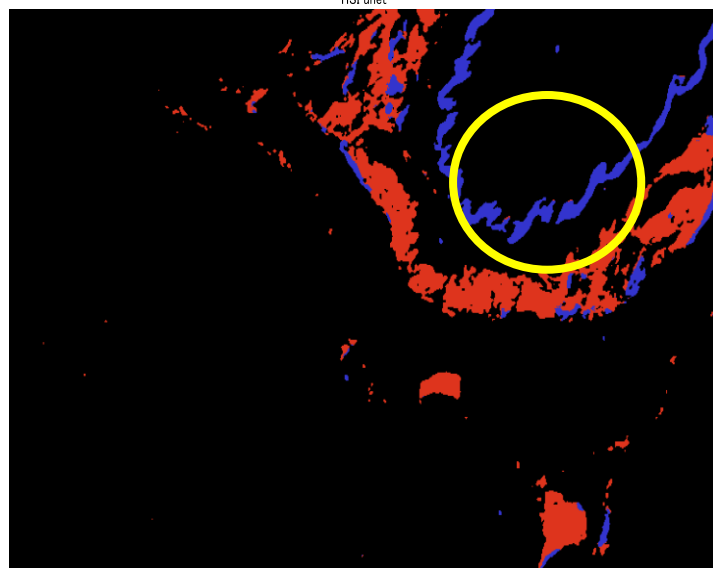
正解



RGB 78.5 %

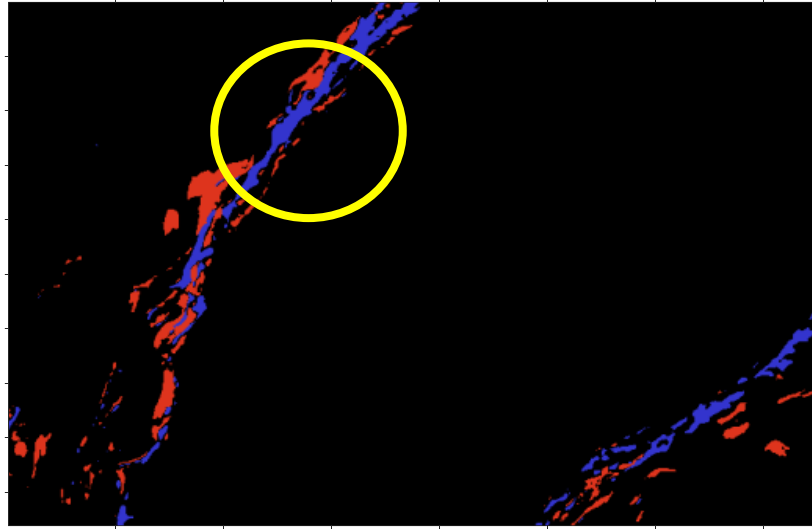


Hyperspectral 86.3%



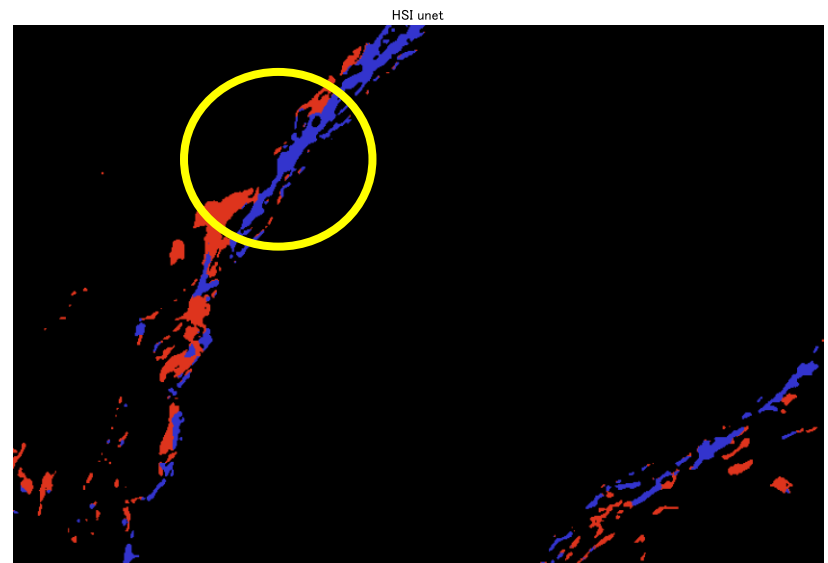
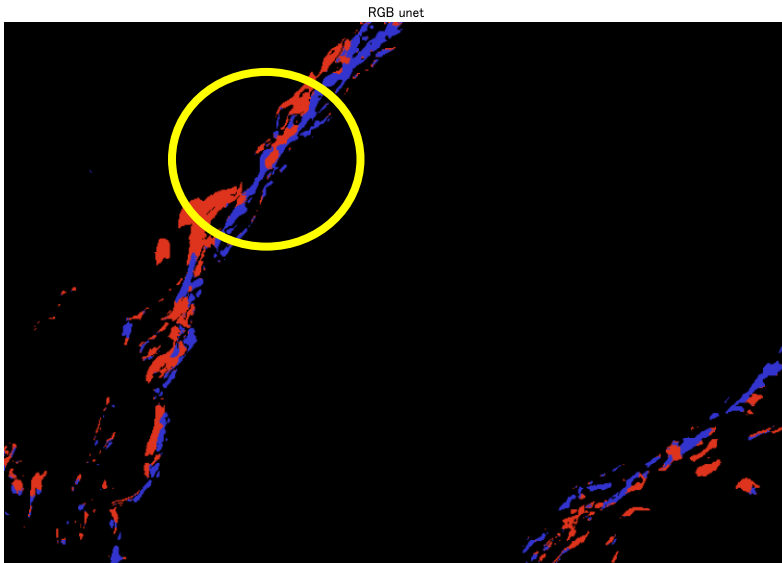
交差検証2: 学習データ: 画像2 画像3 検証データ: 画像1

正解

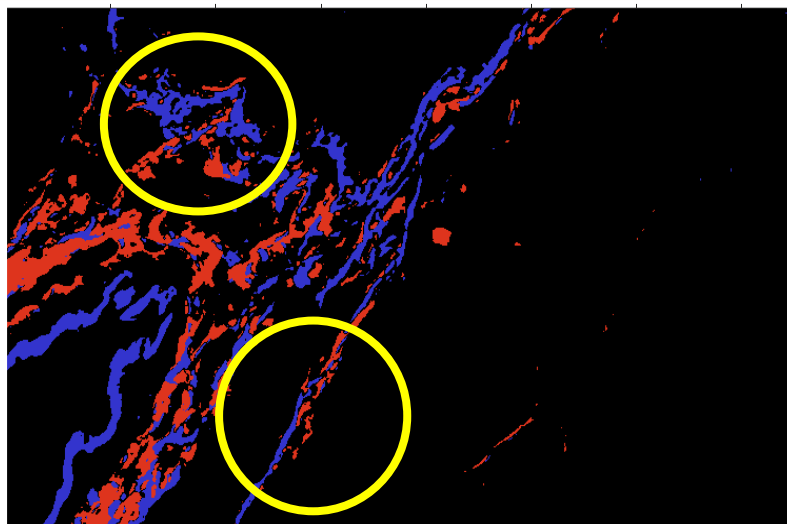


RGB 78.5%

Hyperspectral 82.9%

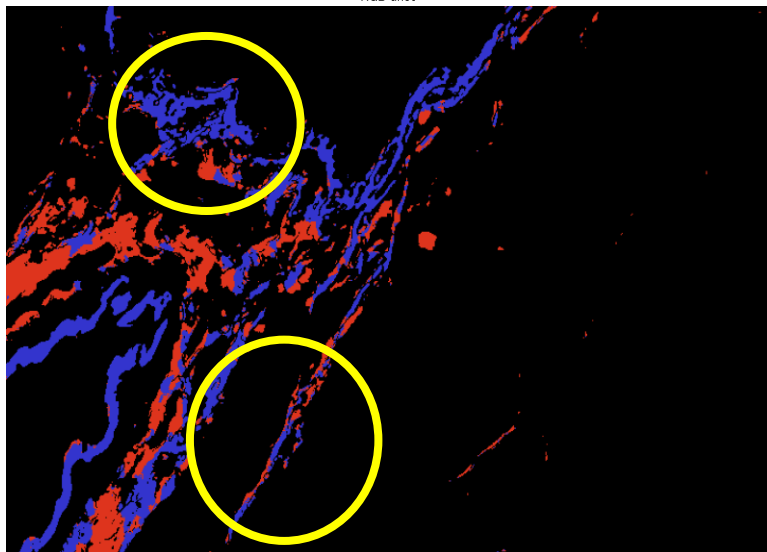


正解



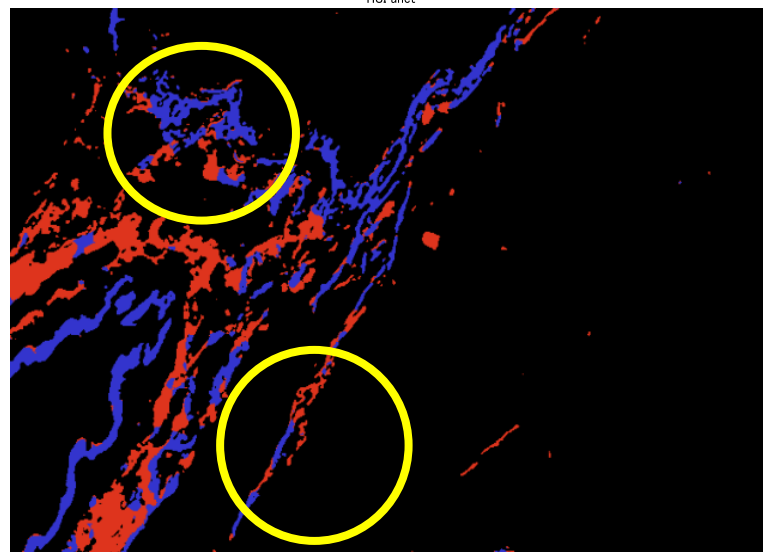
RGB 78.0%

RGB unet



Hyperspectral 81.8%

HSI unet



# 結論

---

- ハイパースペクトル対応U-NETにより、ハイパースペクトルH&E画像により、膠原線維と弾性線維を85%の精度で識別可能なことを示した。



# 機械学習

どの学習方法が適しているか？

近年、深層学習の高精度化に伴い、様々な課題を深層学習で求める傾向がある。

⇒ 問題によっては適さない課題も存在する。

- NN学習（深層学習） ⇒ 入力データで学習（特徴量は自動生成）
  - 人間が目視で確認可能課題 ⇒ 精度良く推定可能
  - 画像から単純に得られないデータ推定 ⇒ 推定精度？
  - 少ないデータ数 ⇒ 推定精度？ 過学習？
- 古典的機械学習 ⇒ 特徴量を求め、特徴空間上で学習
  - データ数が限られている。
  - 人間では見えにくい機能の推定。
  - 有効な特徴量は？ ここが重要

# 関連学会

---

## (遠隔診療・遠隔医療)

- 2022 IEEE 4<sup>th</sup> Global Conference on Life Science and Technologies (Lifetech 2022)
  - March 7-9, 2022 Osaka, Japan (Hybrid)
  - Organized Session: Sensor Technology for Telemedicine (Dead line Dec. 17 2page)
  
- 第26回日本遠隔医療学会
  - 2022年 10月28日～29日
  - 埼玉会館(さいたま市)

