

畳み込みニューラルネットワークを用いた OCT 画像内におけるステント領域の検出 (2018年2月14日版)

森本 宜樹^{†1,a)} 呉 海元^{†1,b)} 赤坂 隆史^{†2} 久保 隆史^{†2}

概要：本論文では循環器 OCT 画像内におけるステント領域を自動検出する方法を提案する。光干渉断層法 (optical coherence tomography: OCT) は、光工学を応用した血管内イメージング技術であり、 $10\sim 20\mu m$ という高解像度の画像が得れる。循環器 OCT 画像内には血管内壁の他に治療に用いられるステントと呼ばれる器具が存在し、ステントと血管壁が反応することで内部に新生内膜と呼ばれる膜が生成されるが、この膜が過剰生成されると再狭窄を起こす危険性があるため、ステントの内側の新生内膜の有無を確認することは、治療において重要な診断材料となりうる。しかし 1 人の患者から得られる循環器 OCT 画像は膨大であり、医師はそれらを全て目視によって検査しなければならないため、すべての異常を確認するのは困難である。そこで本論文では、畳み込みニューラルネットワークを用いて OCT 画像内のステント領域を自動的に検出すると同時に、ステント領域内の新生内膜有無を分類する。

キーワード：循環器 OCT 画像，ステント領域検出，ディープラーニング，畳み込みニューラルネットワーク

Stent area detection in OCT image using convolutional neural network (version 2018/02/14)

MORIMOTO YOSHIKI^{†1,a)} WU HAIYUAN^{†1,b)} AKASAKA TAKASHI^{†2} KUBO TAKASHI^{†2}

1. はじめに

1.1 研究背景

近年、医療分野において画像撮影技術の発達が目覚ましく、診断においても画像を用いた高度な診断が可能となっている [1]。光干渉断層法 (optical coherence tomography: OCT) は、光工学を応用したイメージング技術であり、 $10\sim 20\mu m$ という高解像度の画像が得れる。循環器 OCT 画像上には血管壁の他に治療に用いられるステントと呼ばれる医療器具が存在する。ステントとは経皮的冠動脈形成

術 (Percutaneous Coronary Intervention: PCI) 治療の際に使用される網目状のワイヤーである。ステントと血管壁が反応し内部に新生内膜と呼ばれる膜が生成されるが、この膜が過剰生成されると再狭窄を起こす危険性があるため、ステントの内側の新生内膜の有無を確認することは、重要な診断材料となりうる。しかし 1 人の患者から得られる循環器 OCT 画像は膨大であり、現在では専門家の目視によってそれら全ての異常を検査しなければならない。そのため、すべての異常を目視のみで確認することは困難であり、見落としが発生してしまう可能性がある。したがって、循環器 OCT 画像上のステントをシステムにより自動で検出することで、見落としの発生や専門家の負担が軽減されることが期待される。なお、以降本論文では循環器 OCT 画像を OCT 画像と呼称する。

^{†1} 現在，和歌山大学
Presently with Wakayama University

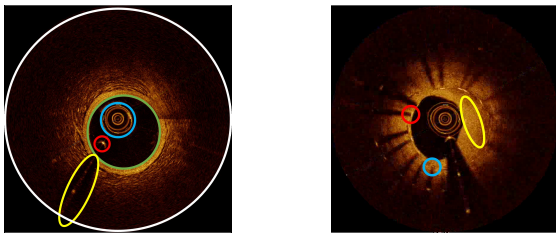
^{†2} 現在，和歌山県立医科大学
Presently with Wakayama Medical University

a) morimoto.yoshiki@g.wakayama-u.jp

b) wuhy@center.wakayama-u.ac.jp

1.2 OCT 画像の特徴

図 1(a) に OCT によって撮影された正常な血管内の画像例を挙げる。



(a) OCT による血管内画像 ([2] 図 2.1.2 を改変) (b) OCT 画像におけるステント

図 1: OCT 画像の特徴の例

白色の丸で囲まれた範囲が OCT によって撮影された範囲である。外周付近にほとんど何も写っていないのは、内側の血管内膜、中膜、外膜がある程度近赤外線を反射しているからである。青色の丸で囲まれた範囲はカテーテルであり、このように写るのは撮影レンズがカテーテル内に存在するためファイバーやカテーテルで近赤外線が反射した結果である。赤色の丸で囲まれているのはガイドワイヤーである。ガイドワイヤーは金属製の為、近赤外線光を全て反射しており周辺よりも輝度値が高くなっている。緑色の丸は血管壁と内腔の境界を示している。黄色の丸で囲まれている範囲は、近赤外線光がガイドワイヤーによって反射した結果できる影である。影の中に幾つかの点や線が写り込んでいるが、これはノイズによる影響である。

図 1(b) にステントと血管内画像にステントが写り込んでいる例を挙げる。ステントとは網目状の医療器具であり、血管を内側から押し広げるために使用される。赤色の丸で囲まれている周辺よりも高輝度値の高い部分がステントである。周辺に比べて輝度値が高いのはガイドワイヤーと同様にステントが金属製であり近赤外線光で反射して減衰しているからである。またガイドワイヤーと同じくステントの外周方向には影が写り込む傾向がある。青色の丸で囲まれた範囲は本来ステントがあるため輝度値が高くなる筈だが周辺と同程度の輝度値となっておりステントの検出が困難である。これは目視での診断の際に見落としの一因となる。最後に黄色の丸で囲まれた範囲は新生内皮である。

1.3 関連研究

Wang A ら [3] は撮影レンズから各角度方向の OCT 画像を帯状に変換し、周辺輝度値とステントによる影の有無を用いてステントの自動検出を行っている (図 2)。Hong L ら [4] は各角度毎の輝度値の合計を用いて特徴を抽出し、決定木を用いたステントの自動検出を行っている (図 3)。津田ら [2] は分離度フィルタを用いたステント候補範囲検出

及びステント点自動識別法を提案している。しかしこの手法では分離度フィルタのしきい値に依存した候補範囲検出を行っており、本来候補範囲として検出されるべき範囲が検出できず、ステント点の識別が行えないといった課題が残っている (図 4)。

これらの関連研究では各角度方向の線上におけるステントの輝度値とステントによる影の関係に依存しており、影が検出できない場合や図 1(b) の青丸のようにステントの輝度値が周辺輝度値と区別がつかない場合にはステント検出が不可能である。また、これらの特徴はステント点内側の新生内膜の有無に対する考慮はされておらず、また石灰化、脂質性プラーク等の他の特徴に対しては別の特徴モデルの構築が必要となる。

そこで本論文では線上の輝度値等の関係性ではなく、ステント周辺領域から特徴を抽出する。方向性としては津田らの手法と同じだが、津田らは 9 次元の点特徴であるのに対し、本論文ではステント周辺領域全体から自動で特徴を抽出する。これによりステントの輝度値や影の影響を抑えた特徴が抽出できると考えられる。また同時にステント周辺の新生内膜の有無も判別する。そのためのツールとして深層学習 (ディープラーニング) を用いる。

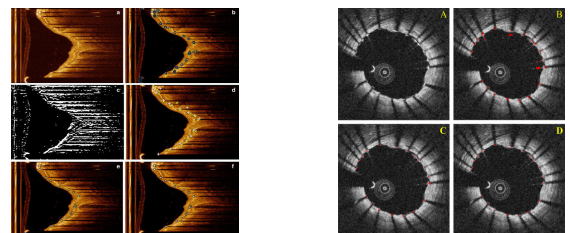


図 2: Wang A らによる自動識別 (Wang A[3] の図 4 を引用) 図 3: Hong L らによる自動識別 (Hong L[4] の図 5 を引用)

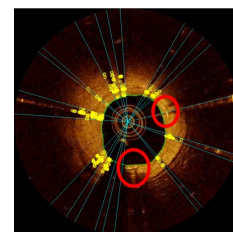


図 4: 候補点として選出されないステント [2]

2. 深層学習を用いた画像分類

深層学習 (ディープラーニング) とは、多層のニューラルネットワークによる機械学習手法である。近年研究が進み、学習に必要な計算機的能力向上、大量のデータを利用しやすくなった等の背景から注目を集めており、音声、画像、自然言語を対象とする問題に対し従来の手法より高い性能を発揮している。

2.1 畳み込みニューラルネットワーク

畳み込みニューラルネットワーク (convolutional neural network: CNN) とは格子状のトポロジーを持つデータの処理に使われる特殊なニューラルネットワークである [5]。畳み込み層 (convolution 層), pooling 層等の特徴抽出部と全結合層 (ニューラルネットワーク) の識別部から構成されており, 出力と教示データとの誤差から最適となる重みを自動的に計算し逆伝播により更新する。また最近では Batch Normalization [6] を各層に用いることで学習を効率的に進めることができる。

2.2 代表的な画像分類モデル

ここでは2つの代表的な畳み込みニューラルネットワークモデルについて述べる。本論文では実験用モデルを構築する際に, これらのモデルを参考にしている。

2.2.1 AlexNet

AlexNet [7] はトロント大学の研究グループ (Alex Krizhevsky, Geoffrey E. Hinton) が 2012 年の ILSVRC (ImageNet Large Scale Visual Recognition Challenge) において提案した畳み込み 5 層と全結合 3 層からなる畳み込みニューラルネットワークである。AlexNet では既存のディープラーニングで広く採用されていた手法ではなく, 後続の畳み込みニューラルネットワークでも広く採用されることになる技術を新たに搭載した。

- 活性化関数: ReLU
- Data augmentation (データ拡張)
- Dropout 法

データ拡張については次章で説明する。Dropout 法は主に全結合層に用いられ, ノードの接続をランダムに選出し, その重みを 0 にすることで過学習を抑制する手法である。

2.2.2 VGG16

VGG16 [8] はオックスフォード大学の研究グループ (Visual Geometry Group^{*1}) が 2014 年の ILSVRC において提案した畳み込み 13 層と全結合 3 層の計 16 層から成る畳み込みニューラルネットワーク。畳み込みニューラルネットワークにおける層の深さについて追求しており, 各層の構造は以下のようにシンプルな構造である。ImageNet と呼ばれる大規模な画像データセットを使って訓練したモデルが公開されている。

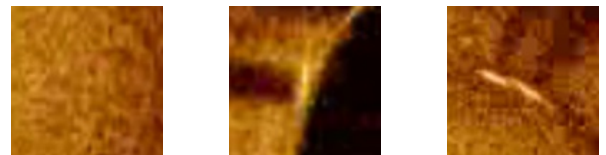
- 畳み込み層
フィルタサイズ: 3×3 , スライド: 1 で全て固定。
- Pooling 層
フィルタサイズ: 2×2 , スライド: 2 で全て固定。複数 (2or3) の畳み込み層後に使用。
- 全結合層
全結合上 2 層のノード数: 4096 出力層: 1000 (分類クラス数)

ラス数) dropout 率は 0.5

3. 畳み込みニューラルネットワークを用いたステント検出の為にモデル構築と評価

3.1 データセット

訓練画像, 評価画像として用いるデータセットについて記述する。各画像は OCT 画像を複数の正方局所領域 (辺長 $n=32$ pixel) に分割することにより生成する。また各画像には学習時と評価時に用いる教師データとなるラベルを付与する必要がある。図 5 にデータセット画像例を示す。図 5(a) のように, 領域画像内にステントが写っていない場合は非ステント領域とみなしラベル "0" を付与する。図 5(b) と (c) には領域画像内にステントが写っており, ステントの内側に新生内幕が写っていない場合はラベル "1", 新生内幕が写っていればラベル "2" を付与する。本研究では前処理として, 全訓練画像の各ピクセル, チャンネル毎に画素値の平均を計算し, 全データセットから差し引く Per-pixel Mean Subtraction を用いている。また平均を引いた後に各要素を 255 で割ることでデータのスケールリングを行っている。また各画像を左右反転させることによってデータセット内の画像数を擬似的に増やしている。この処理により訓練データ 14,984 枚, テストデータ 3,746 枚となった。



(a) 非ステント (b) 新生内幕なし (c) 新生内幕あり

図 5: データセット画像例

3.2 実験用モデル構築

前章で紹介した AlexNet, VGG16 を参考に 3 つのモデルを構築した。実装には Chainer^{*2} を用いている。以下に各モデルの構造について述べる (表 1)。

model1

AlexNet の畳み込み層, 全結合層等の層の数やフィルタサイズ, フィルタ数等のパラメータを参考にして構築。ただし, データセットとして使用している画像サイズが異なるため出力は異なる。

model2

VGG16 の畳み込み層, 全結合層等の層の数やフィルタサイズ, フィルタ数等のパラメータを参考にして構築。入力画像のサイズが小さく圧縮しすぎると特徴の消失が起こる可能性があるため, pooling 層の数を減少。

^{*1} Department of Engineering Science, University of Oxford

^{*2} 株式会社 Preferred Networks が開発・公開している深層学習用フレームワーク

model3

上記 2 つの構築を元に、畳み込み層等の数、パラメータを微調整。データサイズが 32 × 32 と小さく、複雑な輪郭構造を持たないため全結合層のノード数を減少し分解能を圧縮。ステント画像におけるステント以外の要素（ステントの影、新生内膜）が重要であるため、畳み込み層の上層ではフィルタサイズ、パディングサイズを大きめに調整。

表 1: 各階層の設定値

	model1	model2	model3
畳み込み層	1層(Size:7,Out:96,pad:3)	2層(Size:3,Out:64,pad:1) (Size:3,Out:64,pad:1)	2層(Size:7,Out:32,pad:4) (Size:7,Out:32,pad:4)
Pooling	Size:2,Stride:2	Size:2,Stride:2	Size:2,Stride:2
畳み込み層	1層(Size:3,Out:128,pad:1)	2層(Size:3,Out:128,pad:1) (Size:3,Out:128,pad:1)	3層(Size:3,Out:64,pad:1) (Size:3,Out:64,pad:1) (Size:3,Out:64,pad:1)
Pooling	Size:2,Stride:2	Size:2,Stride:2	Size:2,Stride:2
畳み込み層	3層(Size:3,Out:192,pad:1) (Size:3,Out:192,pad:1) (Size:3,Out:128,pad:1)	4層(Size:3,Out:256,pad:1) (Size:3,Out:256,pad:1) (Size:3,Out:256,pad:1) (Size:3,Out:256,pad:1)	3層(Size:3,Out:128,pad:1) (Size:3,Out:128,pad:1) (Size:3,Out:128,pad:1)
Pooling	Size:2,Stride:2	Size:2,Stride:2	Size:2,Stride:2
畳み込み層	x	5層(Size:3,Out:512,pad:1) (Size:3,Out:512,pad:1) (Size:3,Out:512,pad:1) (Size:3,Out:512,pad:1) (Size:3,Out:512,pad:1)	x
pooling	x	Size:2,Stride:2	x
全結合層	1024	1024	128
全結合層	1024	1024	128

これらのモデルの出力層以外の各層には Batch Normalization[6] を用いており、全結合層の間には dropout 層を用いていない。Batch Normalization はミニバッチ内の分散を 1、平均を 0 にし各階層の入力分布の偏りを無くするための正規化手法。過学習を抑制し、学習の収束速度が上がるということが知られている。

3.3 学習と評価結果

学習時の最適化手法は Momentum SGD[9] を用いる。各モデルのテストデータに対する結果を図 6 と図 7、表 2 と表 3 に示す。学習時のバッチサイズは 100、学習回数は 50 回とした。各学習 (epoch) 毎に損失を求めると同時に評価を行っている。また比較検討用として chainer に実装されている VGG16 モデルを使用して実験する。このモデルには Batch Normalization は用いられておらず、全結合層の間に dropout 層が使用されているため、今回の実験では各層に Batch Normalization を適用させ、dropout 層は用いないこととする。

表 2: 実験結果参考値 (損失)

	model1	model2	model3	VGG16
平均	0.069	0.070	0.059	0.069
分散	0.056×10^{-4}	0.146×10^{-4}	0.044×10^{-4}	0.571×10^{-4}
最小	0.064	0.064	0.055	0.059

3 つのモデルの損失の結果について考察する。表 2 に示す参考値から、全体的に類似した結果となっているが、model3 が他の 3 つのモデルと比較して損失の値が抑えられ

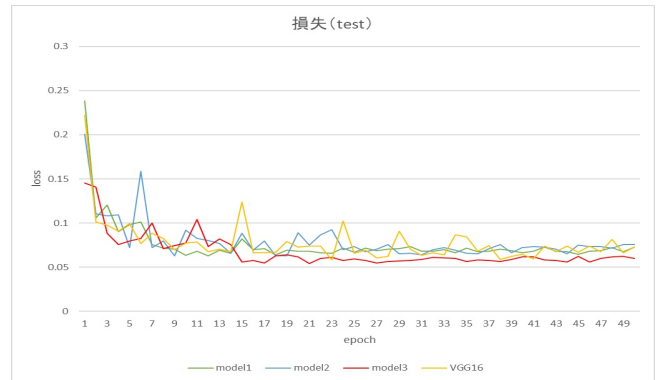


図 6: 実験結果 (損失)

ている事がわかる。これは今回のデータはサイズが 32x32 と小さく、分類クラスも 3 クラスしかないため、全結合層のノード数が model1 と model2 が 1024 に対して model3 が 128 と少ないことにより損失が抑えられていると考えることができる。次に分散の値に関しては model1, model3 は model2, VGG16 よりも低くなっている。model1 は畳み込み層の数が 5、model3 は 8 であり、model2, VGG16 の 13 よりも少なく設計していることが影響していると考えられる。model2 と VGG16 に関しては、ほとんど同じ結果が得られたが分散の値が既存の VGG16 よりも分散の値が低くなった。これはモデル構築のときに述べたとおり、pooling 層を減らすことで入力画像の過度の圧縮を防いだためと考えられる。

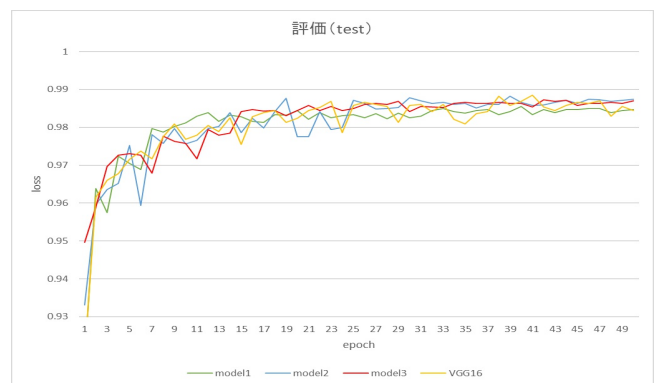


図 7: 実験結果 (評価)

表 3: 実験結果参考値 (評価)

	model1	model2	model3	VGG16
平均	0.984	0.986	0.986	0.985
分散	0.055×10^{-5}	0.054×10^{-5}	0.054×10^{-5}	0.346×10^{-5}
最大	0.985	0.988	0.987	0.988

3 つのモデルの評価の結果について考察する。表 3 に示す結果も概ね損失における考察と同じだが、model2, model3 に関しては評価における参考値がほとんど同じであるが、損失における参考値が model3 の方が全体的に低くなって

いる．損失は出力を確率に計算し直してから教示データとの交差エントロピー誤差を求めているため，model3の方がより正解に近い分類ができていているということがわかる．

表2と表3の結果から今回実験用に構築したモデルは既存のモデルより分散の値が低いため，既存のモデルをそのまま利用するよりも安定した学習が行えているということがわかった．また実験用モデルの中でもmodel3が最も良い結果が得られた．これにより，今回のデータはサイズが 32×32 と小さく，また複雑な輪郭構造持たないためmodel2ほどの深さは必要なくmodel3が適切であったということが推察できる．

4. 実験

4.1 未学習画像を用いたステント領域の検出

学習後のモデル(model3)を用いて，前章で述べたデータセットに用いていないOCT画像に対してステント領域の検出を行った．OCT画像を 32×32 の領域に切り取り，切り取り画像の出力結果がステント領域なら，OCT画像のその領域を色枠で囲む．青もしくは赤の枠で囲まれているならステント領域であり，青は新生内幕なし，赤は新生内幕ありの領域である．検出は，OCT画像上を全探索して行うが，全ての切り取り画像を識別すると時間がかかるため，各画像毎にノルムを求め一定値以上の画像だけ識別を行う．

4.1.1 検出結果

図8はステント領域を検出する実験の結果例である．図8の(a)，(c)は実験結果をわかりやすくするために入力画像のステントとして検出したい領域を色枠で囲んだ図であり，(b)，(d)は実際の検出結果例である．

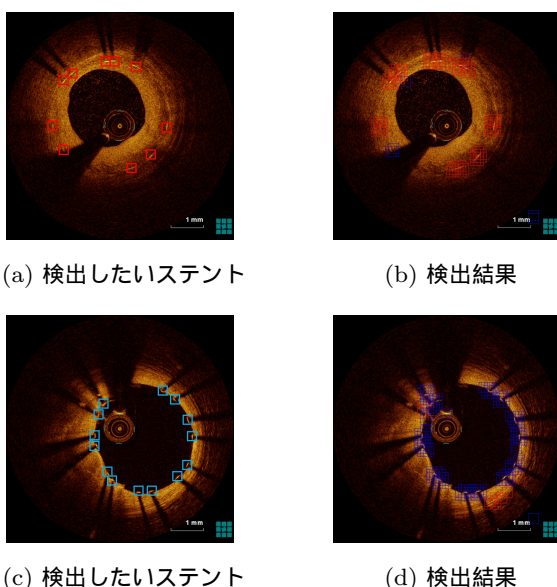


図8: 検出結果例

検出結果からステントによる影が顕著でない部分においても検出したいステント領域はおおむね検出できていることがわかる．しかし本来ステントではない領域をステント領域と誤検出している場合や新生内幕がある領域に対し，新生内幕無しの領域として誤識別される領域が複数箇所存在する．誤検出された場所をみると，ガイドワイヤーの周辺が新生内幕無しの領域と識別されている．これはガイドワイヤーの部分の輝度値が高く，その周辺の輝度値は低くなるという特徴が新生内幕無しのステント領域の特徴と極めて似ているためである．次に(d)の画像4時から6時方向にかけて本来ステントではない領域に対して，新生内幕ありのステント領域として検出されている箇所が幾つか見受けられる．これらはノイズによる影響で一時的に輝度値がステント並に高くなっており，外周方向に向けて輝度値が低くなってしまったため誤検出されたと考えられる．しかし誤検出された内側の部分で新生内幕無しの領域識別が正常にできているため，診断時にはこれらの情報は無視できると考えられる．また11時方向の誤識別もノイズによる影響で，1つはステントの影の部分に高輝度値の箇所が発生している点，もう1つはカテーテル部分と繋がっている部分が発生している点が影響しているからである．

前章で行ったモデル学習の際にこのような画像はデータセットとして与えていなかった．そこで今回の検出結果において誤検出・誤識別されたいくつかの領域を新しいデータセットとして構築し，再び学習を行うことで精度向上を目指す．

4.1.2 検出結果を用いた再学習

前節で誤検出・誤識別されたデータを学習データとし，再び学習を行った．再学習には6枚のOCT画像から誤検出・誤識別した領域を再学習用データセットとして学習を行っている．この再学習したモデルを用いた未使用のOCT画像におけるステント領域検出を行った．図9に結果例を示す．

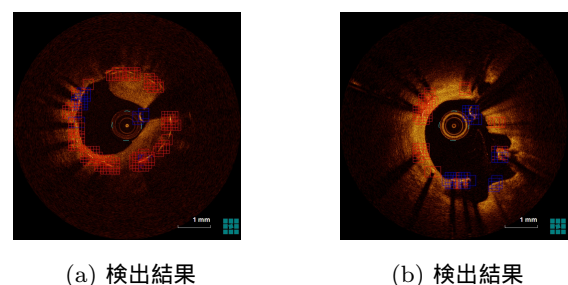


図9: 再学習モデル検出結果例

結果画像(a)のようにOCT画像全体が暗く，ステントによる影がほとんど写っていない箇所においても，ステント領域を検出できている．また(b)のようにノイズがひどく血管壁が歪んでステントの影が途切れている場合でもス

ステントの検出が行えた。しかしガイドワイヤー周辺をみると誤検出された領域が多い。これは畳込みによる特徴抽出がエッジ情報や平行移動に対する不変性を持つが、画像の回転成分に対応しにくいという性質によるもので、再学習に用いた OCT 画像のガイドワイヤーの方向と異なるからだと考えられる。この問題に対してはデータセット構築におけるデータ拡張の際に、入力画像に対してアフィン変換を適応させることで解決できる可能性がある。また (b) はステント無しの領域とステント有りの領域が混在しており、目視しづらい箇所が発生している。この問題については識別判定の際に、周辺の識別結果を反映させる等の対応が必要である。

次に津田ら [2] の手法 (以下関連研究) によるステント検出との比較を行う。図 10 に検出結果を示す。関連研究ではステント候補区間を 2 色の直線で表しており、時計回りに水色 紫色の順に直線が引かれている区間が候補区間である。そして候補区間内の黄色の丸がある位置がステント候補点の場所を示している。結果から関連研究の手法ではステント候補点がほとんど検出できていない、ステントの影が検出できないような画像においても提案手法では検出できており、識別においても概ね可能である事がわかる。しかし図 10(f) を見ると、提案手法でも検出できていないステント領域が存在する。これはモデルの再学習時に新生内膜無しの場合の領域と酷似しているガイドワイヤーの領域を非ステントとして学習した結果だと考えられる。

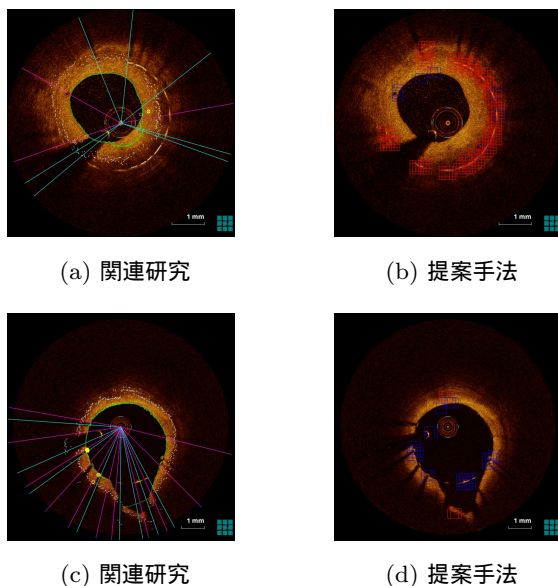


図 10: 関連研究 [2] との比較

5. まとめ

本論文では OCT 画像内のステント領域の検出を目標としており、関連研究の課題を解決するため CNN を用いた

ステント領域検出法を提案し、実験を行った。検出結果から、OCT 画像全体の輝度値が低くステントの影がほとんど見えない場合や、ノイズがひどく影が途切れている場合でも検出したいステント領域は概ね検出できることを確認した。それと同時にステント周辺の新生内膜の有無を分類することが可能である。しかしステントが密集している箇所や、新生内膜の有無が混在している場合に検出結果が視認しづらい場合が確認できた。また、ガイドワイヤー等のステント領域に似た特徴を持つ領域で幾らかの誤検出・誤識別が認められた。そこで、検出結果を用いた再学習を行い、その成果について確認した。ガイドワイヤーと新生内膜が無い場合のステント領域の特徴が酷似していたため、新生内膜が無い場合のステント領域においては必ずしも精度が向上するとは認められなかった。最後に関連研究の手法との比較実験を行い、提案手法の有効性を確認できた。

謝辞

本研究は JSPS 科研費 15K01331 の助成を受けたものです。

参考文献

- [1] 久保隆史, 片山陽介, 江守裕紀, 寺田幸誠, 赤阪隆史. 血管内 oct にかかわる最新の知見. 日本冠疾患学会雑誌, Vol. 23, No. 1, pp. 41–47, 2017.
- [2] 津田嶺雪, 呉海元. Oct 画像からのステントの自動検出 (パターン認識・メディア理解). 電子情報通信学会技術研究報告 = IEICE technical report: 信学技報, Vol. 114, No. 521, pp. 35–40, 2015.
- [3] Ancong Wang, Jeroen Eggermont, Niels Dekker, Hector M Garcia-Garcia, Ravindra Pawar, Johan HC Reiber, and Jouke Dijkstra. Automatic stent strut detection in intravascular optical coherence tomographic pullback runs. 2013.
- [4] Hong L, Madhusudhana Gargasha, Zhao Wang, Daniel Chamie, Guilherme F Attizzani, Tomoaki Kanaya, Soumya Ray, Marco A Costa, Andrew M Rollins, Hiram G Bezerra, et al. Automatic stent detection in intravascular oct images using bagged decision trees. 2012.
- [5] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. <http://www.deeplearningbook.org>.
- [6] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International Conference on Machine Learning*, pp. 448–456, 2015.
- [7] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pp. 1097–1105, 2012.
- [8] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.
- [9] David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. Learning representations by back-propagating errors. 1986.