

2019 年度 卒業論文

論文題目

血管構造に基づく非剛体運動のための
画像マッチング手法

指導教員

舟橋 健司 准教授

名古屋工業大学 工学部 情報工学科
2016 年度入学 28114020 番

名前 江本 峻

目次

第1章	はじめに	1
第2章	背景知識	5
2.1	Neural Network	5
2.2	Convolutional Neural Network	5
2.2.1	畳み込み層 (Convolution Layer)	6
2.2.2	ReLU(Rectified Linear Units)	7
2.2.3	プーリング層 (Poling Layer)	8
2.2.4	全結合層 (Fully Connected Layer)	9
2.2.5	逆畳み込み (Deconvolution)	9
2.3	U-net	10
2.4	特徴ベースマッチング	11
2.5	領域ベースマッチング	12
第3章	提案手法	14
3.1	血管領域抽出	14
3.2	ノイズ除去	16
3.3	モルフォロジー変換	16
3.4	血管構造のグラフ化	17
3.4.1	血管構造の分類方法	17
3.4.2	グラフ抽出アルゴリズム	18
3.5	グラフ構造の補間	20
3.6	グラフの対応付け	21
第4章	実験	23
4.1	シミュレーションによる実験	23
4.2	実画像による実験	27
第5章	むすび	32
	謝辞	33
	参考文献	34

第1章 はじめに

現在の医療現場では、出血、炎症、腫瘍などの異常部発見のために、内視鏡による大腸の診察が行われている。内視鏡とは人体内部を観察することを目的とした医療機器であり、内視鏡本体に光学系を内蔵しており、その先端部分を体内に挿入することによって内部の映像を手元で見ることが可能となるものである。またこれは患者の体内において発見された腫瘍の切除や回収といった処理のための鉗子も備えている。

内視鏡診断で発見される大腸の異常部の一つにポリープがある。異常部とされるポリープが発見された場合、すぐに切除すべきと考えるかもしれないが、実際には内視鏡によって発見されたポリープは、必ずしも全て切除されるべきものではない。大腸で発見されるポリープは大まかに腫瘍性ポリープと非腫瘍性ポリープに大別される。非腫瘍性ポリープには炎症性のポリープや過形成によるポリープなどがあるが、どちらも正常細胞が集まることでイボ状になったものであり、放置しても大腸癌になる可能性は低い。その一方で、腫瘍性ポリープには放置しておくとう癌化する危険な悪性腫瘍とそのまま放置しても癌化しない良性腫瘍がある。発見されたポリープが非腫瘍性ポリープ、または良性腫瘍である場合、悪性腫瘍の場合と同様に切除すると却って患者の負担となってしまう。そのため、診断医師は発見されたポリープの良性、悪性を識別し、それに応じてポリープを切除するか否かを決定する。これはポリープの大きさ、形状、表面のテクスチャを主な判断基準として内視鏡でポリープを観察することで識別を行う。しかし、内視鏡環境が特殊である上に、内視鏡では二次元画像しか得られないため、経験を積んだ診断医師であっても良性、悪性の識別は困難であると言われている。実際に画像中でポリープが発見され、ポリープの大きさが2cm以上である場合、悪性の可能性が高いとされているため、ポリープを切除するのだが、取り出して大きさを確認すると画像内にて判断していた大き

さと異なるということも多い。このような現状から、内視鏡画像からポリープの三次元形状復元を行うことで大きさや形状といった定量的な情報を取得できれば、診断医師の負担を軽減することが可能となり、診断時間の短縮や良性、悪性の識別精度の向上が期待できる。

三次元形状復元を行う手法は大まかに2種類に大別される。一つ目の復元手法は2台以上のカメラを用いて同一物体を撮影しておき、そこから得られる視差によって物体の形状を復元するステレオ法と呼ばれる手法、またはカメラ移動させて取得できる動画フレームを用いてステレオ法により Shape from Motion といった、三角測量の要領で幾何学的に物体の位置を推定する手法である。ステレオ法の欠点として、形状復元の精度が画像間対応点の結果に大きく依存し、多くの対応点が取れない場合に疎な形状復元になってしまう点である。二つ目の復元手法は、画像の陰影情報から物体の形状を復元する SFS(Shape from Shading) と呼ばれる手法であり、これは Horn[1] により提案され、現在でも様々な研究により改良改善がなされている。この手法は、上記で紹介したステレオ法とは異なり、SFS では物体表面の反射特性及び照明条件を既知とした場合において、画像の濃度と形状との関係を表した光学的制約式を三次元形状に関する偏微分方程式を用いて表し、これらから得られた方程式を解くことにより物体の位置を推定している。この光学的な手法では単一画像から形状復元を行うことが可能となるため必然的に画像間対応点を行う必要がなく、密な形状復元が可能である。しかし、SFS の欠点として、単一の画像を用いた場合での形状復元では、画像内の対象物体の大きさがわかる特別な指標が存在しない場合、復元された物体の実際の大きさを求めることが原理的に困難な点である。そこで、内視鏡画像から大きさが既知の物体を参照して、2枚の画像間でのカメラ移動量から、三次元形状復元に用いるパラメータを算出する手法がある。文献 [2] では、ポリープの切除などで用いられる縫合糸の太さを参照し、ポリープの絶対的な大きさを求めている。また文献 [3] では、大腸の内視鏡画像によく含まれる血管を縫合糸に代わる参照物体として応用させた。しかし、カメラ移動量を正確に計算するには2枚の画像間の高精度な対応付けを行う必要がある。ポリープを含む腸内は表面が滑らかであり、エッジなどがほとんど存在しないため、対応付けには血管領域を用いている。つまり、1枚目画像中の血管領域の移動先を2枚目画像中から正しく検

出する必要がある。

パターンや対象物体を検出する手法には図形の幾何学的なモデルを利用した Hough 変換による検出法や Merlin 法などがある。対象物体の幾何学モデルや図形そのものを利用し、パラメータ空間や画像空間での投票によって検出を行う。これらの手法には対象物体の隠蔽や線が不連続となっても検出できるといった特徴がある一方で、対象物体が幾何学モデルや図形で表現できるものでなければならない。そのため、画像認識、コンピュータビジョン、パターン認識などの幅広い分野においては汎用性が低く、画像マッチングには主に特徴ベースマッチングと領域ベースマッチングの2種類が用いられている。特徴ベースマッチングは、画像からのエッジやコーナーなどの特徴点を検出し、その周囲の局所領域に対して局所特徴量を定義し、局所特徴量の距離に基づいて画像間のマッチングを行う。特徴点検出として、Harris point, DoG(Difference of Gaussian) region, Hessian Affine region, MSER(Maximally Stable External Regions) などがあり、特徴抽出として、SIFT(Scale-Invariant Feature Transform)[4], SURF(Speeded-Up Robust Features)[5], HOG(Histograms of Oriented Gradients)[6], ORB(Oriented FAST and Rotated BRIEF), AKAZE(Accelerated KAZE)などが提案されている。これらの多くは、画像変形に頑健なマッチング手法である。特徴抽出には時間はかかるが、マッチングには比較的時間がかからないため、大量の画像マッチングに適している。しかし、特徴点を検出できなかつたり、対応付けられない領域が生じるため、対応付けの結果が疎となってしまうことが問題である。領域ベースマッチングは、対象物体のブロック画像と、入力画像の局所ブロック画像を相違度あるいは類似度の尺度を用いてマッチングを行う。相違度として SAD(Sum of Absolute Differences), SSD (Sum of Squared Differences), 類似度として NCC(Normalized Cross Correlation), POC(Phase-Only Correlation) が用いられている。対象物体のブロック画像に密な基準点を設定することで、画像全体に分布する密な対応点を得ることが可能である。しかし、特徴ベースマッチングと比較して計算時間が多いことと、大きな画像変形には対応できないことが問題である。文献 [3] では血管領域の移動先推定にテンプレートマッチングを用いているが、カメラの回転や肉壁の非剛体運動が含まれやすい内視鏡環境において不向きであり、正しい移動先を推定出来ない場合がある。

そこで本手法では血管領域から血管構造をグラフとして抽出し，2枚の画像間でグラフ構造の同定をすることで，血管領域の移動先を推定する．グラフ構造でのマッチングができれば，カメラの回転や対象物体の非剛体運動を含む場合でも線単位で対応が取れ，結果として密な対応点が得られる．これにより，内視鏡カメラの回転運動，並進運動といった移動量のパラメータ推定の精度向上が期待できる．

第2章 背景知識

2.1 Neural Network

ニューラルネットワーク (NN:Neural Network) は人間の神経細胞の学習メカニズムをモデルにして作られたアルゴリズムである。NN の構造は入力層，隠れ層，出力層から構成され，層と層の間には，ニューロン同士の繋がり**の強さを示す重み**がある。NN において学習とは正しい答えに近づけるよう，パラメータ (重みとバイアス) を調整する作業を指す。この学習を繰り返すことで精度を向上させ，高精度な予想を可能にする。

2.2 Convolutional Neural Network

畳み込みニューラルネットワーク (CNN:Convolutional Neural Network) は画像認識や音声認識などで多く使われている NN である。一般的な順伝搬型の NN とは異なり，畳み込み層 (Convolution Layer) やプーリング層 (Pooling Layer) から構成されているという特徴がある。CNN の入力には RGB の要素を持ったカラー画像やグレースケール画像とする。その画像に対して $n_x \times n_y$ のフィルタを畳み込むことで特徴マップを得る。その際に得られた値に対し活性化関数を適用することで次の層の入力となる。そしてプーリング層によって特徴マップのダウンサンプリングを行う。このような処理がいくつか重なることでネットワークが深くなっていき，画像から特徴を抽出することが出来る。そしていくつか処理を重ねた後にそれまでの特徴マップを1次元にし全結合層への入力とする。最後の出力層でのユニット数により最終的な結果が変わり，2クラス分類や多クラス分類など様々な問題に対応することが出来る。CNN をベースとした代表的なアーキテクチャには LeNet[7](図 2.1) や AlexNet[8] などが存在する。CNN の基本的な概要図を図 2.2 に示す。

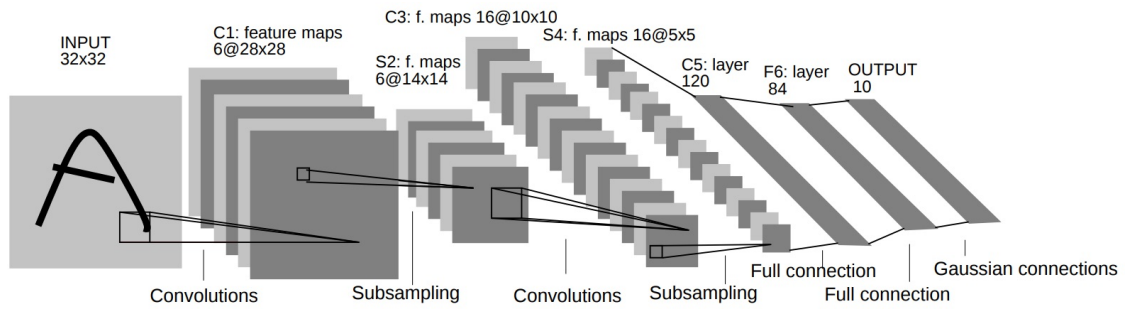


図 2.1: LeNet のアーキテクチャ(文献 [7] より引用)

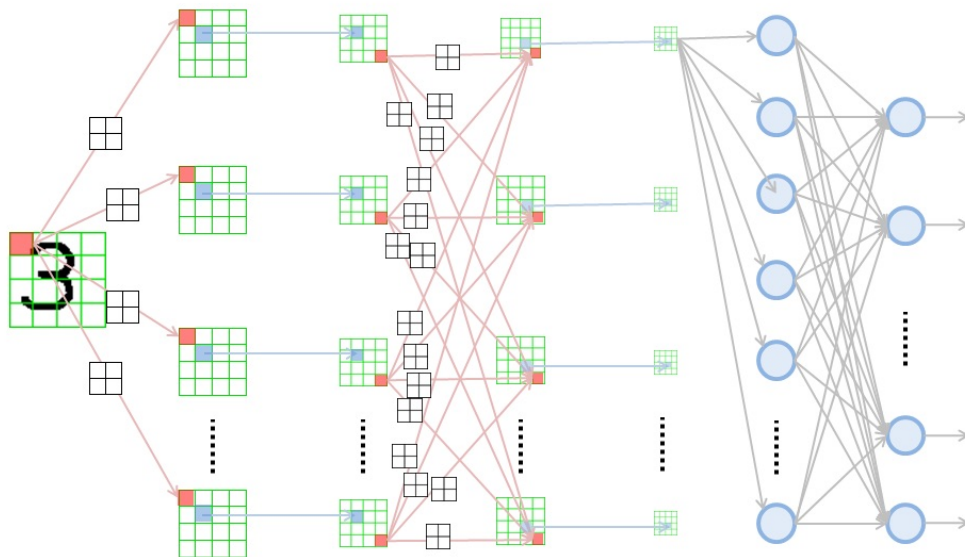


図 2.2: CNN の概要図

2.2.1 畳み込み層 (Convolution Layer)

畳み込み層は画像全体に対し各フィルタごとに畳み込み演算された特徴マップを出力する。一つの層に複数のフィルタが存在し、フィルタの数の深さが出力となる。フィルタは画像全体に対して適用されるため画像の特徴がどこにあっても抽出することが出来る。これを位置不変性という。このような処理を行うことにより、画像から全体的な漠然とした特徴を抽出することが出来る。この出力に対し次に示す活性化関数を適用することにより最終的な出力とし、次の層への入力となる。

2.2.2 ReLU(Rectified Linear Units)

次の層への入力とするために得られた特徴マップから活性化関数を適用するが、畳み込み層においてはその活性化関数に一般的に Rectified Linear Units(ReLU) 関数が用いられている。ReLU 関数の式を式 (2.1) に示す。ReLU 関数は0以上の入力に対してはそのままの入力を返し、0より小さい場合は0を返す。これにより勾配消失問題を回避することが可能なため、一般的に用いられている。ReLU 関数を図 2.3 に示す。

$$a_{ij} = \text{ReLU}(x_{ij}) = \max(0, x_{ij}) \quad (2.1)$$

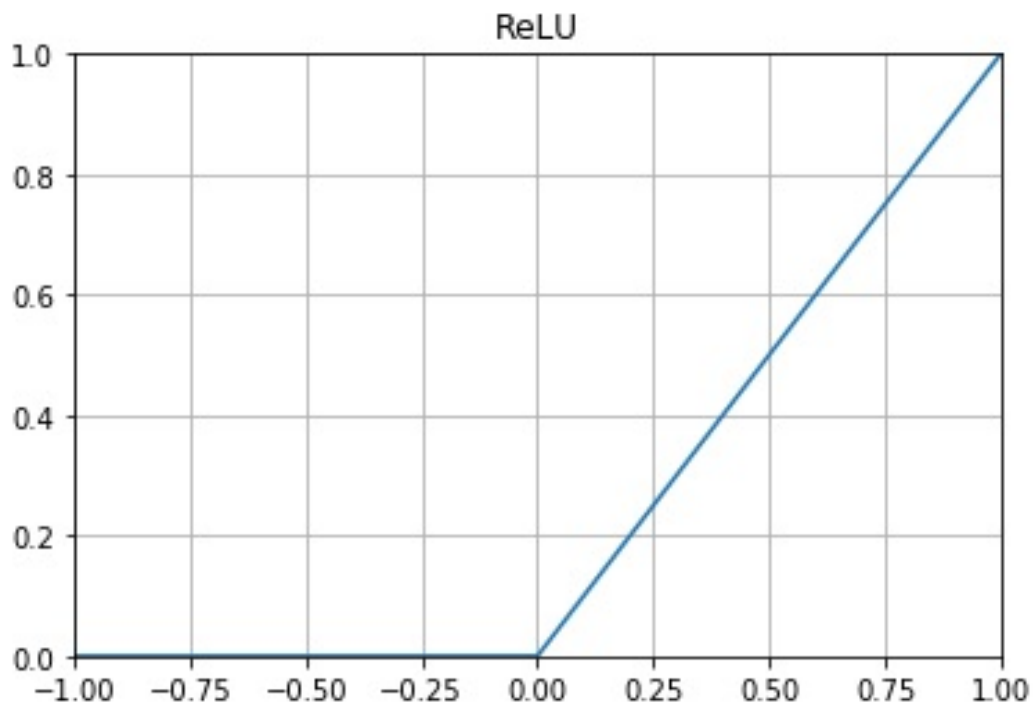


図 2.3: ReLU 関数

2.2.3 プーリング層 (Pooling Layer)

プーリング層は畳み込み層の後に適用されることが多く、入力されたデータに対しダウンサンプリングを行う。プーリング層ダウンサンプリングにより画像サイズが小さくなり情報が圧縮される。それにより微小な変化に頑健になったり計算コストを防いだりすることが可能になる。プーリング層の処理にはいくつか種類が存在し、ここではMax PoolingとAverage Poolingについて記述する。Max Poolingは指定された領域の中の最大値を結果とし、Average Poolingは指定領域中の平均値を結果とする手法である。それぞれの流れを図2.4に示す。

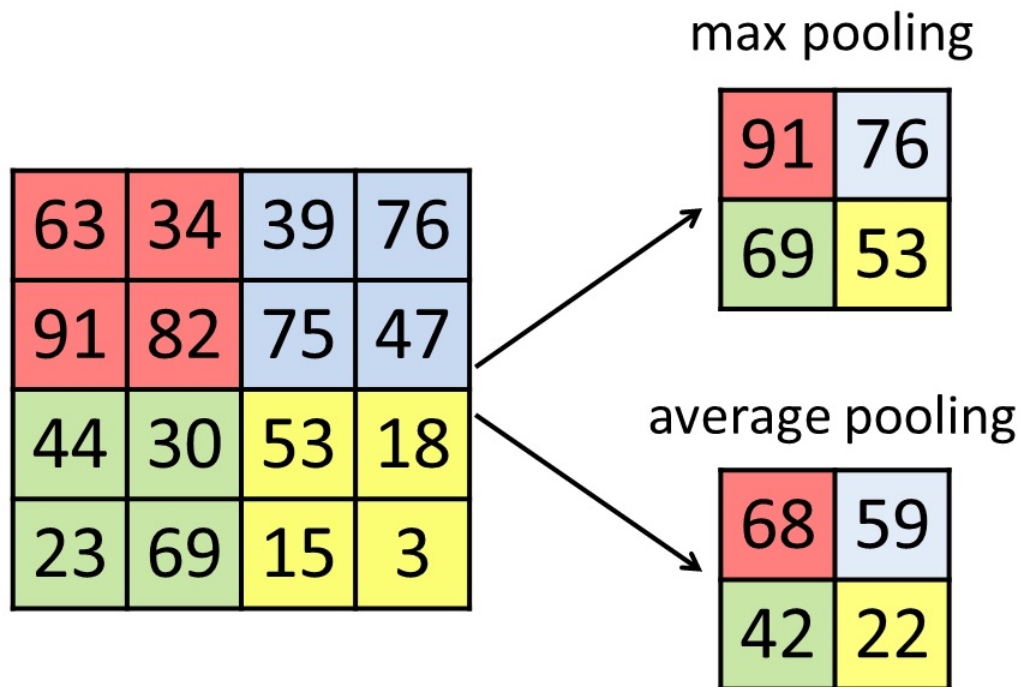


図 2.4: Max Pooling, Average Pooling

2.2.4 全結合層 (Fully Connected Layer)

これまでで得られた特徴を抽出した画像データを1次元のベクトルに変換し一つのノードに結合する。さらに活性化関数によって変換された特徴変数を出力とする。ノードの数が増えると各領域を特徴付ける特徴変数の数が増えることになる。

2.2.5 逆畳み込み (Deconvolution)

逆畳み込みは転置畳み込み (Transposed Convolution) とも言われ、元の特徴マップを拡大した後に畳み込みを行う処理である。図2.5のように 3×3 の特徴マップの間や周囲へゼロパディング処理をしたあとフィルタを適用することで 5×5 の特徴マップが得られていることが分かる。この処理を行うことで局所的な特徴をグローバルなものへと変化させることが可能になる。

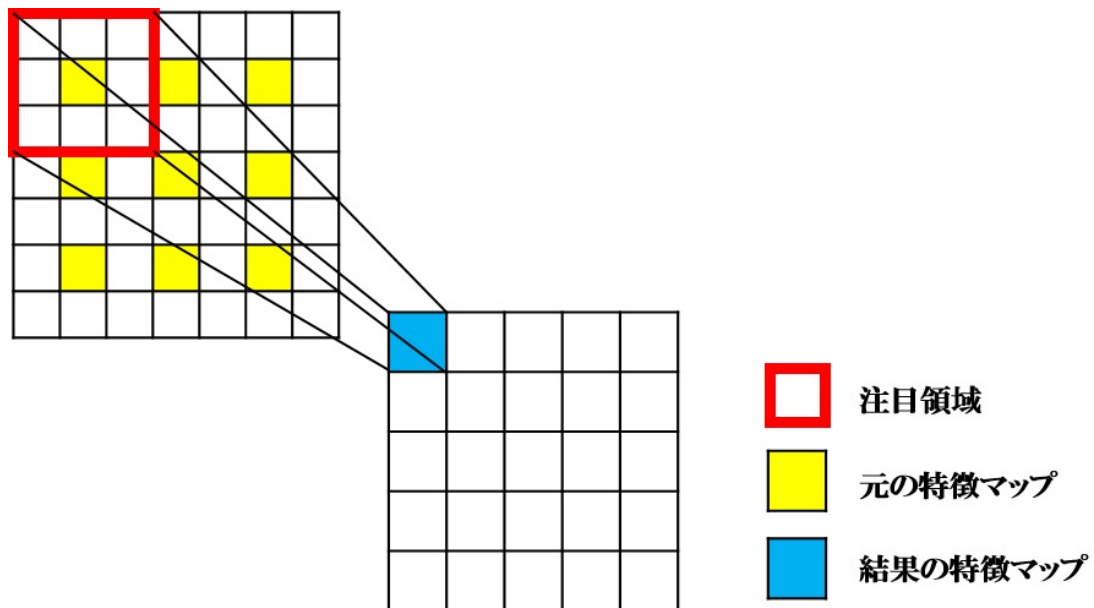


図 2.5: 逆畳み込み

2.3 U-net

U-net[9] は全てが畳み込み層で構成されている Fully Convolutional Network(FCN)[10] の一種である。画像をピクセル単位で検出するセグメンテーションに特化しているネットワークで図 2.6 のように U 字型のネットワークになることから U-net と呼ばれている。青色のボックスの上部の数字はチャンネル数である。

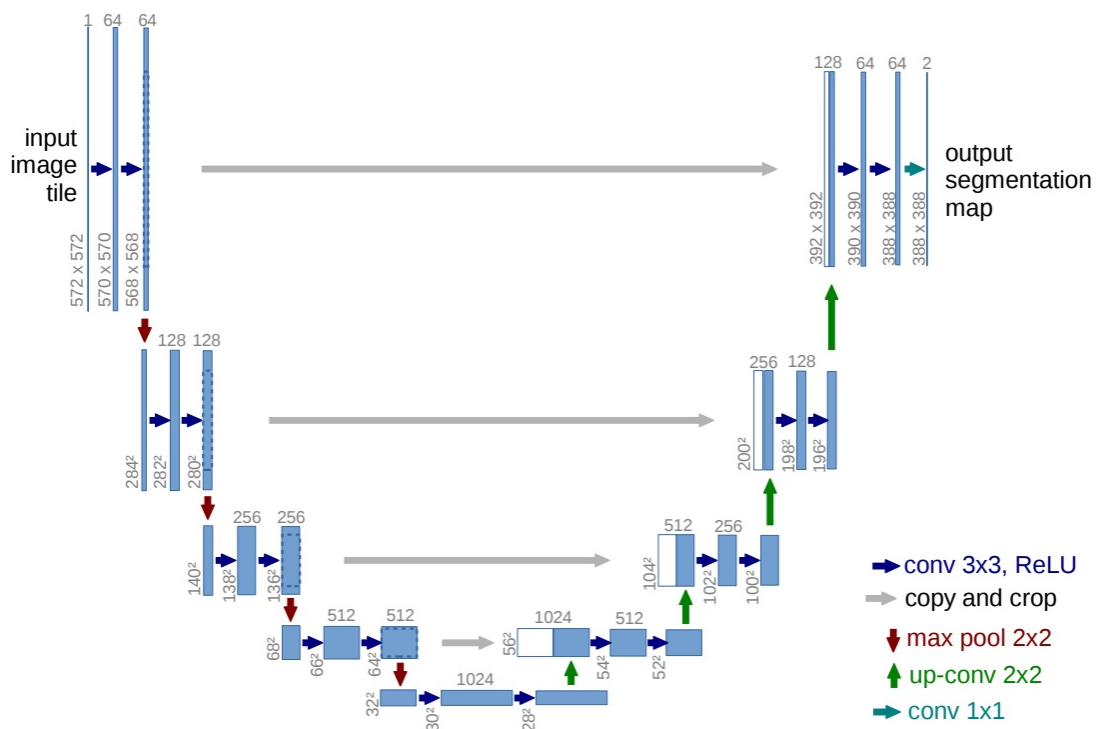


図 2.6: U-net のアーキテクチャ(文献 [9] より引用)

U-net の大きな特徴としてアップサンプリングと Merge 処理が存在することと全結合層が存在しないことが挙げられる。同じく畳み込み層を使用している CNN では畳み込み層とプーリング層によって局所的な特徴を抽出するため、特徴の情報はネットワークが深くなっていくほどさらに局所的なものになっていく。そこで U-net ではその後に逆畳み込み処理を行うことで局所的な特徴からグローバルな特徴へと変化させている。この処理をアップサンプリングと呼び、図 2.6 では緑の上矢印で表されている。さらにネットワークを深くしていく処理の中で得られた特徴マップそ

のものをアップサンプリングした後に転送することによって位置に関する情報を補完している。この処理は Marge 処理と呼び、灰色の矢印で表されている。最終的な出力は画像となるため、CNN で扱われている全結合層は必要としない。このような処理を行うことで局所の特徴を保持したまま位置的な特徴も学習することが出来る。

2.4 特徴ベースマッチング

特徴ベースマッチングとは、画像からのエッジやコーナーなどの特徴点を検出し、その周囲の局所領域に対して局所特徴量を定義し、局所特徴量の距離に基づいて画像間のマッチングを行う。これまでに提案されている特徴ベースマッチングについて述べる。

- SIFT(Scale-Invariant Feature Transform)

1999 年に Lowe が提案して以来、画像処理やコンピュータビジョンなどの幅広い分野で用いられているマッチング手法である。特徴点検出に DoG を用いることで拡大縮小に不変な特徴点を検出し、輝度勾配のヒストグラムを用いることで特徴点近傍の回転を求める。そして、特徴点の周辺を 4×4 のブロックに分割し、ブロックごとに 8 方向の勾配ヒストグラムを求め、128 次元の特徴ベクトルとする。

- SURF(Speeded-Up Robust Features)

SURF は、SIFT よりも高速な特徴ベースマッチングとして知られている。SIFT との違いは、Hessian 行列と Integral image を用いることで高速な特徴検出を、Haar wavelet を用いることで高速な特徴量抽出を実現している点である。SURF は、単なる高速化手法ではなく、SIFT よりも識別性能が高いと言われている。

- ORB(Oriented FAST and Rotated BRIEF)

ORB は、2 点の画素値の比較によってバイナリコードを生成するという BRIEF の考え方にスケール不変性と回転不変性を導入したものである。SIFT、SURF 同様に回転や拡大縮小に頑健でありながら、高速 (SIFT の 100 倍、SURF の 10 倍) である。ORB の特徴点検出には FAST を用いる。また、スケール不変性を

導入するために、入力画像を多段階に縮小した画像ピラミッドを作成する。それぞれのレイヤに対して FAST を適用して全ての画素に関して特徴点らしさのレスポンス値を算出する。レスポンス値が閾値以上であり、かつ x , y 方向およびスケール方向の三次元空間において局所最大値となる点を特徴点として抽出する。

- AKAZE(Accelerated KAZE)

AKAZE のもととなっている KAZE のアイデアは、SIFT や SURF で使われている Gaussian filter によるスケールスペースは、Gaussian filter が等方的であるため、オブジェクトのエッジもぼかしてしまい、局所的な特徴をうまくとれないことがある。それを解決するために、非線形で非等方的なスケールスペースを使った手法である。

2.5 領域ベースマッチング

領域ベースマッチングの代表例としてテンプレートマッチングがある。テンプレートマッチングとは、入力画像からテンプレート画像(部分画像)と最も類似する箇所を探索する処理であり、入力画像の一部とテンプレート画像の類似度を求め、類似度が最も高い場所を探索する。その類似度の計算方法にはいくつか種類があり、その特徴を述べる。

- SAD(Sum of Absolute Difference)

テンプレート画像と探索対象の同じ位置の画素の輝度値の差の絶対値の和を用いて、相違度を計算する。そのため、相違度が小さいほど、類似していると評価する。SAD は後述する SSD と比較して照明の影響をかなり受けやすいが、計算量が少ない点や外れ値の影響を受けにくいといった利点がある。

- SSD(Sum of Squared Difference)

テンプレート画像と探索対象の同じ位置の画素の輝度値の差の2乗の和を用いて、相違度を計算する。そのため、SAD 同様相違度が小さいほど、類似していると評価する。

- NCC(Normalized Cross Correlation)

SAD や SSD と異なり, NCC は画像間の類似度を計算する手法. -1 から 1 の値を取り, 1 なら同一画像, -1 ならネガポジ反転画像となる.

- POC(Phase-Only Correlation)

POC は, 画像をフーリエ変換して得られる情報のうち, 位相情報のみを用いて画像をマッチングする手法. NCC と同様, 画像間の類似度を計算する手法. 2 つの画像が類似している場合, POC 関数はデルタ関数に近い極めて鋭いピークを有する. この相関ピークの高さは画像の類似度の尺度として有用であり, 一方, 相関ピークの座標は 2 つの画像の相対的な位置ズレに対応する.

第3章 提案手法



図 3.1: 内視鏡画像の例

本手法では非剛体運動する対象物体でも2枚の画像間で密な対応付けをすることを目的とする。一般的な内視鏡環境下にて撮影された大腸の内視鏡画像の例を図3.1に示す。図3.1では、大腸の肉壁に散らばる血管を確認することができる。これらの血管構造は2枚の画像間で非剛体運動が起こったとしても変化しないことに着目した。そこで本章では、血管領域から血管構造をグラフとして抽出し、2枚の画像間でグラフ構造の同定を行うことで、血管領域の移動先を推定する。

3.1 血管領域抽出

一部の血管領域を検出する手法に、Frangi Filter[11]を用いた文献[12]がある。しかし、本手法で画像間のマッチングを行うに当たって画像全体の血管領域を取得する必要がある。そこで、医用画像処理分野において、実応用されているディープニューラルネットワーク(DNN:Deep Neural Network)を用いる。DNNの代表的なアーキテク

チャの一つに畳み込みニューラルネットワーク (CNN:Convolution Neural Network) がある。通常の CNN によって行われる画像のクラス分類では、畳み込み層が対象の局所的な特徴を抽出する役割を担い、プーリング層が対象の全体的な位置情報をぼかす役割を担う。そのため、より深い層ほど抽出される特徴はより局所的になり、またその特徴の全体的な位置情報はより曖昧になる。CNN による検出処理の多くは画像中の対象を発見し、対象の大きさや位置を同定することを目的としている。一方の領域分割は、画像中の個々の対象の領域をピクセルの精度で同定することを目的とする。つまり、プーリング層でぼかされた局所的特徴の位置情報を元画像上でピクセル単位で復元する必要がある。そこで、対象の局所的特徴と全体的位置情報の両方を結合して学習させるために開発された U-Net[9] を用いて血管領域分割を行う。U-Net の中でも、文献 [13] で用いられる Retina U-Net によって血管領域分割を行う。本研究の目的は血管構造のマッチングであり、血管抽出はあくまで手段であるため、変更は必要最小限とし、今回扱う内視鏡画像データ (OLYMPUS 製) では、畳み込み層を 5 層から 3 層への変更のみが妥当であった。

以下に手順を示す。

Step 1. 図 3.2(a) にある学習用の画像から図 3.2(b) にあるマスク画像を作成する。

Step 2. 学習用の画像と Step1 のマスク画像から 48×48 ピクセルの 9500 個のランダムパッチを抽出する。図 3.2(c) にある Field Of View(FOV) 境界と血管領域を区別するために、一部または完全に FOV の外側にあるパッチも抽出する。また、元の画像の同じ部分を含む複数のパッチが抽出される場合もあるが、それ以上のデータ拡張は行わない。

Step 3. Step2 のパッチから血管領域を特徴とさせて学習を行う。

Step 4. Step3 の学習データを元にテストデータから学習させた特徴を検出する。

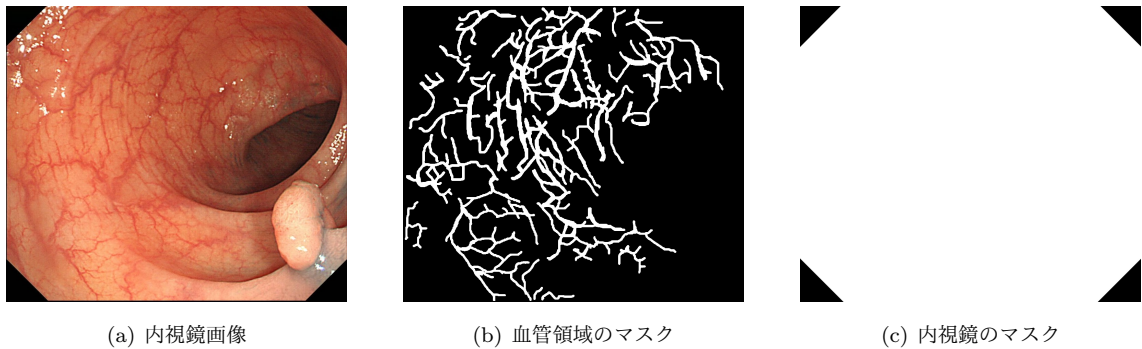


図 3.2: 血管構造の分類例

3.2 ノイズ除去

内視鏡画像からの血管検出には Olaf Ronneberger らの U-net[9] を用いている。U-Net による血管検出画像は 256 階調のグレースケール画像であるため、血管検出画像から直接血管構造を取得することは困難である。そこで、血管検出画像に対して 2 値化処理を行うことで、白と黒の 2 階調に変換し、血管領域を明確にする。また、血管検出画像には細かいノイズが含まれているため、ラベリングにより注目領域のみを抽出する。

以下に手順を示す。

Step 1. 血管検出画像に対して、OpenCV のライブラリである大津の 2 値化処理をする。

Step 2. Step1 から得られた画像に対して、OpenCV のライブラリであるラベリング処理をする。

Step 3. Step2 から得られた画像から最も大きい血管領域のみを取り出すことでノイズを取り除く。

3.3 モルフォロジー変換

本手法では画像間のマッチングを行うため、2 枚の画像からグラフ同定に適した血管構造を取得する必要がある。処理目的に適した構造要素を取得できれば、通常の

画像処理よりも効果的な処理手法が実現できる。そこで、OpenCVのライブラリであるオープニング処理とクロージング処理をする。これにより、画像中の小さなノイズを取り除き、血管領域中の小さな穴を埋める効果があり、画像間の血管構造グラフの差異を減らす。

3.4 血管構造のグラフ化

血管構造のグラフ化の前処理として、scikit-imageのライブラリであるSkeletonizeを用いてモルフォロジー変換後の画像を細線化する。細線化処理により、血管構造を保持した状態の幅1ピクセルの線画像となり、血管の端点や分岐点などの特徴が抽出しやすくなる。

以下に細線化画像から血管構造のグラフを抽出する手順を示す。

Step 1. 細線化画像からグラフの端点の座標を1つを線形探索する。

Step 2. Step1により探索された端点からグラフ構造を抽出する再帰的アルゴリズムを呼び出す。

Step 3. 再帰的アルゴリズムにより、血管構造の端点と分岐点をノードとした無向グラフを取得する。

Step 4. Step1~Step3を細線化画像中のグラフ数分繰り返す。

3.4.1 血管構造の分類方法

細線化処理を適用した血管領域から端点や分岐点の判別方法について示す。血管領域の画素値を1とした場合、血管領域上の注目画素の8近傍に存在する画素値1の領域がいくつ存在するかによって判別する。ただし、隣接する画素値が同じ場合、その領域を1つの領域とみなす。図3.3に示すように画素値1の領域が1つ存在した場合は端点、2つ存在した場合は連結点、3つ以上存在した場合は分岐点として分類を行う。

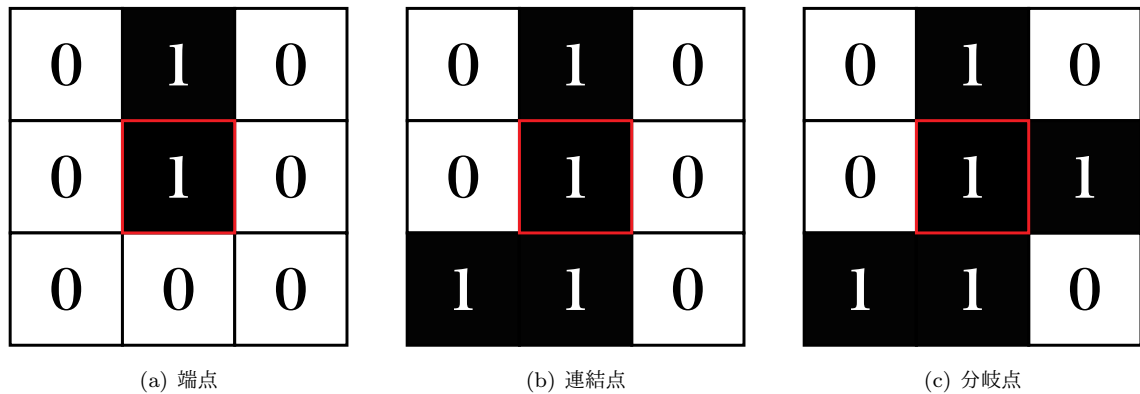


図 3.3: 血管構造の分類例

3.4.2 グラフ抽出アルゴリズム

血管領域からグラフ構造を定義する再帰的アルゴリズムを示す。

Step 1. 注目画素が血管構造グラフの端点，分岐点である場合，注目画素に一意の数字を割り当て，グラフのノードとして定義する。

Step 2. ノード間に繋がりがある場合，そのノードの組み合わせをグラフのエッジとして定義する。

Step 3. 注目画素の隣接画素かつ未確認の血管領域の画素を次の注目画素とし，再び自身のアルゴリズムを呼び出す。

以上の手順によって血管領域からグラフ構造を抽出できる。注目領域が移動する順番，またその結果得られるグラフ構造を図 3.4 に示す。

Step3 において，注目画素の 8 近傍から未確認の血管領域の画素を探索し，次の注目画素と定義するが，該当する画素が複数存在する場合がある。そこで，複数の該当する画素に対して，血管構造の分類に基づき分岐点，連結点，端点の順に優先順位を定義する。具体的には，次の注目画素の候補点に端点と連結点を含む場合，連結点が優先され，分岐点と連結点を含む場合，分岐点が優先される。また，候補点がどちらも連結点の場合，図 3.5 の番号順に優先順位付けをする。候補点の優先度順位例を図 3.6 に示す。

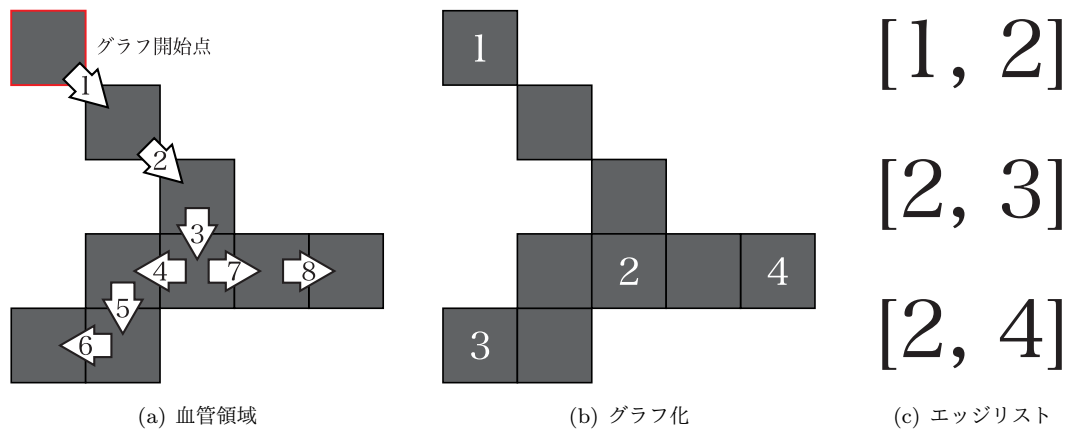


図 3.4: 血管領域のグラフ化の例

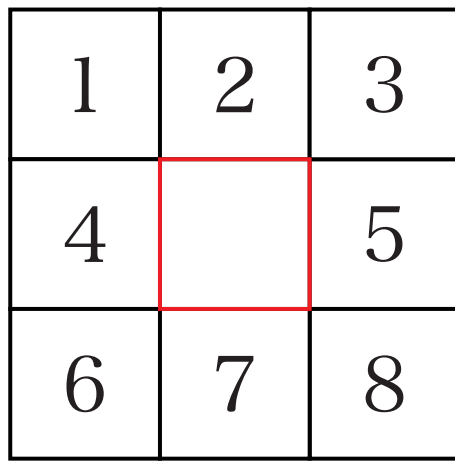


図 3.5: 連結点の優先順位

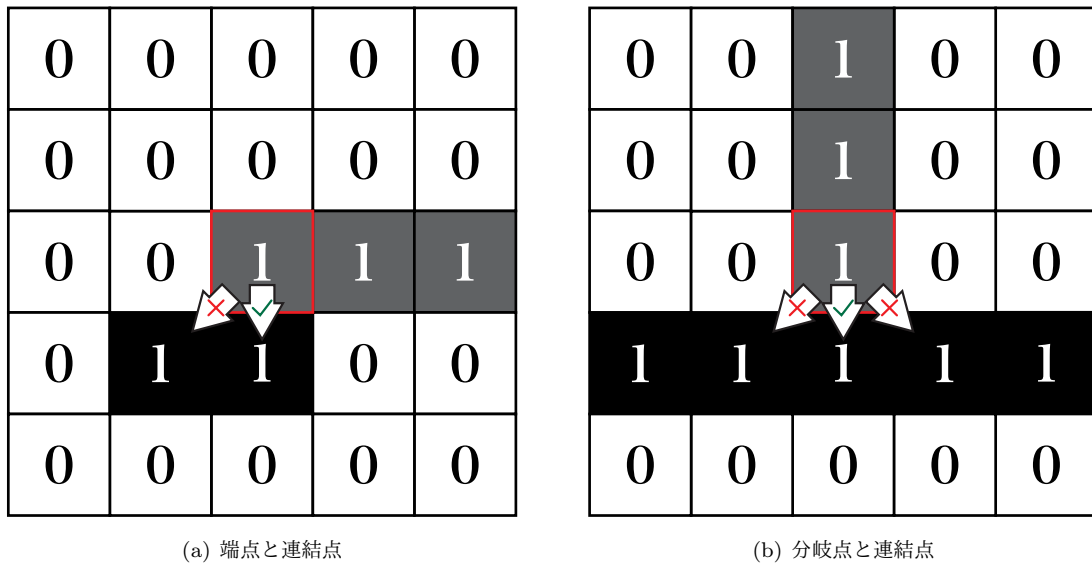


図 3.6: 候補点の優先順位例

3.5 グラフ構造の補間

図 3.1 に示すように、腸内の肉壁は表面を水分に覆われていることが多いため、内視鏡の光源を鏡面反射してしまう。血管領域上で鏡面反射が起こる場合、内視鏡画像からの血管領域分割には Retina U-Net を用いているため、鏡面反射箇所を血管と認識しないことがある。その結果、鏡面反射箇所では血管領域が分断され、抽出される血管領域の血管構造も分断されてしまう。本手法では、より広い範囲の血管領域を対応付けたいため、抽出される血管構造が分断されることは適切ではない。そこで、血管領域が分断された可能性がある箇所を検出し、その箇所をエッジで結ぶことでグラフ構造を補間する。血管領域が分断されるとき、グラフ構造では本来エッジであった箇所に新しいノードが2つできることになる。図 3.7 に例を示すように、本来のエッジ $[1,2]$ に 3, 4 という新しいノードができてしまう。このときに、本来のエッジ $[1,2]$ を追加することでグラフ構造を補間する。

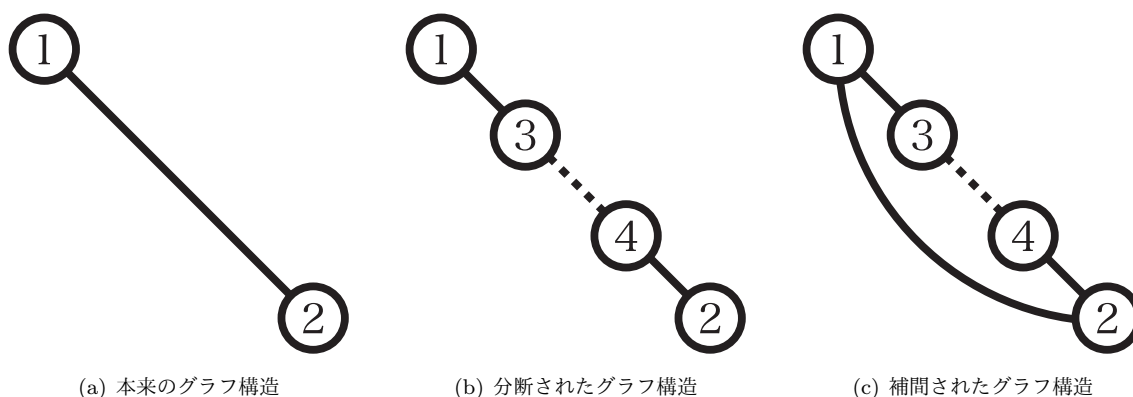


図 3.7: グラフ構造の補間例

以下にグラフ構造の全ての端点から分断された端点同士であると判定する条件を述べる.

- 2つの端点の親ノードが異なるノードである.
- 2つの端点の親ノード間に既にエッジが存在する.
- 2つの端点を結ぶエッジが他の全てのエッジと交差しない.

この条件を満たす2つの端点の親ノード間にエッジを結ぶ. これにより, 次の節で説明するグラフの対応付けでより多くの対応が取れるようになる.

3.6 グラフの対応付け

2枚の画像から得られたグラフ構造を同定するアルゴリズムについて述べる. 2つのグラフが同型であるかどうかを判定し, また同型である場合にはグラフの対応付けを行うライブラリが NetworkX にある. しかし, NetworkX でグラフの対応付けを行うには対応付けする2つのグラフが同型でなければならないという条件がある. 本研究では2枚の画像から血管領域のグラフ構造を抽出するが, 抽出されるグラフ構造は同型ではないことが多いため利用できない. そこで, 2枚の画像から抽出された血管領域のグラフ構造に共通するグラフ構造を取り出し, 対応付けを行うアルゴリズムを考える. 2つのグラフ構造に共通するグラフ構造を取り出すということは, 言い換えると共通する部分グラフを取り出すということである.

発見したいパターンを定義したグラフ (以下, パターングラフ) と探索対象のグラフ (以下, データグラフ) が与えられたときに, データグラフに含まれるパターングラフと同型の部分グラフを全て見つけるという部分グラフ同型問題は広い分野において大きな課題として存在し, 様々な手法により解かれている. 部分グラフ同型問題に関する先行研究として, Ullmann によって提案された文献 [14] がある. さらに, Ullmann のアルゴリズムを拡張した, VF2[15], QuickSI[16], GraphQL[17] なども提案されている. しかし, 部分グラフ同型問題は NP 完全であり, 上記の手法を用いたとしても, この問題を解くには最悪の場合指数時間を要する. また, 本手法はパターングラフとデータグラフが与えられる問題ではなく, 2つのデータグラフが与えられる問題であり, そこから共通するパターングラフを見付け出さなければならぬ. そこで, 2つのデータグラフに対して対応するノードを一つ決め, そこから出来るだけ多くの対応が取れたパターングラフを返すアルゴリズムを考える.

以下に手順を示す.

- Step 1.** 2つのグラフから一つのノード同士を対応付ける.
- Step 2.** 2つのグラフそれぞれから, 対応付けられたノードに隣接する未対応のノードを取得する.
- Step 3.** 未対応のノードが一つずつの場合, そのノード同士を対応付ける.
- Step 4.** 未対応のノードが複数の場合, 未対応のノード同士の全組み合わせ通りの対応付けを考える.
- Step 5.** Step2~Step4 を対応が取れなくなるまで繰り返す.
- Step 6.** 対応づけられたグラフのうち最も多くの対応が取れたグラフを取り出す.

第4章 実験

前章で示した提案手法の有効性を確認するため、血管領域のシミュレーション画像と実際の内視鏡画像を用いて実験を行う。なお、本手法では式(4.1)を用いて対応付け精度の評価を行う。式(4.1)の分母であるエッジの総数は、2枚の血管領域画像から抽出されるグラフからエッジ数を比較し、エッジ数が多いものを用いる。

$$\text{対応付け成功率} = \frac{\text{対応が取れたエッジ数}}{\text{エッジの総数}} \quad (4.1)$$

また、2枚の血管領域画像から抽出されるグラフ構造両方に対してグラフ構造の補間を行う場合、グラフ対応付けの候補が爆発的に増えてしまう。そのため、2つのグラフ両方に対してグラフ構造の補間を適応するのではなく、1つのグラフに対してのみグラフ構造の補間を適応する。本実験では、2つのグラフのノード数を比較し、ノード数の多いグラフのみにグラフ構造の補間を適応する。

本実験に用いた計算環境は表 4.1 の通りである。

表 4.1: 計算環境

CPU	2.3 GHz Intel Core i5
OS	MacOS Mojave 10.14.5
メモリ	16 GB

4.1 シミュレーションによる実験

本節では、血管領域からグラフ構造を抽出した結果を示し、グラフ構造の補間を行う場合と行わない場合でのグラフ対応付け成功率を比較し、グラフ構造の補間の有効性を示す。そこで、画像サイズを 994×863 とした2枚の血管領域のシミュレーション画像を作成し、画像間での対応付けを行う。シミュレーションでは画像間に血管領域が大きく変形し、鏡面反射などで一部の血管領域が分断されることを想

定する。また、シミュレーション画像は2値化，モルフォロジー変換，ノイズ除去が既に行われているものとする。実験に用いた画像を図4.1に示す。

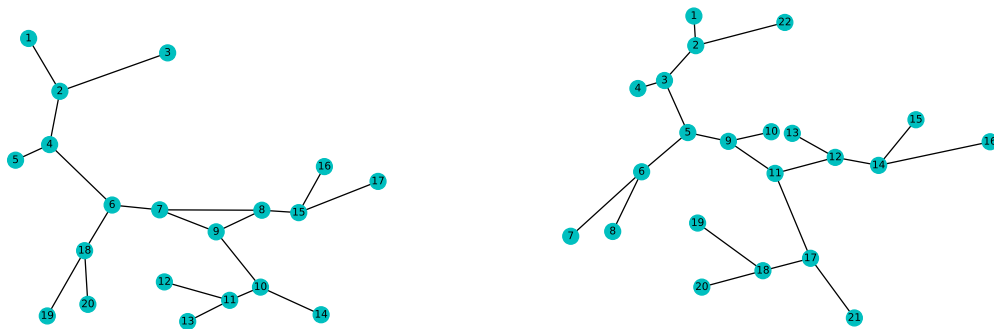
次に，2枚のシミュレーション画像に対して，血管領域のグラフ抽出を行う。得られたグラフ構造を図4.2に示す。ただし，グラフの描画はNetworkXのライブラリを用いて行い，グラフのノードの座標は元画像の座標と一致するように配置されている。また，グラフ構造の補間を行ったグラフを図4.3に示す。



(a) 1枚目の画像

(b) 2枚目の画像

図 4.1: シミュレーション画像



(a) 1枚目の画像

(b) 2枚目の画像

図 4.2: グラフ補間なし血管領域のグラフ構造

表 4.4: 対応付け精度評価

	対応が取れたエッジ数	エッジの総数	対応付け成功率 (%)
グラフ補間なし	15	21	71.4
グラフ補間あり	21	21	100

表 4.5: 計算時間 (秒)

	グラフ化	グラフ対応付け
グラフ補間なし	2.64	0.107
グラフ補間あり	2.39	0.188

図 4.2, 図 4.3 に対して, ノード 1 からグラフ対応付けを行った結果を表 4.2, 表 4.3 に示す. また, 表 4.4 に本手法の精度評価を示す. 表 4.2 について, エッジ [7,8] において血管領域が分断されているため, ノード 8 の正しい対応が取れていないことがわかる. ノード 8 が正しく対応されないことで, ノード 15, 16, 17 の対応も取れていない. しかし, 表 4.3 にあるように, グラフ構造の補間を行った結果, ノード 8 の正しい対応が取れ, その先の対応が取れるようになっている. このことから, グラフ補間を行うことで対応付けできる範囲を広げることができたといえる.

表 4.3 に本実験の計算時間を示す. グラフ構造の補間は, 血管領域からグラフ構造を抽出した後に行うため, グラフ化計算時間に大きな差は見受けられなかった. また, グラフ補間を行う分岐先のグラフ構造が少ないため, グラフ対応付けの計算時間も大きな差は見受けられなかった.

4.2 実画像による実験

本節では、2枚の内視鏡画像を用いて、画像間での対応付けを行う。また、OpenCVのライブラリであるSIFT特徴量を用いた画像マッチング手法と比較し、本手法の有効性を示す。実験に用いた画像を図4.4に示す。なお、画像サイズは994×863であり、画像間のフレーム差は6フレームである。

文献[13]のRetina U-Netによる血管領域分割結果を図4.5に示す。次に、図4.5の画像に対して、ノイズ除去、モルフォロジー変換、スケルトン化処理を行った結果を図4.6に示す。

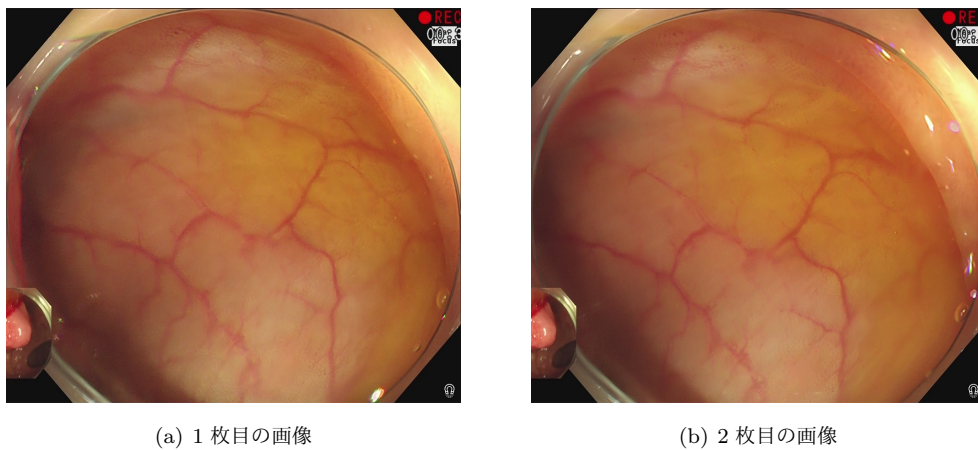


図 4.4: 内視鏡画像

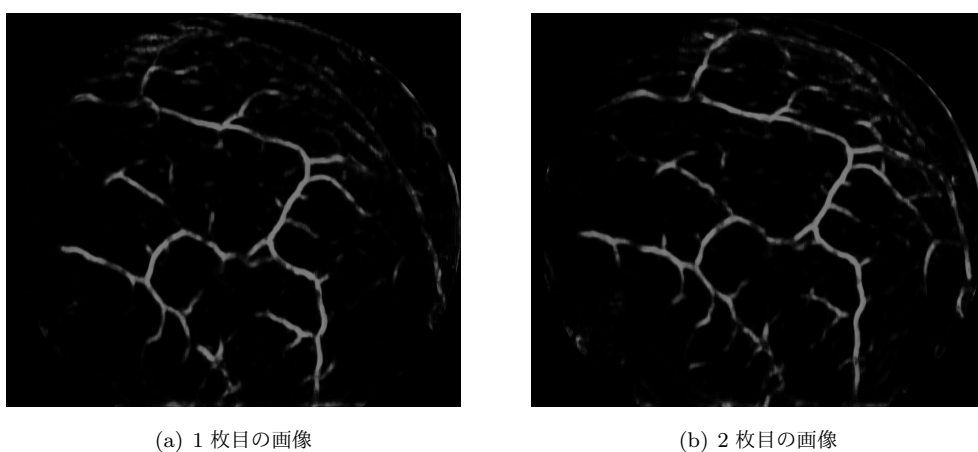


図 4.5: 血管領域画像

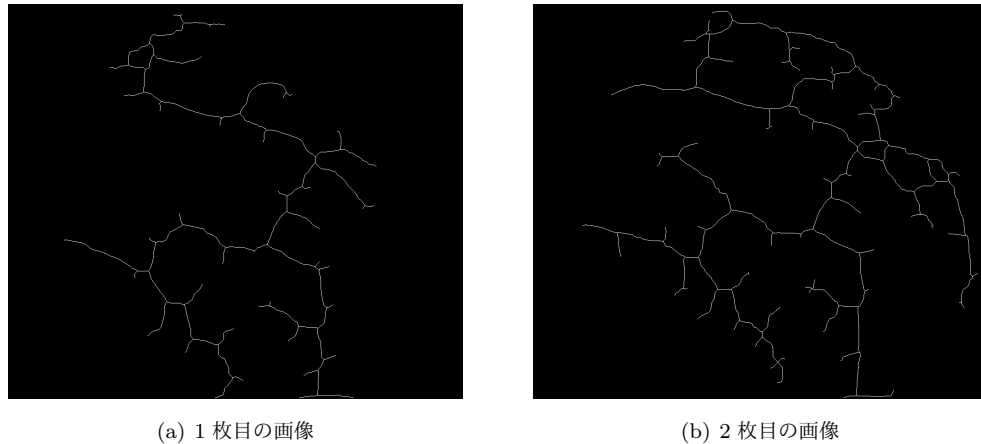


図 4.6: スケルトン化画像

まず、図 4.6 に対して、本手法を用いてグラフ構造の対応付けを行い、ノード間でマッチングした結果を図 4.7 に示す。加えて、本手法による画像マッチングで正しく対応付けできた部分グラフの範囲を図 4.9 に示す。次に、SIFT 特徴量を用いて画像マッチングを行った結果を図 4.8 に示す。図 4.4 の内視鏡画像には右上の文字や右下の記号など血管領域以外の要素も含まれているため、図 4.4 に対して SIFT 特徴量による画像マッチングを行うと、血管領域以外に多くの特徴点を検出してしまう。そこで、SIFT 特徴量による画像マッチングには、図 4.4 ではなく、図 4.5 の血管領域画像を用いる。また、より正しいマッチングを得るために、文献 [4] で説明されている割合試験 (ratio test) を行う。以下に割合試験について述べる。

x, y をそれぞれ 1 枚目の画像の SIFT 特徴量, 2 枚目の画像の SIFT 特徴量とすると, 128 次元の SIFT 特徴量間のユークリッド距離は式 (4.2) のように表せる。ここで, x に対応する点 y_i を特定するため, y_1 が x と最も距離が近い点, y_2 が x と 2 番目に距離が近い点とする。このとき, $dist(x, y_1)$ は x と最も距離が近い点との距離, $dist(x, y_2)$ は x と 2 番目に距離が近い点との距離となり, 割合試験は式 (4.3) を満たす特徴点のみ対応付けする。

$$dist(x, y) = \sqrt{\sum_{i=1}^{128} (x_i - y_i)^2} \quad (4.2)$$

$$\frac{dist(x, y_1)}{dist(x, y_2)} < R_{distance\ ratio} \quad (4.3)$$

文献 [4] によると $R_{distance\ ratio}$ を 0.8 とすることで，正しい対応付けの 5% 未満が取り除かれてしまうが，誤った対応付けの 90% を取り除けることが示されている．そこで，本実験でも $R_{distance\ ratio}$ を 0.8 とした割合試験を行い，SIFT 特徴量による画像マッチングの精度を高める．

図 4.7，図 4.8 の結果から，提案手法は SIFT 特徴量を用いたマッチングより血管領域に対して誤った対応付けしている範囲が狭い．加えて，図 4.9 の範囲ではグラフのノード間だけでなくエッジ間についても対応付けが可能となり，広い範囲で密な対応付けが可能となっていることから，提案手法の有効性を確認できる．

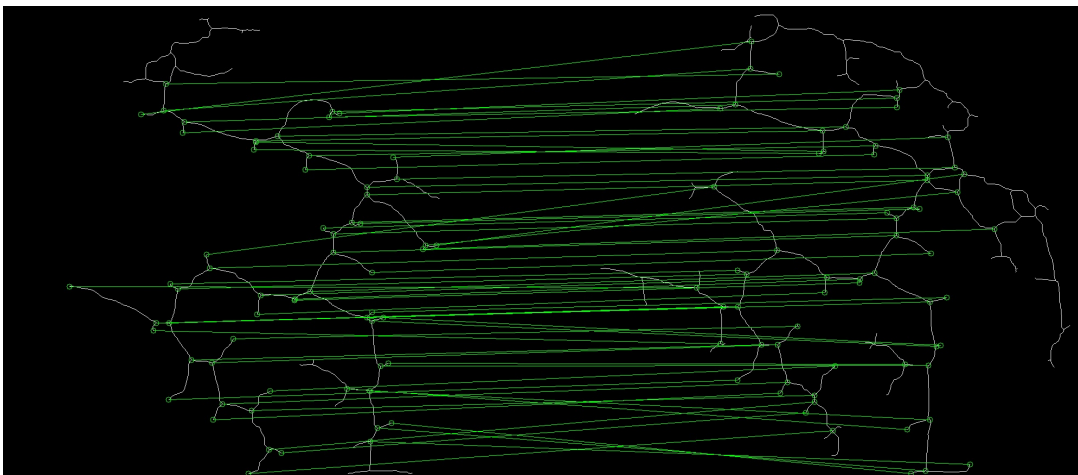


図 4.7: 本手法による画像マッチング

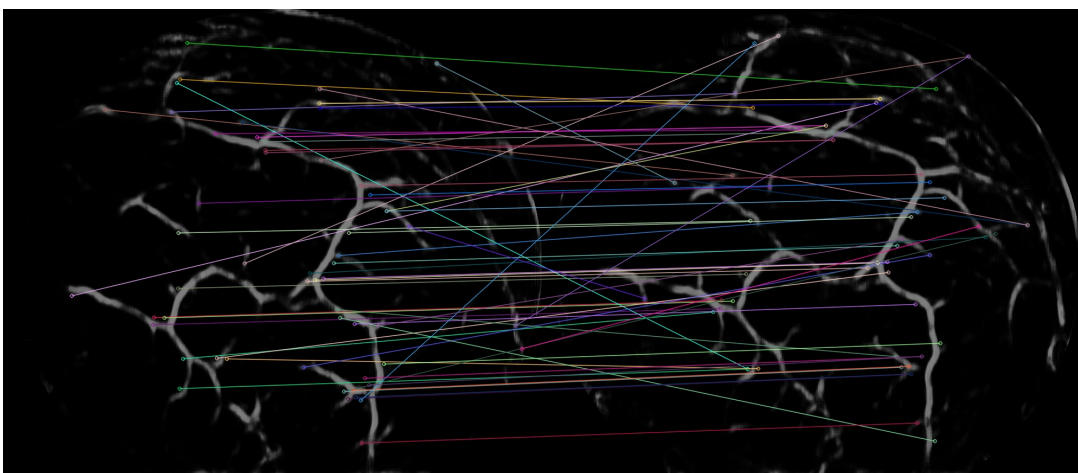


図 4.8: SIFT 特徴量による画像マッチング

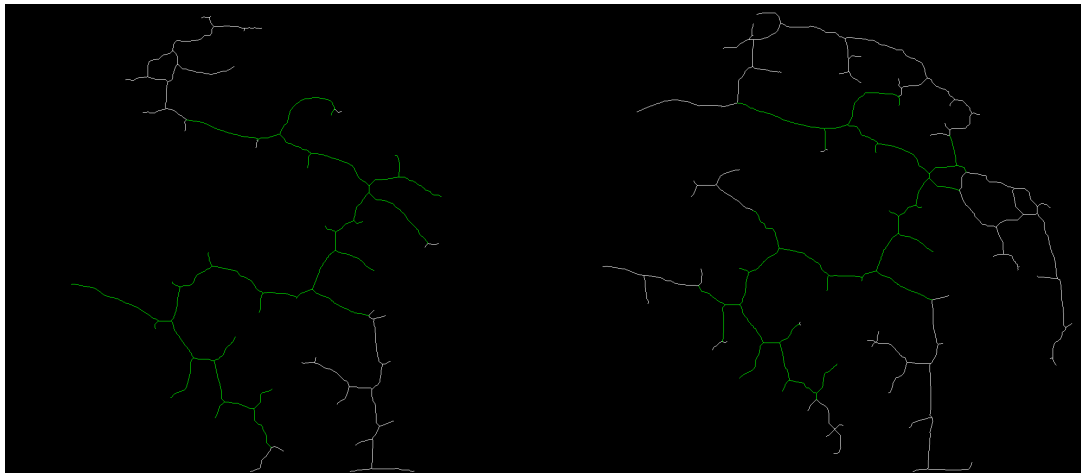
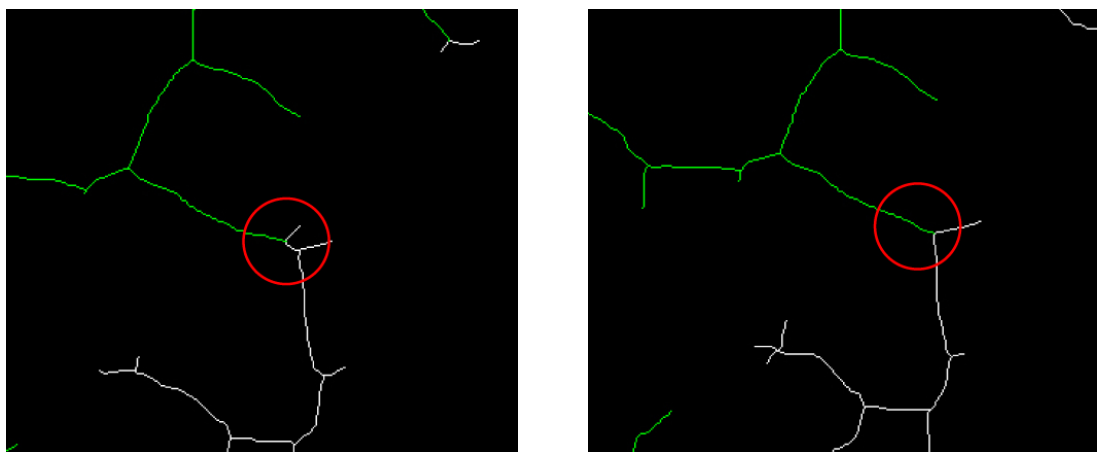


図 4.9: 正しく対応付けられたエッジ



(a) 1 枚目の画像

(b) 2 枚目の画像

図 4.10: 対応付け出来ないグラフ構造

表 4.6: 対応付け精度評価

対応が取れたエッジ数	エッジの総数	対応付け成功率 (%)
42	129	32.6

表 4.7: 計算時間 (秒)

グラフ化	グラフ対応付け
2.82	125

表 4.6 に本手法の精度評価を示す。対応付け成功率が低い原因として、画像中の血管領域が変化してしまった点が挙げられる。図 4.6 にあるように、1 枚目の画像中には存在しない血管領域が 2 枚目の画像でフレームインすることで血管領域の構造が変化してしまっている。本実験では載せていないが、1 枚目の画像中の血管領域が 2 枚目の画像でフレームアウトすることでも同様に血管構造が変化し、対応付け成功率を下げる要因となっている。また、画像間でグラフ構造が異なってしまうことも対応付け成功率が低くなる原因として挙げられる。図 4.10 に示すように、赤い円で囲まれる箇所が画像間においてグラフ構造が異なり、その先の正しい対応付けが出来なくなってしまう。

表 4.7 に本手法による計算時間を示す。表 4.5 に比べ、グラフ化計算時間に大きな差は見受けられなかった。これは、シミュレーションに用いた画像と内視鏡画像の画像サイズが同じであり、グラフ抽出アルゴリズムの再帰数に差がないためであると考えられる。その一方で、本実験での内視鏡画像は分岐先のグラフ構造が複雑なため、グラフ対応付けの計算に多くの時間がかかっている。

第5章 むすび

三次元形状復元を行う手法の一つであるステレオ法は，形状復元の精度が画像間対応点の結果に大きく依存し，多くの対応点が取れない場合に疎な形状復元となってしまうため，画像間で密な対応点を取る必要がある．また，文献 [2]，文献 [3] のような2枚の画像間でのカメラ移動量から，三次元形状復元に用いるパラメータを算出する手法でも画像間で密な対応点を取る必要がある．しかし，特徴ベースマッチングや領域ベースマッチングなどの従来手法では，内視鏡画像の血管領域のように対象物体が非剛体運動をすると，密に対応点を取ることが出来ないという問題があった．これらの課題に対して本研究では，血管領域のグラフ構造を用いることで，非剛体運動する対象物体にも線単位で対応が取れ，結果として密な画像マッチングを可能なものにした．これにより，文献 [3] に用いられる内視鏡カメラの移動量パラメータ推定の精度向上に役立つことが期待できる．

実験により発見された課題として，画像間でグラフ構造が異なる場合，その先の正しい対応付けができなくなる．より広い範囲で対応点を取るために，画像間でのグラフ構造を近づけることが今後の課題として挙げられる．また，本手法では可能な限り多くの対応点を取るようなグラフ対応付けアルゴリズムのため，正しい対応付けができなかった場合でも，その先も誤った対応付けを続けてしまう．そこで，誤った対応付けを自動で検知することが今後の課題として挙げられる．誤った対応付けを検知することができれば，その先の対応付けを行う必要がない為，誤った対応付けを減らし，計算時間を減らすことが期待される．ただし，現状では対応点の信頼性評価方法が確立されていないため，信頼性評価も今後の課題となる．

謝辞

本研究を行うにあたり，終始熱心にご指導を頂いた名古屋工業大学准教授 舟橋 健司先生に深く感謝いたします。また，日頃よりご指導いただきました中部大学教授 岩堀 祐之先生，中部大学助手 宇佐美 裕康先生に感謝いたします。

なお，本研究において，内視鏡画像を提供をはじめ，ご討論，ご協力を頂きました愛知医科大学教授 春日井 邦夫先生と准教授 小笠原 尚高先生に感謝いたします。

さらに，本研究を進めるにあたり多くの助言を頂き，御協力いただいた名古屋工業大学 舟橋研究室の皆様，および中部大学 岩堀研究室の皆様に深く感謝し，厚く御礼申し上げます。

参考文献

- [1] B.K.P. Horn, “Obtaining shape from shading information” MIT Press, pp.123-171, 1989.
- [2] Y. Iwahori, D. Yamaguchi, T. Nakamura, K. Kunio, “Estimating Reflectance Parameter of polyp Using Medical Suture Information in Endoscope Image”, Proceeding of ICPRAM 2016, pp.503-509, 2016.
- [3] Y. Iwahori, T. Suda, K. Funahashi, H. Usami, A. Wang, M.K. Bhuyan, K. Kasugai, “Shape Recovery of Polyp from Endoscope Image Using Blood Vessel Information”, Computational Science/Intelligence and Applied Informatics, pp.165-184, 2017.
- [4] D. Lowe, “Distinctive image features from scale-invariant keypoints”, International Journal of Computer Vision 60, pp.91–110, 2004.
- [5] H. Bay, A. Ess, T. Tuytelaars, L. Gool, “Supeeded-up robust features (SURF)”, Computer Vision and Image Understanding, vol.110, pp.346–359, 2008.
- [6] N. Dalal, B. Triggs, “Histograms of Oriented Gradients for Human Detection”, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005.
- [7] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, “Gradient-based learning applied to document recognition”, Proceedings of the IEEE, vol.86, no.11, pp.2278–2324, 1998.

- [8] A. Krizhevsky, I. Sutskever, G. E. Hinton, “Imagenet classification with deep convolutional neural networks”, *Advances in neural information processing systems*, pp.1097–1105, 2012.
- [9] O. Ronneberger, P. Fischer, T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation”, *International Conference on Medical image computing and computer-assisted intervention*, pp.234–241, 2015.
- [10] J. Long, E. Shelhamer, T. Darrell, “Fully convolutional networks for semantic segmentation”, *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp.3431–3440, 2015.
- [11] A.F. Frangi, W.J. Niessen, K.L. Vincken, M.A. Viergever, “Multiscale vessel enhancement filtering”, *Medical Image Computing and Computer-Assisted Intervention*, pp.130-137, 1998.
- [12] M. Golhar, Y. Iwahori, K. Funahashi, M.K. Bhuyan, K. Kasugai, “A Robust Method for Blood Vessel Extraction in Endoscopic Images with SVM-based Scene Classification”, *ICPRAM 2017*, pp.1-10, 2017.
- [13] orbix Srl, “Retina blood vessel segmentation with a convolutional neural network”, 2016. <https://github.com/orobix/retina-unet> (2020/1/31)
- [14] J.R. Ullmann, “An algorithm for subgraph isomorphism”, *Journal of the ACM (JACM)*, vol.28, no.1, pp.31-42, 1976.
- [15] L.P. Cordella, P. Foggia, C. Sansone, M. Vento, “A (sub) graph isomorphism algorithm for matching large graphs”, *IEEE Trans. PAMI*, vol.26, no.10, pp.1367–1372, 2004.
- [16] H.He and A.K. Singh, “Closure-tree: An index structure for graph queries”, *Proceedings of ICDE*, pp.38–49, 2006.
- [17] H. He, A.K. Singh, “Graphs-at-a-time: query language and access methods for graph databases”, *Proceedings of SIGMOD*, pp.405–418, 2008.