

病理画像のがんらしさと正常らしさを 情報量で評価する情報密度法

金沢工業大学 工学部 情報工学科

教授 金道 敏樹

2023年2月28日

従来技術とその問題点

【従来研究】



- 病理画像の自動診断に、深層学習を応用
- 診断性能は、専門医の診断以上

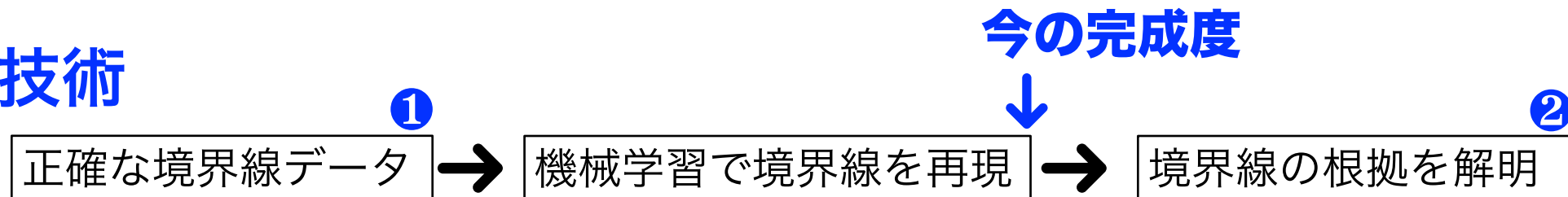
Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer
B. E. Bejnordi, et.al.; and the CAMELYON16 Consortium [JAMA. 2017;318\(22\):2199-2210. doi:10.1001/jama.2017.14585](https://doi.org/10.1001/jama.2017.14585)

【その問題点】

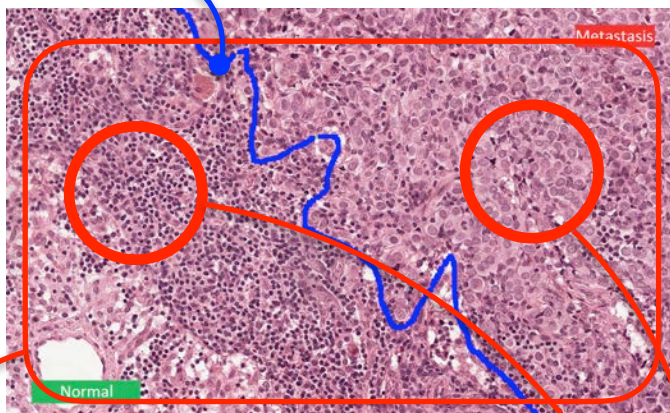
- 深層学習では、**診断根拠が分からない**
→ 結果の根拠づけの研究が必要

新技術の特徴と従来技術との比較

従来技術



①②が開発のボトルネック



<https://camelyon16.grand-challenge.org/Data/>

①②のボトルネックを解消

新技術



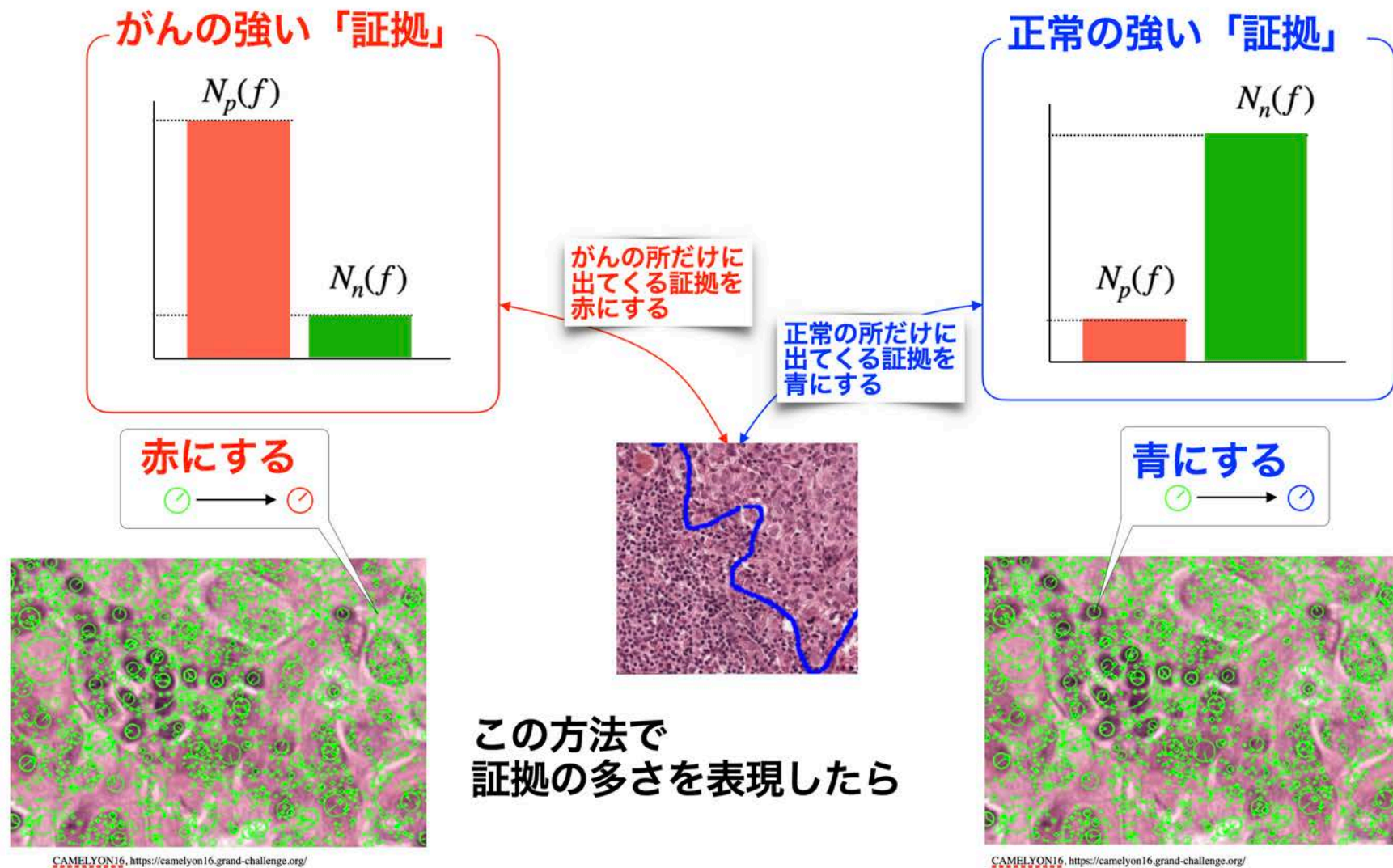
正確な境界線データは不要

画像著作権情報

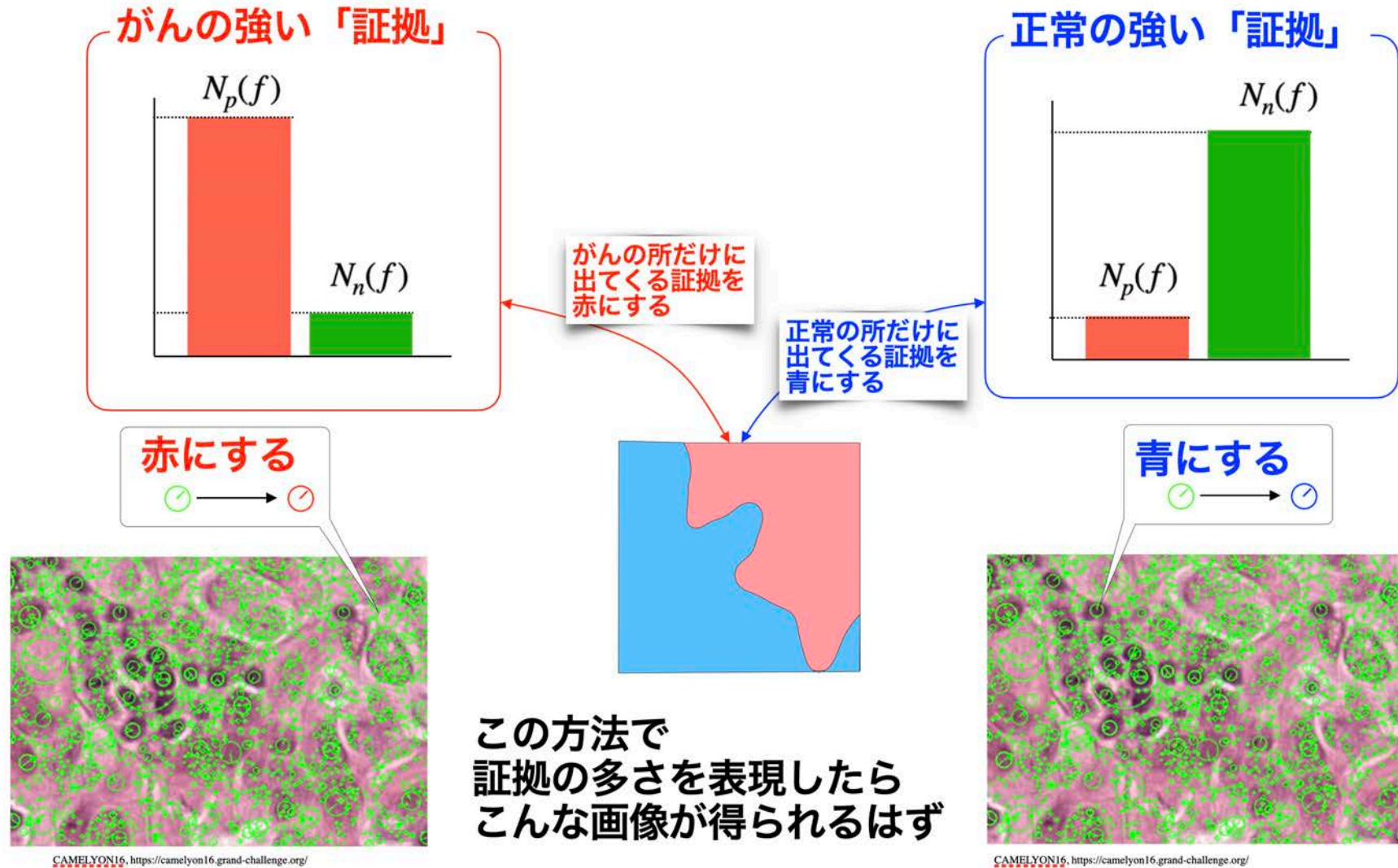
URL: <https://camelyon16.grand-challenge.org/data>

License: <https://creativecommons.org/publicdomain/zero/1.0>

強い「証拠」を見つけ、色付けする



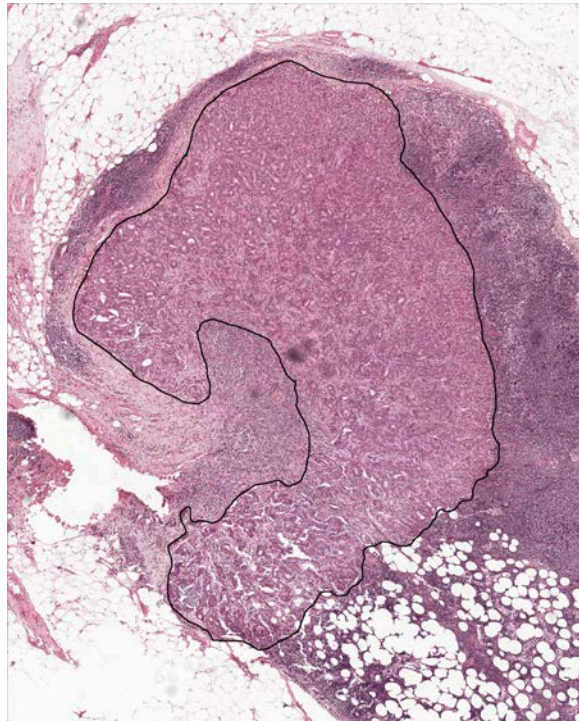
強い「証拠」を見つけ、色付けする



新技術の現状

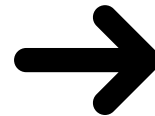
正確な境界線データは不要

入力画像

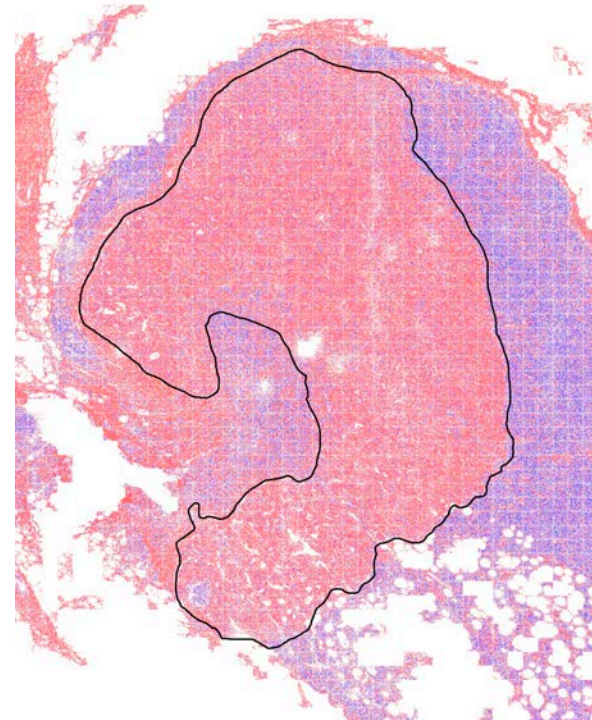


バーチャルスライド (医用画像)

<https://camelyon16.grand-challenge.org/Data/>



新技術で、識別根拠となる
特徴量分布を可視化した画像

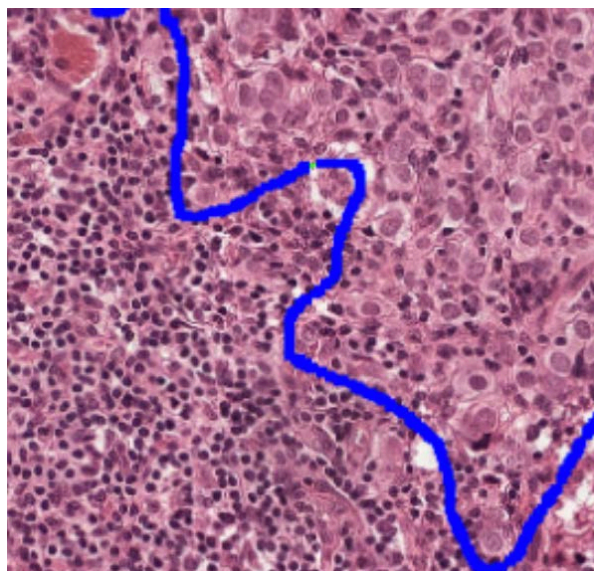


「がん」の情報を持った特徴量の分布
「正常」の情報を持った特徴量の分布

新技術の第1の問い 証拠は十分にあるのか？

病理画像を情報量で評価する情報密度法

そもそも
がんと正常をみて区別でき、
こんな境界線を弾くことができるとは...

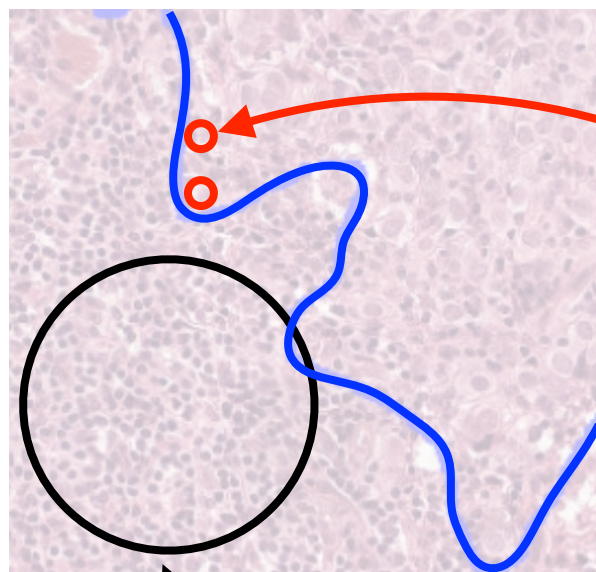


CAMELYON16, <https://camelyon16.grand-challenge.org/>

1) 画像中に、がんと正常を区別できる証拠がある

病理画像を情報量で評価する情報密度法

そもそも
がんと正常をみて区別でき、
こんな境界線を弾くことができるとは...



CAMELYON16, <https://camelyon16.grand-challenge.org/>

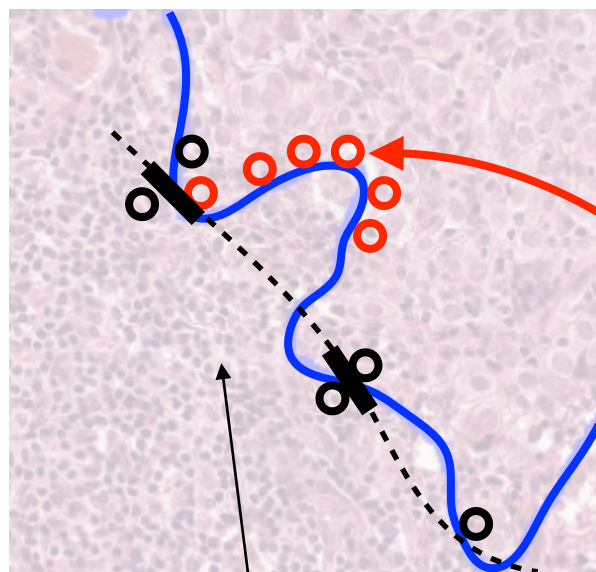
1) 画像中に、がんと正常を区別できる証拠がある

2) その証拠の大きさは、この境界線の曲がり具合を説明できるほど小さい

証拠がこんなに大きいと、
境界線はもっとゆったりとしてるはず

病理画像を情報量で評価する情報密度法

そもそも
がんと正常をみて区別でき、
こんな境界線を弾くことができるとは...



CAMELYON16, <https://camelyon16.grand-challenge.org/>

1) 画像中に、がんと正常を区別できる証拠がある

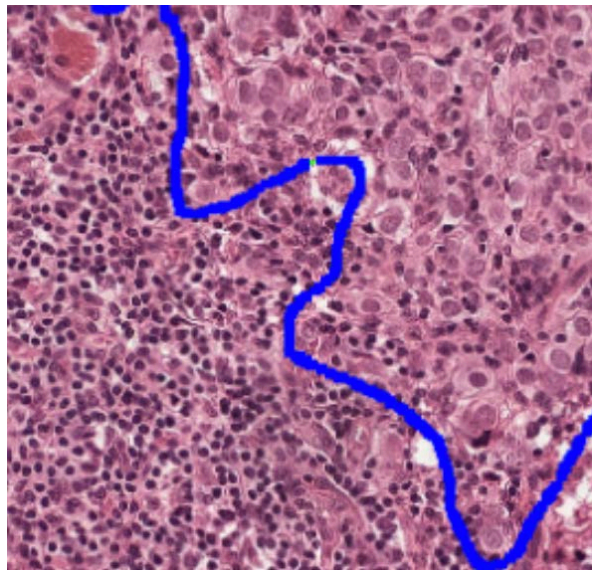
2) その証拠の大きさは、この境界線の曲がり具合を説明できるほど小さい

3) 一筆で境界線を引けるほど、証拠は沢山ある

証拠がこんなに少ないと
境界線はところどころでしか引けないはず

病理画像を情報量で評価する情報密度法

そもそも
がんと正常をみて区別でき、
こんな境界線を弾くことができるとは...



CAMELYON16, <https://camelyon16.grand-challenge.org/>

- 1) 画像中に、がんと正常を区別できる証拠がある
- 2) その証拠の大きさは、この境界線の曲がり具合を説明できるほど小さい
- 3) 一筆で境界線を引けるほど、証拠は沢山ある

証拠は十分沢山ある
おそらくそれを見つけることは難しくない

新技術の第2の問い

何を証拠の候補とし、 どうやって絞りこむか？

何を証拠の候補にするか？

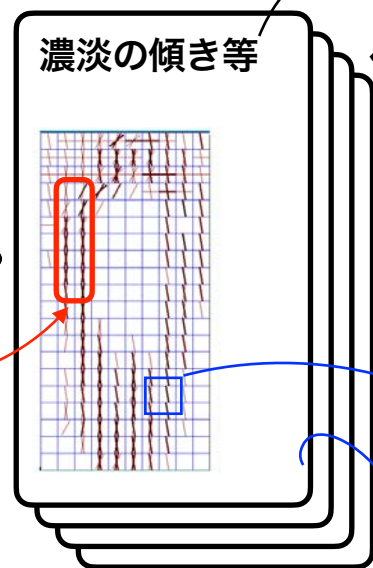
f : 画像の特徴

- 性質がよく分かっている
- 画像処理技術で実績がある

画像処理では、処理の最初に



Histograms of oriented gradients for human detection,
N. Dalal; B. Triggs, et.al, CVPR2005



「画像特徴量」と呼ばれる

四角の一つ一つが画像特徴量
山ほどの提案がある



濃淡の傾き使うものの一つ

SIFT を使用する

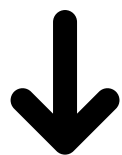
Scale-Invariant Feature Transform

この例だと
画像中の腕の輪郭線が、
特徴的な縦線として
現れている

何を証拠の候補にするか？

f : 画像の特徴

- 性質がよく分かっている
- 画像処理技術で実績がある

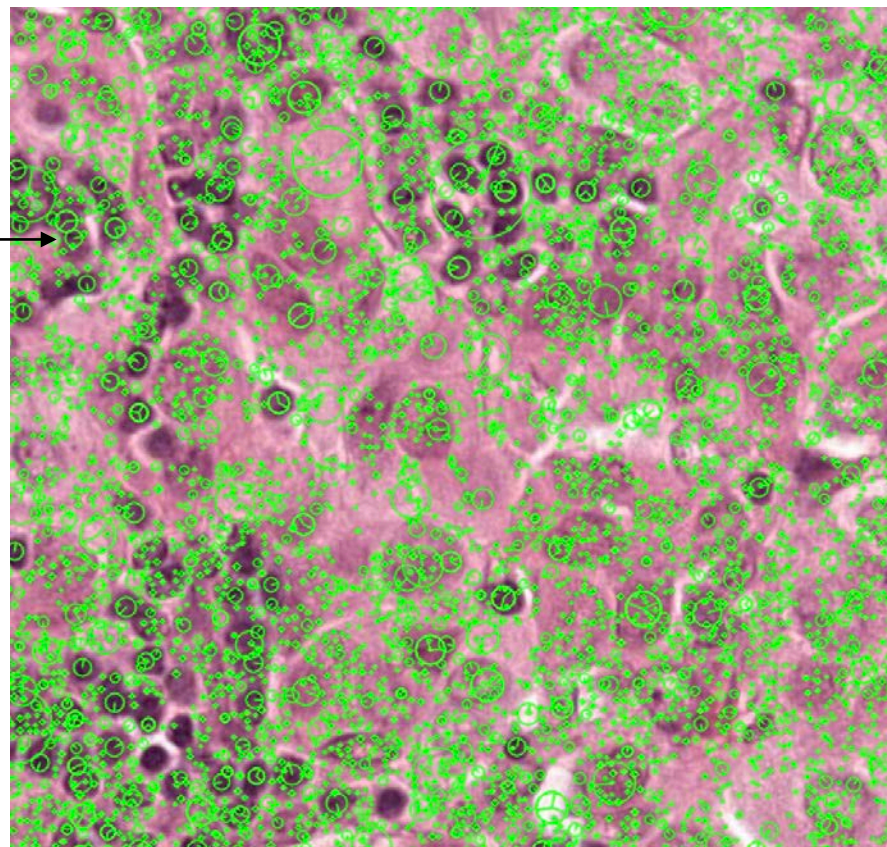


SIFT

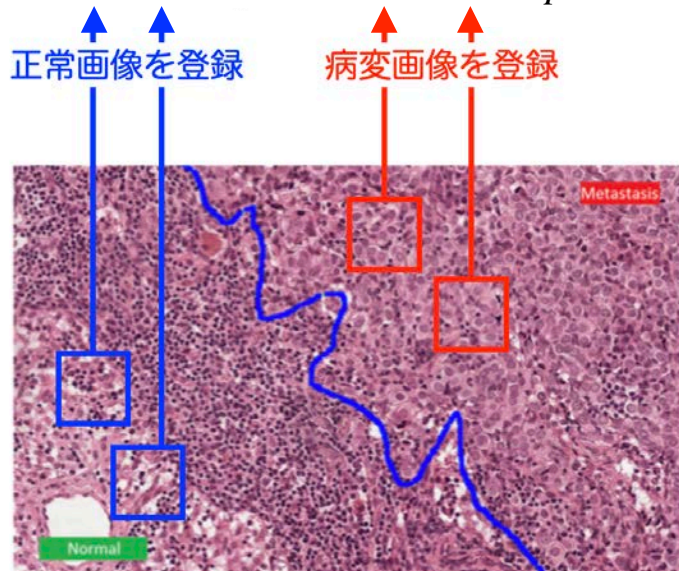
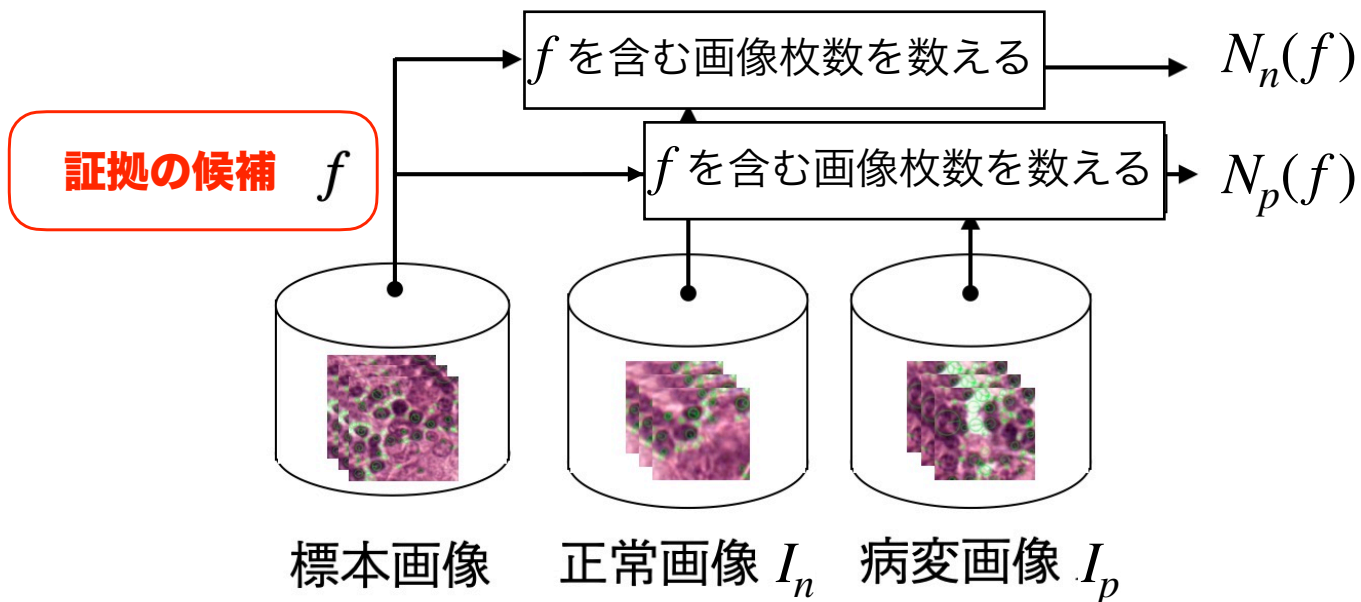
Scale-Invariant Feature Transform

- 1) 画像中に、がんと正常を区別できる証拠がある
- 2) その証拠の大きさは、この境界線の曲がり具合を説明できるほど小さい
- 3) 一筆で境界線を引けるほど、証拠は沢山ある

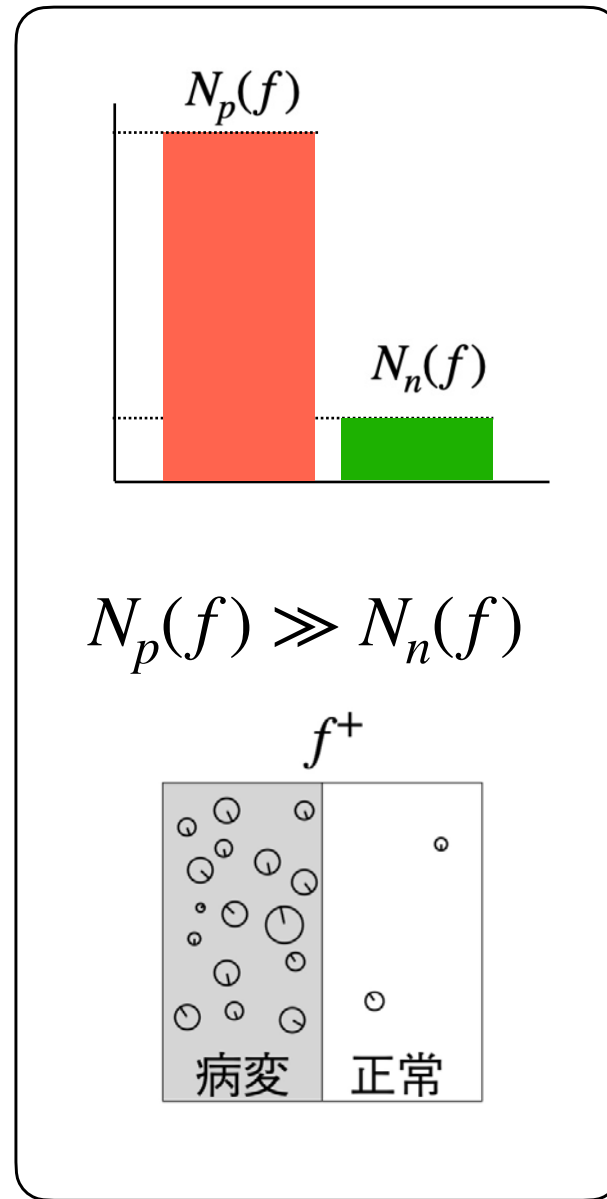
OK!



証拠の見つけ方

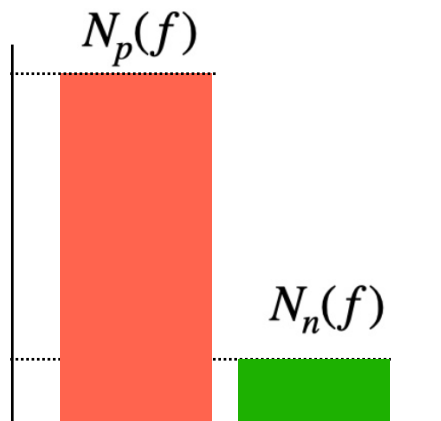


<https://camelyon16.grand-challenge.org/Data/>



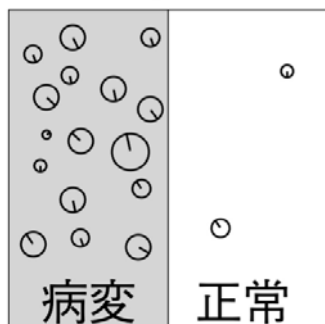
強い「証拠」と弱い「証拠」

強い「証拠」

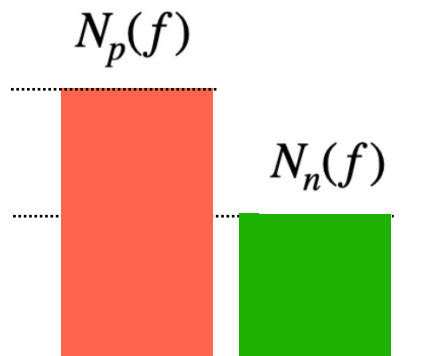


$$N_p(f) \gg N_n(f)$$

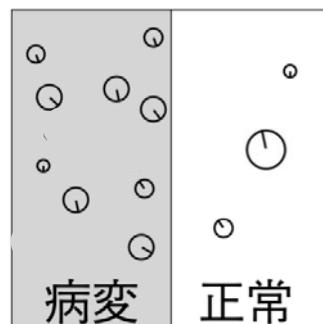
f^+



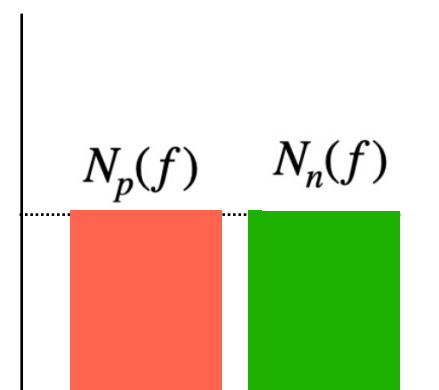
弱い「証拠」



$$N_p(f) > N_n(f)$$

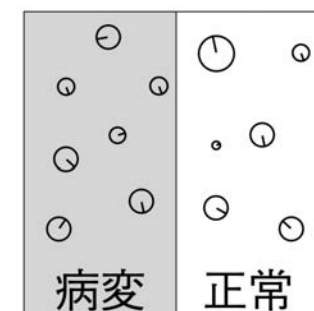


ゴミ



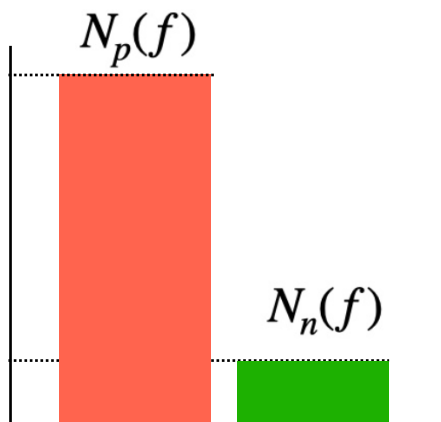
$$N_p(f) = N_n(f)$$

f^0



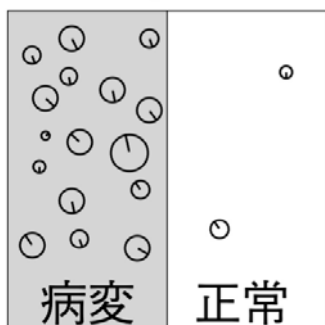
2種類の強い「証拠」

がんの強い「証拠」

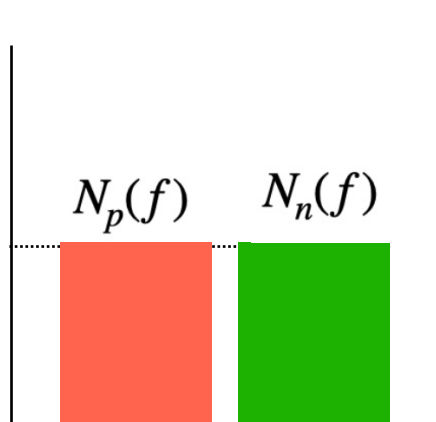


$$N_p(f) \gg N_n(f)$$

f^+

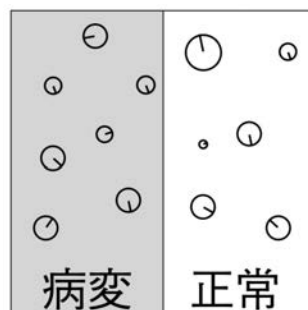


ゴミ

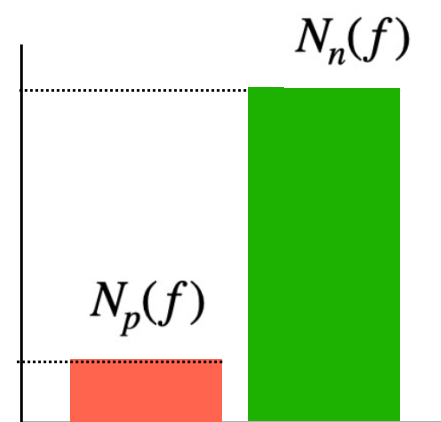


$$N_p(f) = N_n(f)$$

f^0

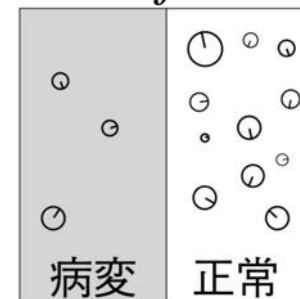


正常の強い「証拠」



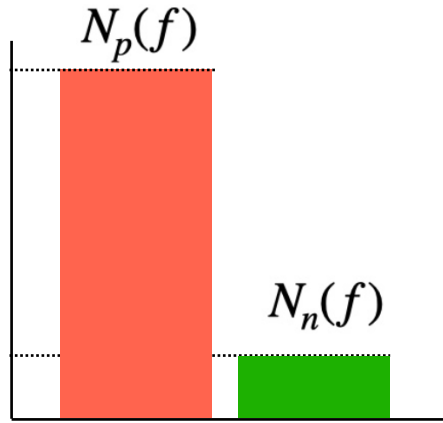
$$N_p(f) \ll N_n(f)$$

f^-



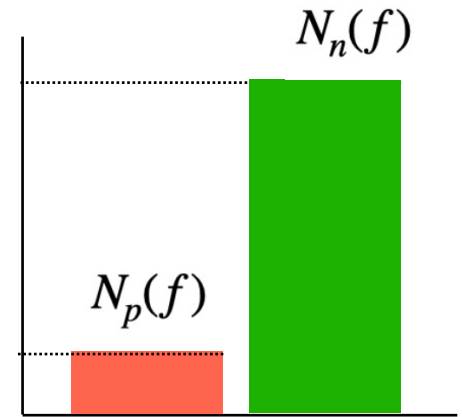
強い「証拠」で色付けする

がんの強い「証拠」



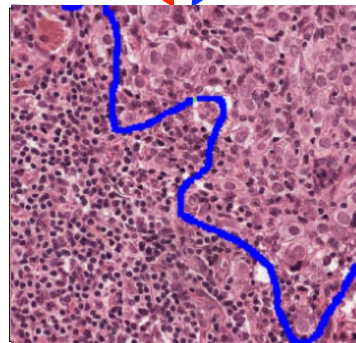
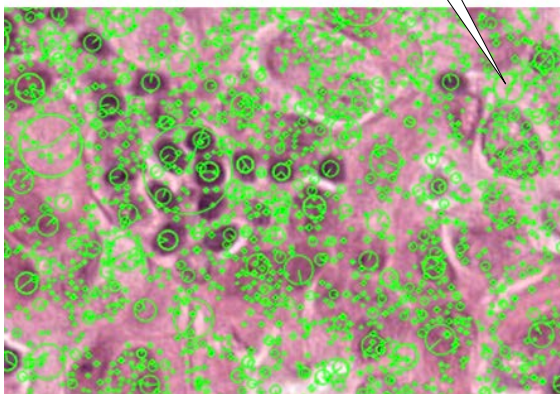
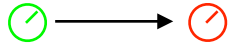
赤にする

正常の強い「証拠」



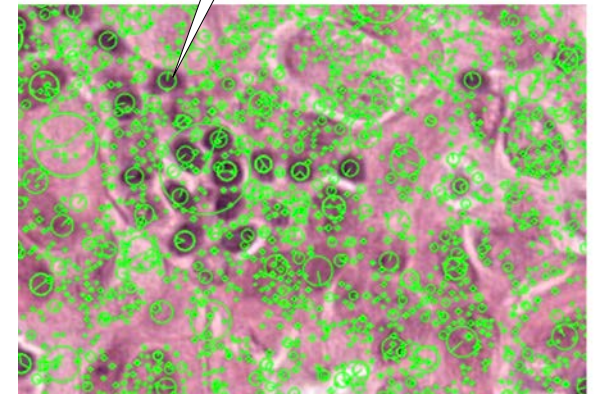
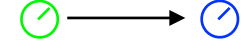
青にする

赤にする



CAMELYON16, <https://camelyon16.grand-challenge.org/>

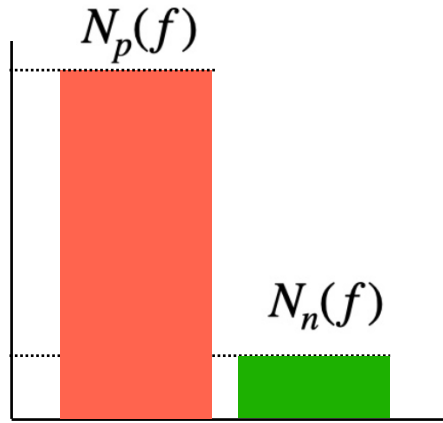
青にする



CAMELYON16, <https://camelyon16.grand-challenge.org/>

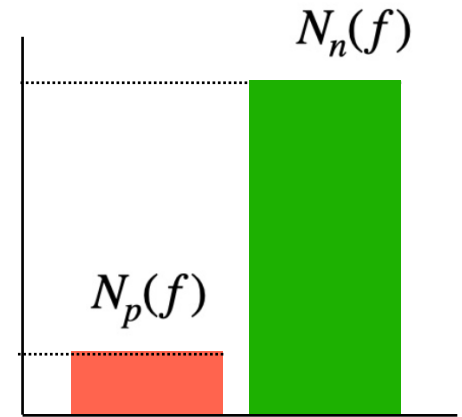
強い「証拠」で色付けする

がんの強い「証拠」



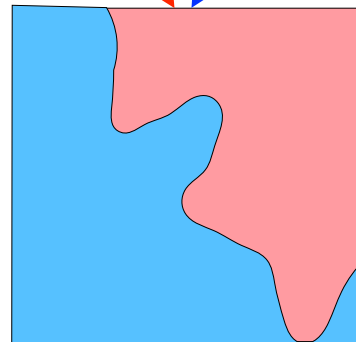
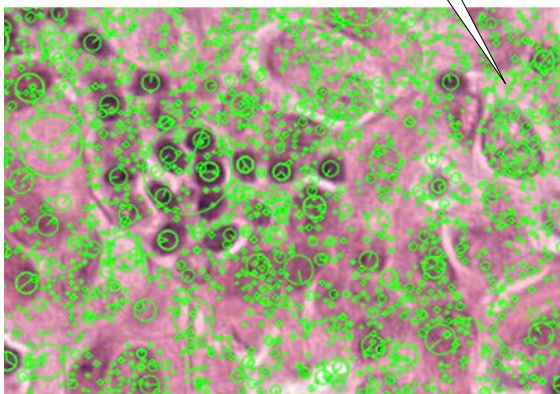
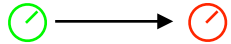
赤にする

正常の強い「証拠」

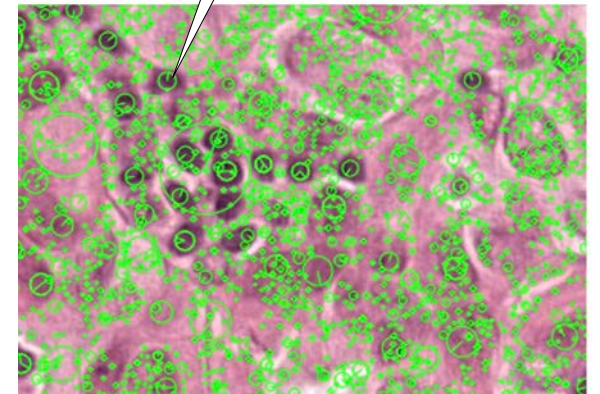
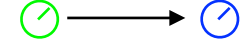


青にする

赤にする

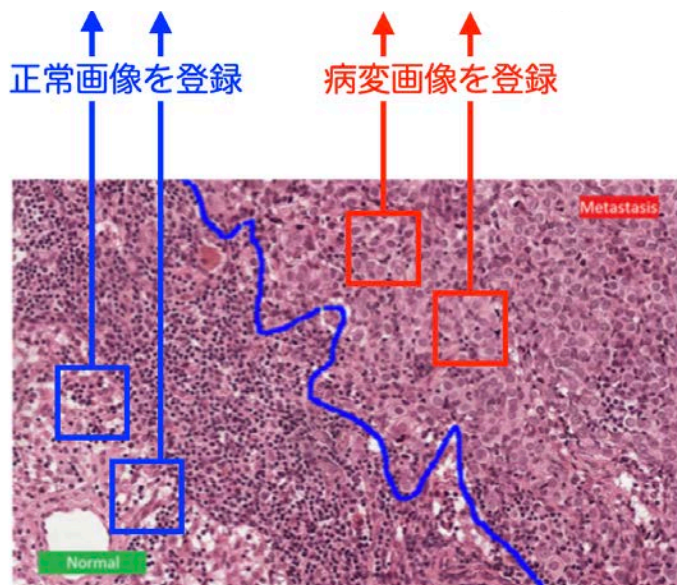
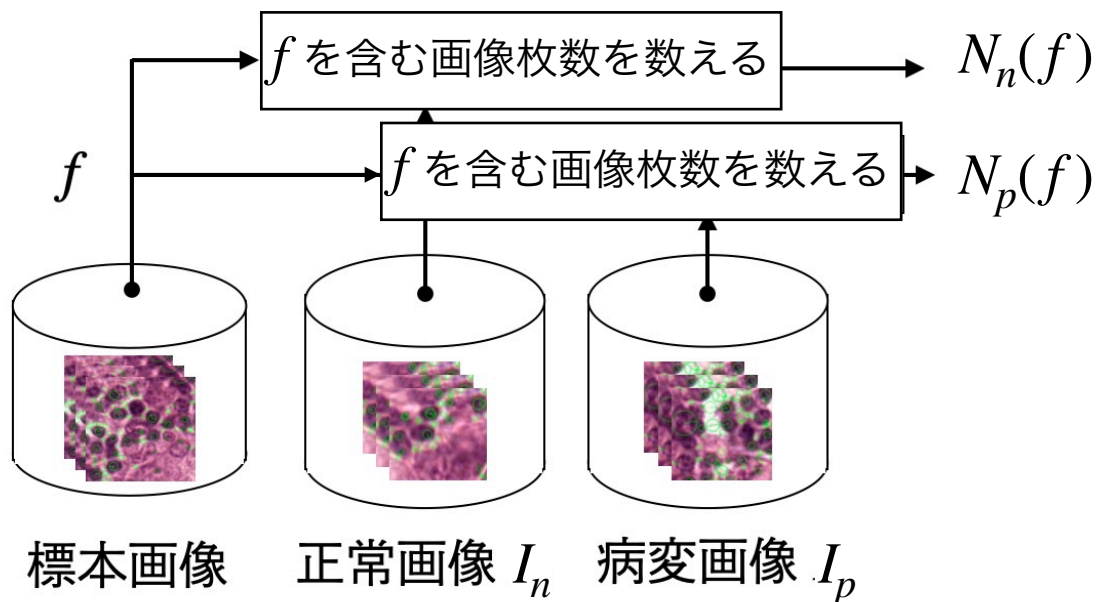


青にする



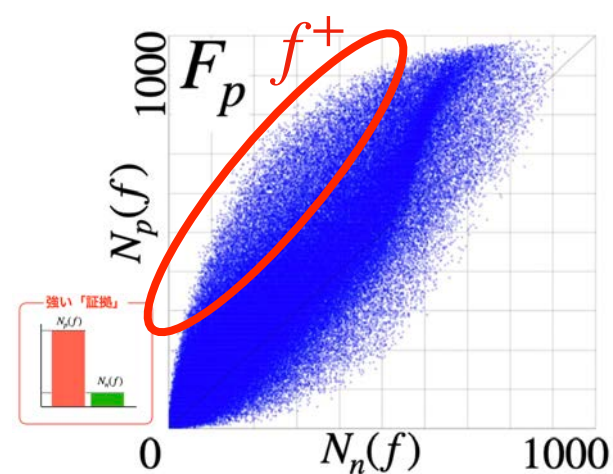
この方法で
証拠の多さを表現した
こんな画像が得られるはず

証拠候補の絞り込み

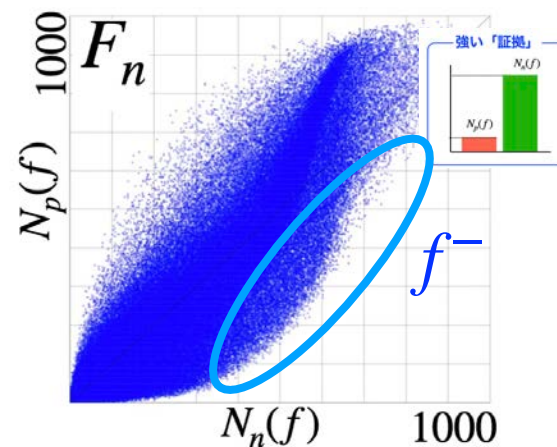


<https://camelyon16.grand-challenge.org/Data/>

標本画像ががんの場合



標本画像が正常の場合



f^+ f^- の「証拠」としての強さ

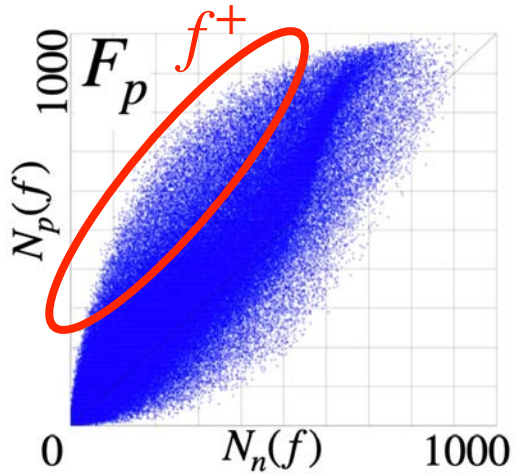
識別情報量

$$C(f, f^0) = P(f) \log \frac{P(f)}{P(f^0)} - (1 - P(f)) \log \frac{1 - P(f)}{1 - P(f^0)}$$

ガンらしさ

正常らしさ

標本画像ががんの場合

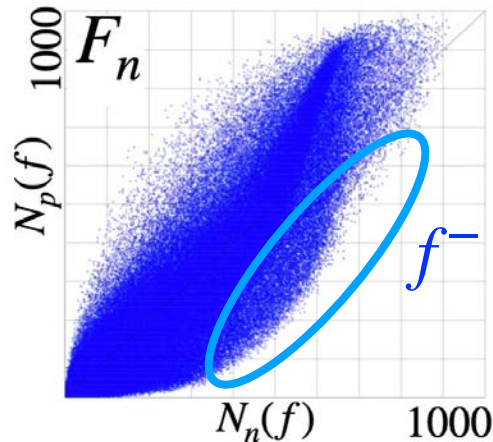


ガン特徴 $C(f^+, f^0) \gg 0$

中立特徴 $C(f^0, f^0) = 0$

正常特徴 $C(f^-, f^0) \ll 0$

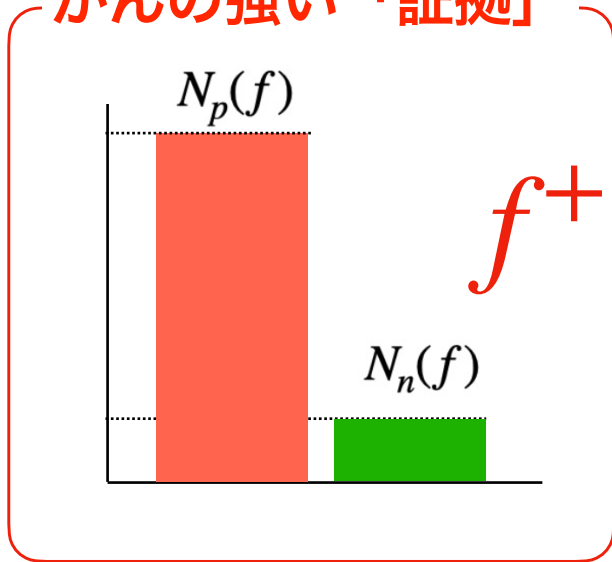
標本画像が正常の場合



強い証拠 $P(f) = \frac{N_p(f)}{N_p(f) + N_n(f)}$
~ 1 or 0

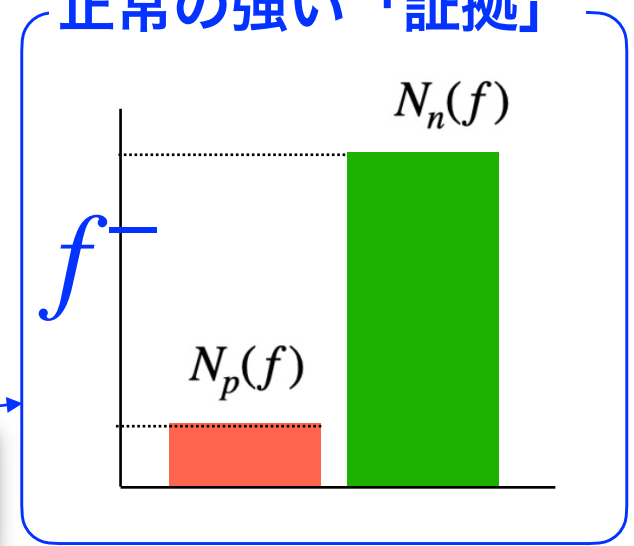
強い「証拠」で色付けする

がんの強い「証拠」



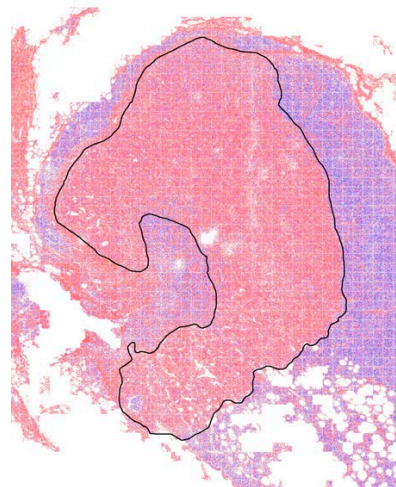
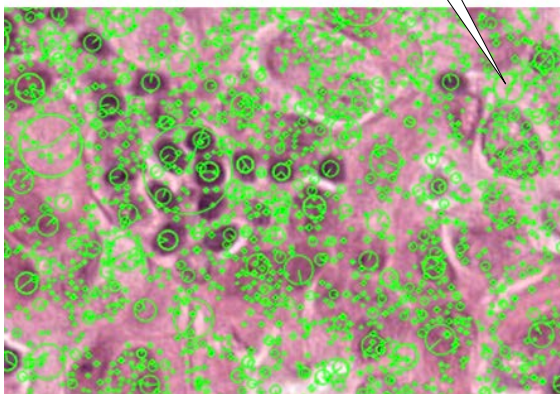
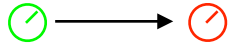
がんの所だけに
出てくる証拠を
赤にする

正常の強い「証拠」

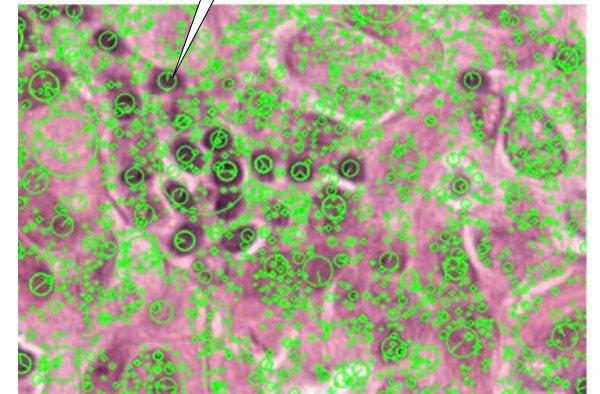
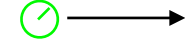


正常の所だけに
出てくる証拠を
青にする

赤にする



青にする

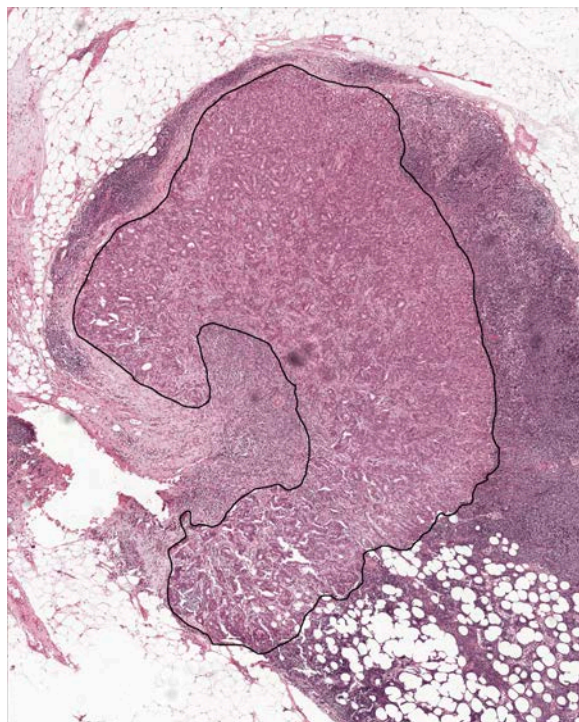


新技術の現在位置

新技術の現状

正確な境界線データは不要

入力画像

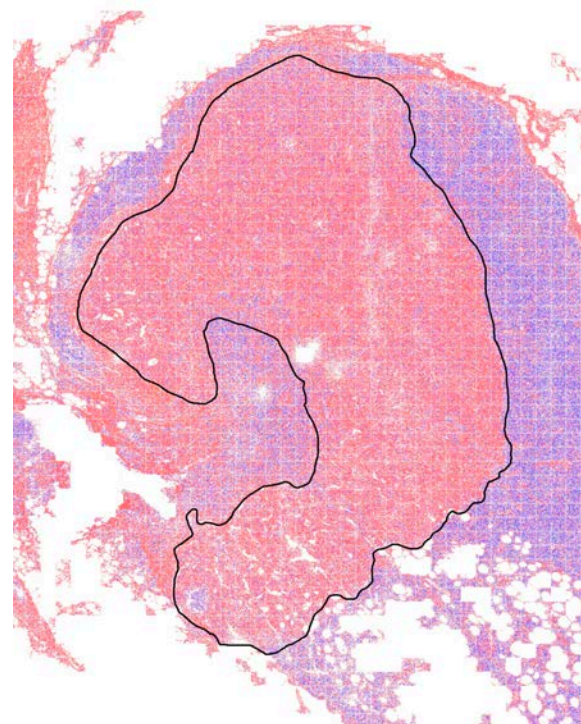


バーチャルスライド (医用画像)

<https://camelyon16.grand-challenge.org/Data/>



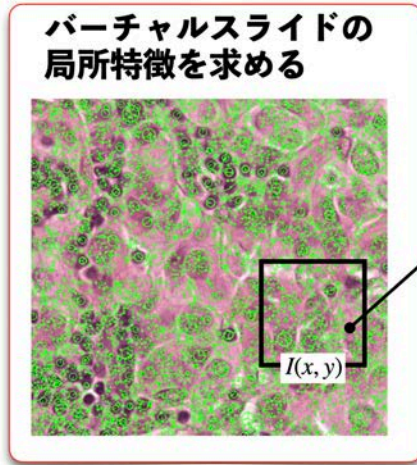
新技術で、識別根拠となる
特徴量分布を可視化した画像



「がん」の情報を持った特徴量の分布

「正常」の情報を持った特徴量の分布

新技術のまとめ：情報密度法（VOID）



情報密度法（VOID）

この範囲のガンらしさを、ここに含まれる
特徴の識別情報量の和で評価する

$$I(x, y) = \sum_{i=1}^n C(f_i, f^0)$$

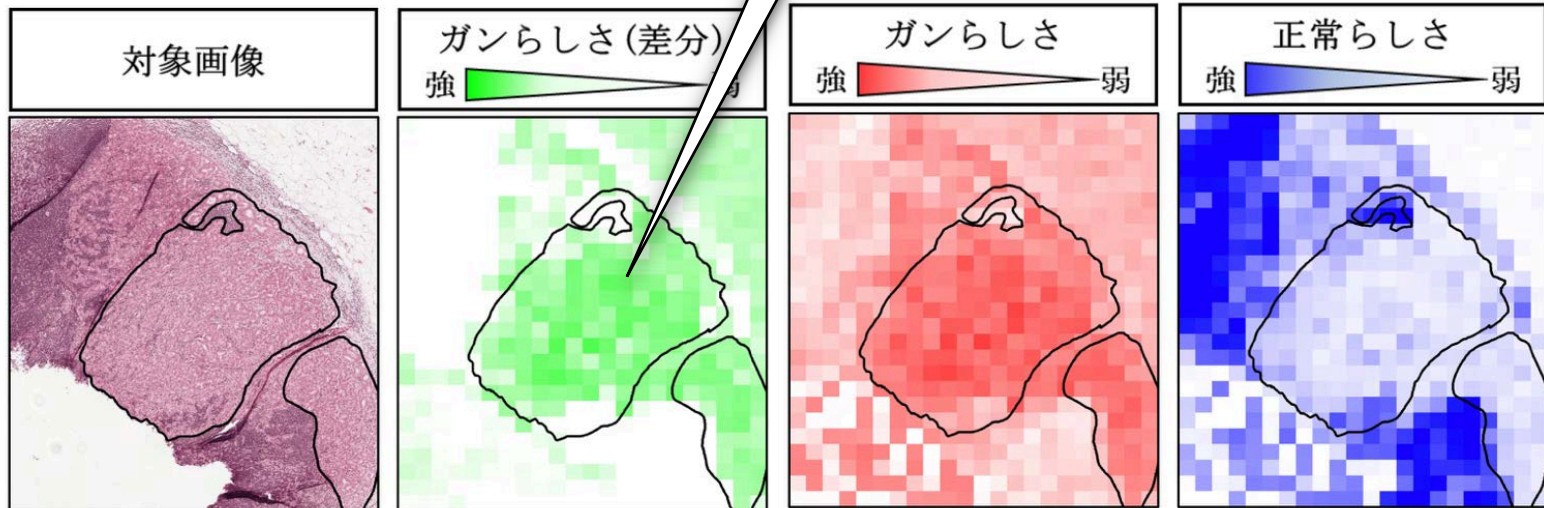
ガンらしさ

$$I^+(x, y) = \sum_{i=1}^n C^+(f_i, f^0), \quad C^+(f_i, f^0) = P(f_i) \log \frac{P(f_i)}{P(f^0)}$$

正常らしさ

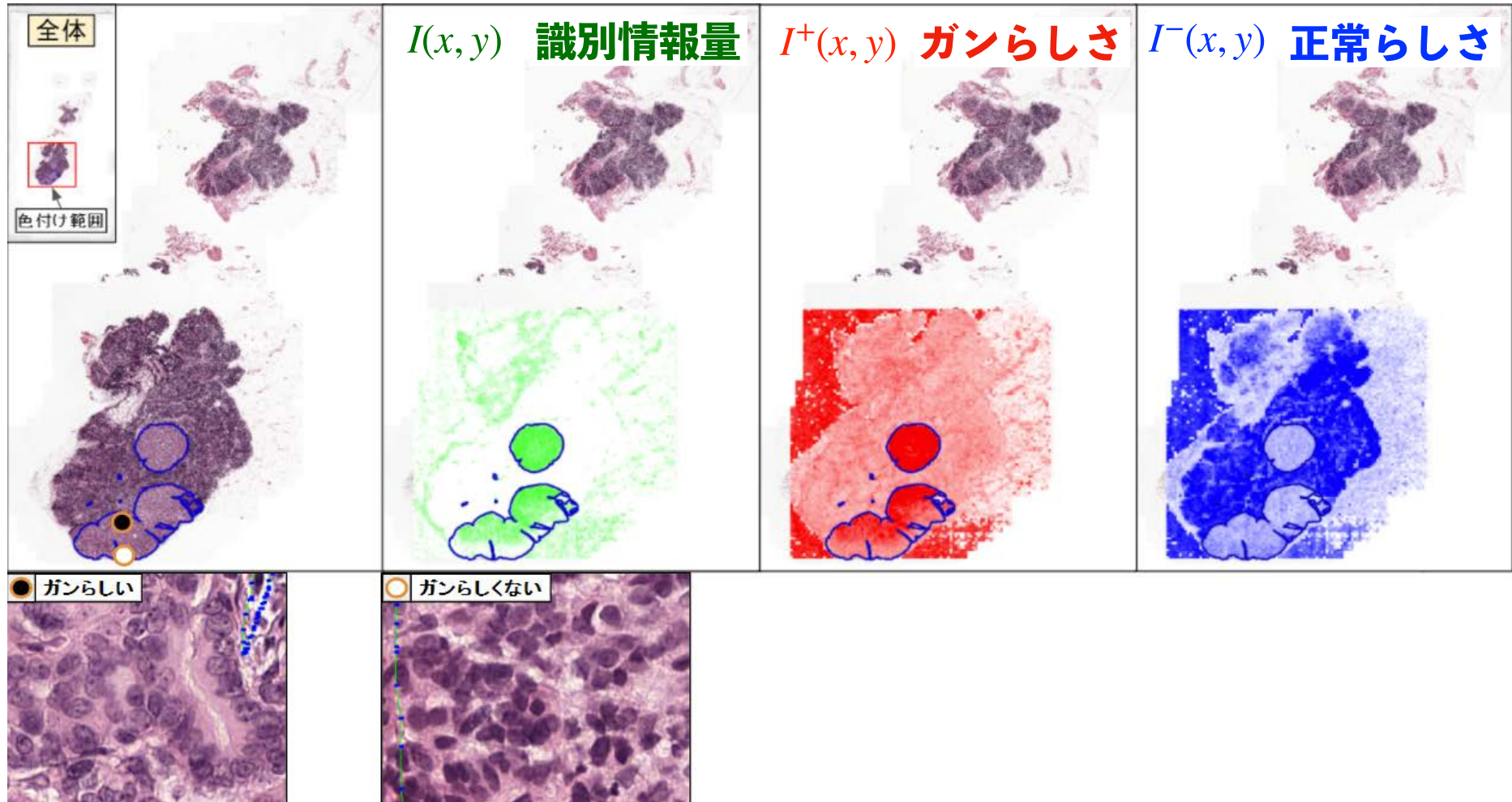
$$I^-(x, y) = \sum_{i=1}^n C^-(f_i, f^0), \quad C^-(f_i, f^0) = (1 - P(f_i)) \log \frac{1 - P(f_i)}{1 - P(f^0)}$$

判断根拠の定量化：
ガンと判断すべき
根拠が沢山ある



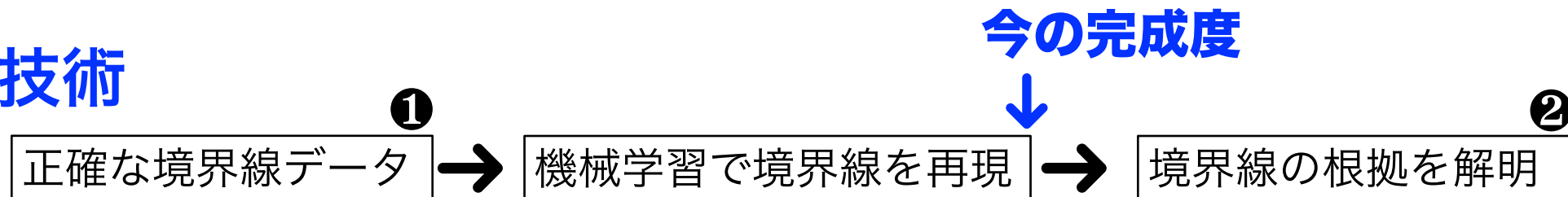
標本の広範囲を見た結果

広範囲を見ても、良い結果である。

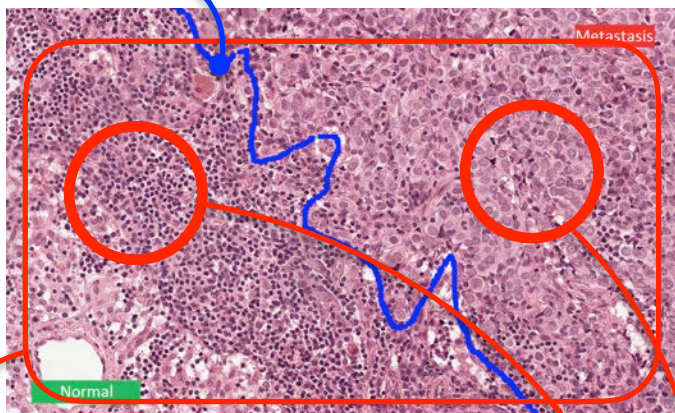


新技術の特徴と従来技術との比較

従来技術



①②が開発のボトルネック



<https://camelyon16.grand-challenge.org/Data/>

①②のボトルネックを解消

新技術



正確な境界線データは不要

想定される用途

- 現ターゲットは、病理画像診断支援システム
 - 技術的には、対象は病理画像に限定されない
- 画像で良否を判断するシステムで下記の要件がある場合に特に適合
 - アノテーション（教師信号作成）の負荷が大きい
 - 診断/判断根拠が強く求められる

実用化に向けた課題

- 現在、大学で性能研究をするためのコードはあるが、実用に耐えるコードはない
- 証拠の抽出、証拠の分布の可視化、境界線の抽出など基本技術については、出願済み
- その他の性能向上のアイデアは多数あるが、大学研究室のマンパワーの限界もあり、未着手

【参考】今年度より、金沢医科大学と共同研究を実施中。
現状、特許は金沢工業大学単独出願のみ。

企業への期待

- ・ 具体的事例を想定して、データを共有しながら現状の実用に耐えるコード&システムを共に開発して下さる企業との連携を希望
- ・ その他の性能向上のアイデアは多数あるので、大学研究室のマンパワーの限界を補い、共同研究をして下さる企業との連携を希望
- ・ 出願済の特許のライセンス供与の相談にも応じます
 - ・ 非排他的使用权（自社実施用のサブライセンス権付）の許諾が原則
 - ・ 一括支払を原則（ランニングロイヤリティは特段希望しない）

本技術に関する知的財産権

- **発明の名称** : 生産装置、画像処理装置、生産方法、
画像処理方法、およびプログラム
- **公開番号** : 特開2022-112542
- **出願人** : 学校法人金沢工業大学
- **発明者** : 金道敏樹、田島和正、田村一希

出願済、未公開の関連特許

特願2022-116962

発明の名称：生産装置、画像処理装置、生産方法、画像処理方法、およびプログラム

産学連携の経歴

- 2003年-2019年 トヨタ自動車（株）の自動運転技術リーダーとして、国内外の多数の大学、研究機関及び企業と共同研究実施
- 2020年-2021年 金沢工業大学教授として、某企業と市販商品の技術課題を解決する共同研究を実施

共同研究における企業と大学の双方の立場の経験から、現実的で妥当な連携契約を提示する用意があります。

パナソニック（株）、トヨタ自動車（株）において、200件（含む共願）を超える特許出願、中間処理対応を経験しており、企業の知財業務にも精通していますので、共同研究の際には成果のスムーズな権利化に協力できます。

お問い合わせ先

金沢工業大学

産学連携局 産学連携東京分室 担当/高田

TEL 03-5777-1964

FAX 03-5777-1965

e-mail iuctky@mlist.Kanazawa-it.ac.jp