

Deep Learning を用いた内視鏡画像の分類と血管 を利用したポリープ形状復元

岩堀祐之, 宇佐美裕康, 井上智司*, 海老正秀*, 小笠原尚高*, 春日井邦夫*

Classification of Endoscope Image using Deep Learning and Shape Recovery of Polyp Based on Blood Vessel

Yuji Iwahori, Hiroyasu Usami, Satoshi Inoue*, Masahide Ebi*, Naotaka Ogasawara* and Kunio Kasugai*

By realizing the 3D-shape recovery of the image inside of the human body taken by the endoscope, the quantitative evaluation of polyps becomes possible. Polyp detection, classification, and shape recovery methodologies are required based on endoscope images taken by "general" endoscope images. Against "general" endoscope images, this research proposes an efficient polyp classification method using a convolutional neural network trained in advance for other tasks as a feature extractor and construct multiple SVM (Support Vector Machine) as a classifier. Furthermore, for polyps which are difficult to classify benign or malignant, a robust polyp shape recovery method is proposed based on blood vessel information.

Keywords : *Deep Learning, CNN, Transfer Learning, Shape from Shading, U-net*

1. はじめに

内視鏡から得られた人体内部の画像を用いて、胃壁などの形状復元を行うことで、腫瘍などの形状に関する定量的な情報を得ることが可能になる。内視鏡画像診断の立場からは通常内視鏡でポリープ検出、分類、形状復元を行う手法が望ましい。本研究では、通常内視鏡を用いて撮影された内視鏡画像に対して、他のタスクで事前に学習済みの畳み込みニューラルネットワーク (Convolutional Neural Network。以下、CNN)¹⁾を活用することで、比較的少ない内視鏡画像から Deep Learning を用いたポリープの分類を実現する。先の手法で分類の困難な画像に対しては血管情報をもとに、2次元である内視鏡画像からポリープの3次元形状復元を獲得可能にする手法を開発することで、医師の診断支援を目的としたコンピュータ支援診断を実現する。

1.1. ポリープ分類手法について

ポリープの分類手法として、機械学習をベースとした超拡大内視鏡を用いた複数の手法²⁾³⁾では、高精度なポリープの分類を実現している。しかしながら、これらの手法は特殊な内視鏡を必要とし、今日の医療の現場での汎用

* 愛知医科大学医学部消化管内科

性は低い。一般的な内視鏡で撮影可能な画像は、白色光源を用いた画像、染色液を用いた染色画像、血管の領域を明瞭にするための狭帯域光源を用いた画像の三種類であり、これら汎用的な内視鏡で撮影される画像からの高精度なポリープ分類手法の実現が医療の現場では望まれている。画像処理の分野では、CNN に端を発する Deep Learning が演算装置の高性能化とともに「大規模かつ良質なデータセット」⁴⁾が提供されることで飛躍的高精度での画像分類が実現可能であることが証明されている^{5) 6) 7)}。ポリープ分類を目的とした内視鏡画像のデータセットは少なく、特殊な装置で撮影した拡大 (Narrow Band Imaging) NBI 画像における毛細血管のパターン特徴量を利用することで少量のデータセットから効果的なポリープの良性・悪性分類を実現する手法⁸⁾が提案されているが、やはり一般的な内視鏡で撮影された内視鏡画像を対象としたデータセットはほとんどない。画像処理の分野では、このような問題を解決するために、全く異なるドメインのタスクで事前学習済みの CNN を特徴抽出器とし、対象となるドメインのタスクへ適応し、抽出された特徴量に対して分類を目的とする (Support Vector Machine) SVM を構築することで、少量の学習データセットでの効果的な分類を実現する手法^{9) 10)}が提案されている。

1.2. ポリープの 3 次元形状情報について

ポリープ診断ガイドラインは内視鏡医療における診断精度向上を目的として 1993 年 (初版)¹¹⁾から複数のエビデンス^{12) 13) 14)}をベースとして改訂が行われているが、90 年代末では、ポリープの大きさとその形状情報はポリープを良性・悪性と診断する際の重要な情報であることが明確に示されている。

画像処理の技術を用いて、内視鏡画像からポリープの 3 次元形状復元を獲得する場合、内視鏡で撮影される物体がポリープ等の比剛体物体で構成されるため、その手法としては、Shape from Shading (SfS)¹⁵⁾等の光学ベースでの手法が内視鏡撮影環境では妥当であると考えられる。SfS は、画像の輝度と陰影情報をもとに対象物体の表面の傾きとカメラからの奥行きマップを獲得する。内視鏡画像を対象とした SfS ベースの形状情報獲得手法としては、内視鏡光源が 1 つの光源で構成される撮影環境を仮定した光学的制約式と幾何学的制約式を用いる手法¹⁶⁾、実際の内視鏡が光源を 2 つ有することからより実際の撮影環境に制約式を拡張した手法¹⁷⁾がポリープの 3 次元形状復元として提案されている。また、手法¹⁸⁾では SfS をベースとした手法により背骨に対する 3 次元形状復元を実現している。一見、SfS ベースの手法は内視鏡撮影環境における 3 次元形状復元手法として高い親和性を持つが、手法を適用させるためには、内視鏡カメラから撮影物体までの奥行き Z 、物体表面の反射係数 C など複数の初期パラメータを必要とする。

そこで、本研究では、撮影された内視鏡画像に写る血管情報を活用することで、SfS を適用するために必要となる初期パラメータの獲得を実現することで、一枚の内視鏡からのポリープの大きさ形状復元を行うより汎用的な手法を実現する。

2. 提案手法

本研究では、一般的な内視鏡で撮影される内視鏡画像を対象に、CNN を特徴抽出器、SVM を分類器として使用することで、一般的な内視鏡で撮影された少量の内視鏡画像データから効果的に良性・悪性ポリープ分類を行う手法を提案する。また、ポリープの良性・悪性分類が困難な内視鏡画像に対しては、血管情報を活用することで、一枚の内視鏡からのポリープの大きさ形状復元を行うより汎用的な手法を提案する。以下に、それぞれの詳細について述べる。

2.1. CNN を用いたポリープ分類

提案手法の概念を図 1 に示す。内視鏡撮影環境とは異なるドメインで、一般物体認識を目的とし事前に学習を行なった CNN を画像特徴量抽出器として用いることで内視鏡画像から特徴抽出を行い、得られた特徴量に対してカーネル関数を適用し、良性・悪性ポリープの特徴空間への特徴量を写像することで SVM により良性・悪性ポリープの分類を実現する。また、ポリープ分類は、同一ポリープに対して撮影した 3 種類の内視鏡画像それぞれから

CNN により特徴抽出を行い、3 種類の画像の組合せから 7 種類の SVM を構築し、7 つの SVM の投票処理により最終的な良性・悪性ポリープを実現する。

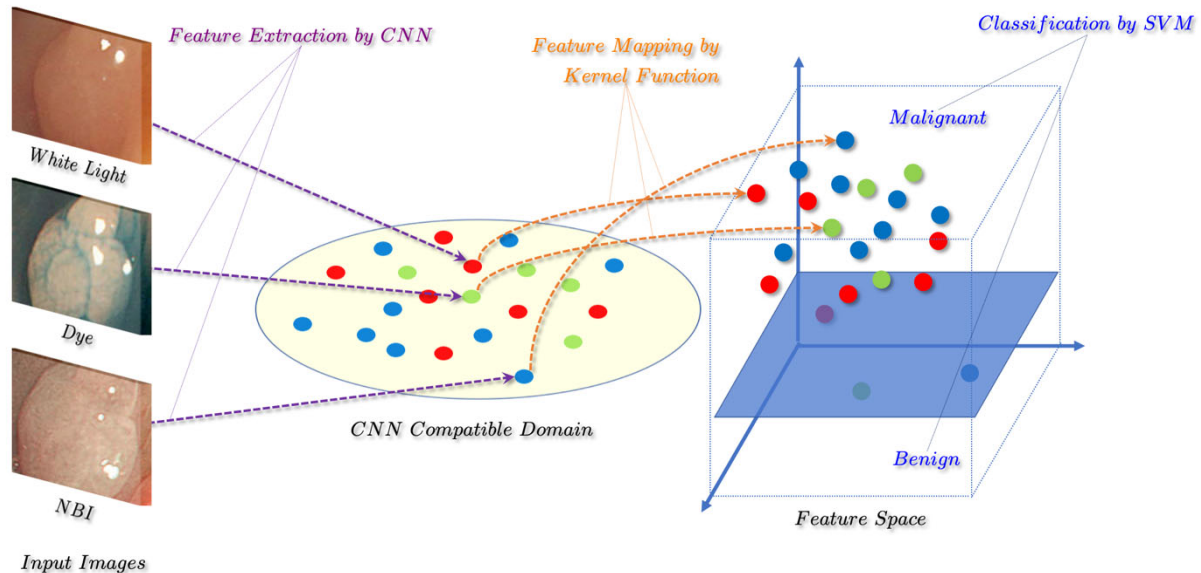


Figure 1. Concept of The Proposed Method

2. 1. 1. 3 種類の内視鏡画像

ポリープの分類は、対象ポリープの 3 種類の内視鏡画像（白色光、染色、狭帯域光）を用いて行う。各タイプの内視鏡画像は、一般的な内視鏡を使用して撮影でき、以下の特徴を有する。

白色光 (White Light) : 通常モードで撮影された白色光源画像

染色 (Dye) : ポリープの病変を観察するためのインディゴカルミン色素による染色画像

狭帯域光 (NBI) : 狭帯域イメージングによる画像。特定の青と緑の波長の光を使用して、血管の細部を強調

2. 1. 2. CNN 特徴抽出器

本研究では、内視鏡画像の特徴抽出器として、AlexNet¹⁹⁾と VGG16²⁰⁾の 2 種類の CNN ネットワークアーキテクチャを用いた。加えて、ImageNet⁴⁾を使用して各 CNN は事前学習済みのネットワークを用いる。

AlexNet は ImageNet Large Scale Visual Recognition Challenge (ILSVRC)²¹⁾2012 で提案された CNN であり、8 層の畳み込みニューラルネットワークであり、本研究では、全結合層の 7 層目の特徴抽出器として活用する。

VGG16 は ILSVRC2014 で提案された 13 層の畳み込み層と 3 層の全結合層の計 16 層からなる Deep Convolutional Neural Network (DCNN)であり、本研究では、15 層目の全結合層の特徴抽出器として活用する。

2. 1. 3. 各特徴量に最適化された SVM 分類器の構築

CNN を使用した特徴抽出は、3 種類の内視鏡画像それぞれに適用する。この研究では、CNN により抽出した各内視鏡画像の特徴空間に対してカーネル関数を適用し、新たに良性・悪性分類を目的とした特徴空間へ写像することにより、SVM によるポリープの良性・悪性分類を行う。カーネル関数を適用した内視鏡画像の組み合わせには、7 つの組合せ（白色光、染色、NBI、白色光+染色、白色光+ NBI、染色+ NBI、白色光+染色+ NBI）があり、それぞれの組合せに対して、3 種類のカーネル関数を適用し、より効果的な条件の検証を行った。

2.1.4. 複数 CNN-SVM の投票処理によるポリープ分類

ポリープ分類は、2.1.3.で構築された 7 つの CNN-SVM 分類器の投票処理にもとづき行う。投票処理に使用する値は式 (1) により算出する。ここで、式 (1) 内の各「CNN-SVM」は式 (2) の条件により与えられる。

$$Score = \sum_{n=1}^7 CNN - SVM_n \quad (1)$$

$$CNN - SVM_n = \begin{cases} 1 & \text{if } Decision = Malignant \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

最終的に、式(1)から導出される値が過半数である 4 を超える場合に、該当ポリープを悪性ポリープと判断し、分類を行う。

2.2. 血管を利用したポリープ形状復元

本研究では、ポリープの良性・悪性分類が困難な内視鏡画像に対しては、血管情報を活用することで、一枚の内視鏡からのポリープの大きさ形状復元を行う、より汎用的な手法を提案する。形状は以下の手順で復元する。

- Step1. 2 光源内視鏡による観測システムを仮定
 - Step2. カメラキャリブレーションによる内視鏡の内部パラメータの取得
 - Step3. 鏡面反射領域除去
 - Step4. U-net を用いた血管領域の抽出
 - Step5. 血管領域の水平面を利用し、血管の反射係数 C ならびにカメラから血管までの奥行き Z の獲得
 - Step6. 血管と腸内壁の境界面を利用し、腸内壁の反射係数 C ならびにカメラから血管までの奥行き Z の獲得
- それぞれのステップの詳細を以下、2.2.1.から 2.2.6.の項に示す。

2.2.1. 2 光源内視鏡における観測システム

形状復元は、対象とする内視鏡が 2 光源を有し、図 2 に示す点光源透視投影の観測システムを仮定して行う。この過程のもとで、撮影物体の輝度 E ならびに反射係数 C の関係は、式(3)から(11)の光学的制約式で表すことができ、カメラレンズから撮影物体までの奥行き Z を導出することで形状復元を実現する。

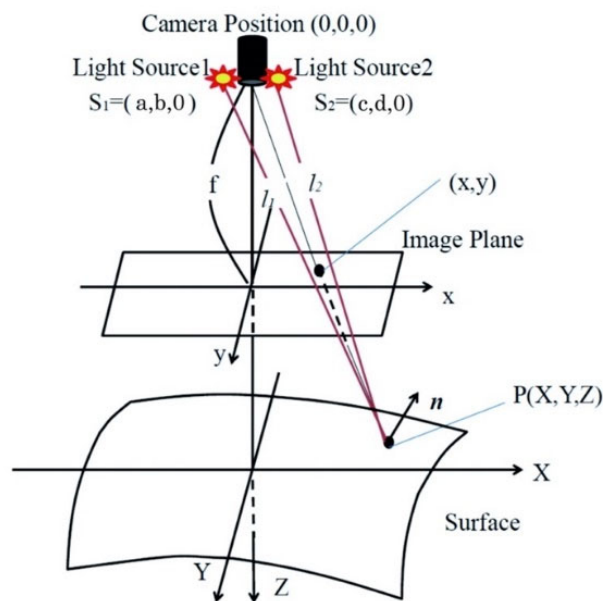


Figure 2. Observation System of Two Light Sources Endoscope

以下に、撮影物体の輝度 E ならびに反射係数式 C とカメラレンズから撮影物体までの奥行 Z の関係式を示す。

カメラレンズの座標を $(0, 0, 0)$ とし、第1光源に対する光源方向ベクトルを $\mathbf{s}_1$ 、第2光源に対する光源方向ベクトルを $\mathbf{s}_2$ 、第1光源から撮影物体表面までの距離を l_1 、第2光源から撮影物体表面までの距離を l_2 、撮影物体表面の法線ベクトルを $\mathbf{n}$ とすることで、撮影物体の輝度 E は、その反射係数式 C を用いて式(3)のように表すことができる。

$$E = C \left(\frac{\mathbf{n} \cdot \mathbf{s}_1}{l_1^2} + \frac{\mathbf{n} \cdot \mathbf{s}_2}{l_2^2} \right) \quad (3)$$

撮影物体表面の法線ベクトル $\mathbf{n}$ は、物体表面の傾き p, q を用いることで式(4)のように表すことができる

$$\mathbf{n} = \frac{(p, q, -1)}{\sqrt{p^2 + q^2 + 1}} \quad (4)$$

ここで物体表面の傾き p, q はそれぞれ、式(5)に示すように Z に対する X 軸面、 Y 軸の偏微分により導出できる。

$$(p, q) = \left(\frac{\partial Z}{\partial X}, \frac{\partial Z}{\partial Y} \right) \quad (5)$$

また、第1光源から撮影物体表面までの距離 l_1 、第2光源から撮影物体表面までの距離 l_2 はそれぞれ式(6)、(7)のよう表される。

$$l_1 = \sqrt{(a - X)^2 + (b - Y)^2 + Z^2} \quad (6)$$

$$l_2 = \sqrt{(c - X)^2 + (d - Y)^2 + Z^2} \quad (7)$$

第1光源に対する光源方向ベクトル $\mathbf{s}_1$ 、第2光源に対する光源方向ベクトル $\mathbf{s}_2$ はそれぞれ式(8)、(9)のよう表される。

$$\mathbf{s}_1 = \frac{(a - X, b - Y, -Z)}{\sqrt{(a - X)^2 + (b - Y)^2 + Z^2}} \quad (8)$$

$$\mathbf{s}_2 = \frac{(c - X, d - Y, -Z)}{\sqrt{(c - X)^2 + (d - Y)^2 + Z^2}} \quad (9)$$

焦点距離を f とするとカメラ座標系の x, y は式(10)により世界座標系の X, Y に変換できる。

$$X = \frac{x}{f}Z, \quad Y = \frac{y}{f}Z \quad (10)$$

式(3)に式(4)から式(10)を代入することで、撮影物体の輝度 E は、その反射係数 C とカメラレンズから撮影物体までの奥行 Z を用いて式(11)のように表すことができる。

$$E = \frac{C}{\sqrt{p^2 + q^2 + 1}} \left(\frac{-p \left(a - \frac{x}{f}Z \right) - q \left(b - \frac{y}{f}Z \right) + Z}{\left(\left(a - \frac{x}{f}Z \right)^2 + \left(b - \frac{y}{f}Z \right)^2 + Z^2 \right)^{\frac{3}{2}}} + \frac{-p \left(c - \frac{x}{f}Z \right) - q \left(d - \frac{y}{f}Z \right) + Z}{\left(\left(c - \frac{x}{f}Z \right)^2 + \left(d - \frac{y}{f}Z \right)^2 + Z^2 \right)^{\frac{3}{2}}} \right) \quad (11)$$

本研究では、先の観測システムならびに光学的制約式を用いてポリープの形状復元を行う。

2.2.2. カメラキャリブレーション

使用する内視鏡の焦点距離等の内部パラメータは、図3に例を示すような内視鏡で撮影したチェッカーボードに対して、カメラキャリブレーション手法^{22) 23)}を実行することで取得する。



Figure 3. Examples of Checker Board

また、加えて図 2 に示した内視鏡の 2 光源の光源位置 $S_1 = (a, b, 0)$ 、 $S_2 = (c, d, 0)$ は、反射率一様な球画像を撮影することで事前に取得する。

2.2.3. 鏡面反射領域の除去

反射光は鏡面反射と拡散反射の 2 つの成分で構成され、物体表面からの光の反射スペクトルエネルギーの分布は 2 色性反射モデル²⁴⁾ を用いることで、照明のスペクトルエネルギー分布と反射率との積として表すことができる。鏡面反射成分は陰影情報を利用した SfS において形状復元結果に影響を与え、内視鏡画像においても頻繁に発生する。鏡面反射成分に対してインペインティング処理を行うことで鏡面反射成分の除去を行う手法²⁵⁾ が提案されているが、インペインティング処理で補完された領域は情報が欠落し、本研究では血管と腸内壁の境界面を形状復元に利用するため、よりエッジを残しつつ鏡面反射領域除去を実現する手法が望ましい。本研究では、よりエッジ情報を残しつつ鏡面反射除去を実現するバイラテラルフィルタ^{26) 27)} を利用した鏡面反射領域の除去を行った後、提案する形状復元手法を適用する。

2.2.4. 血管領域のセグメンテーション

提案手法では、はじめに血管情報を用いて血管領域の奥行き Z と反射係数 C を算出する。そのため、血管領域のセグメンテーションを行う。画像のセグメンテーション手法としては、複数の手法^{28) 29) 30)} が提案されているが、医用画像セグメンテーション用に開発された U-net³⁰⁾ を基礎として図 4 に示すダウンサンプリングとアップサンプリング層を有する畳み込みニューラルネットワークを構築し、血管領域のセグメンテーションを行う。

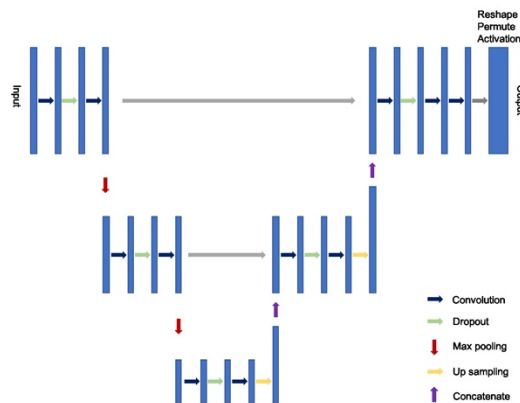


Figure 4. Network Architecture

ここで、提案するネットワークでは、損失関数をクロスエントロピー、最適化に確率的勾配降下を採用し、畳み込み層の各層の活性化関数に ReLU 関数を適用し、過学習防止を目的として、畳み込み層が連続する2層間に 0.2 のドロップアウトを適用する。

2.2.5. 奥行きマップの取得

内視鏡画像は、血管と腸内壁の領域から構成され、それぞれ表面の反射率が異なるため、カメラから物体までの奥行き Z を血管、腸内壁とそれぞれの領域に対して取得することで形状復元を実現する。

2.2.5.1. 水平面

血管は円柱状の形状を有しており、その幾何学的な特徴を活用することで血管領域における奥行き Z 、反射係数 C を導出する。その手順ではまず、レンズに対する血管の水平面を求めることで幾何学的制約を適用可能とする。血管は円柱状の形状を有していることから図 5 に示すように、レンズに対して水平面で切り取られた円柱の幅は、レンズに対して水平な区間では同じ幅となる。この特性を用いて、血管の水平面を取得する。

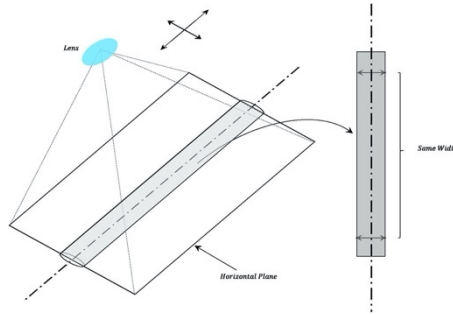


Figure 5. Horizontal Plane of Columnar against lens

血管の水平面取得手順を内視鏡画像に適用した例を図 6 から図 10 に示す。はじめに図 6 のオリジナル画像に対して、血管領域の抽出を行い図 7 に示す画像を得る。次に、血管領域の細線化処理を施すことで図 8 に示す画像を得る。また、エッジ抽出処理を行うことで図 9 の画像を得る。血管領域を細線化した図 8 は血管の中心線を表し、水平面の取得はこの中心線をトレースしながら中心線に対して直交する線を引き、血管の両端のエッジで切り取られる幅を取得し、同じ幅が連続する区間を水平面として抽出する。

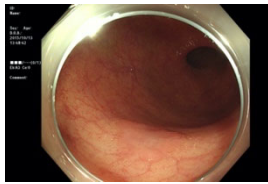


Figure 6. Original Image

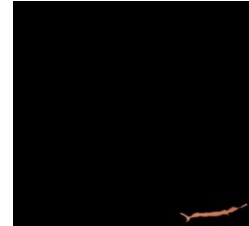


Figure 7. Example of Extracted Blood Vessel Region



Figure 8. Example of Line Thinning Processing



Figure 9. Example of Edge Extraction

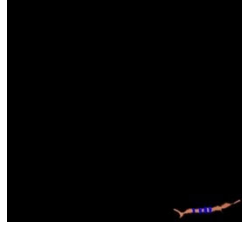


Figure 10. Example of Estimation of Horizontal Plane

2.2.5.2. 血管

得られた血管の水平面を利用して血管の奥行 Z を取得するための観測モデルを図 11 に示す。この幾何学制約を利用した奥行 Z の導出手順を以下に示す。

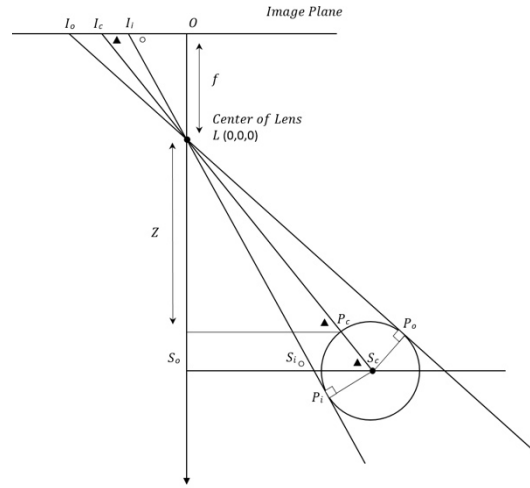


Figure 11. Observation System of Horizontal Plane

$\triangle LOI_i \sim \triangle LS_oS_i$ と $\angle LS_iS_o$ が $\triangle LS_oS_i$ の外角であることから $\angle LS_iS_o$ は式(10)のように示すことができる。

$$\angle LS_iS_c = \pi - \angle LI_iO \quad (12)$$

$\triangle LOI_c \sim \triangle LS_oS_c$ から $\angle LS_cS_i$ は式(13)のように表すことができる。

$$\angle LS_cS_i = \angle LI_cO \quad (13)$$

$\angle S_iLS_c$ は、 $\angle LS_cS_i$ の外角であることから式(14)のように表すことができる。

$$\angle S_iLS_c = \pi - \angle LS_cS_i \quad (14)$$

$\angle LS_cP_i$ は、 $\angle S_iLS_c$ を用いることで式(15)から導出できる。

$$\angle LS_cP_i = \pi - \frac{\pi}{2} - \angle S_iLS_c \quad (15)$$

$\triangle LS_oS_c$ の斜辺 LS_c は $\angle LS_cP_i$ を用いて式(16)のように表すことができる。

$$LS_c = \frac{P_iS_c}{\cos \angle LS_cP_i} \quad (16)$$

P_cS_c が血管の半径であることから LP_c は式(17)のように表すことができる。

$$LP_c = LS_c - P_cS_c \quad (17)$$

最終的に Z は辺 LP_c と $\angle LI_cO$ を用いて式(18)のように表すことができる。

$$Z = LP_c \sin \angle LI_cO \quad (18)$$

ここで、傾き p 、 q は xz 面、 yz 面に注目することでそれぞれ以下の式により導出できる。

傾き p の場合では、 xz 面に注目し、点 $L = (0,0)$ と点 $P_c = (\frac{xI_c}{f}Z, Z)$ を通る線に直交する線の傾きを求めることで得られる。点 L と P_c を通る線の傾きは式(19)のように表すことができる。

$$\frac{\frac{Z}{xI_c}}{\frac{f}{Z}} = \frac{f}{xI_c} \quad (19)$$

点 L と P_c を通る線と直交する線の傾きを式(20)のように求めることで傾き p を導出することができる。

$$p = -\frac{xI_c}{f} \quad (20)$$

傾き q の場合では、 yz 面に注目し、点 $L = (0,0)$ と点 $P_c = (\frac{yI_c}{f}Z, Z)$ を通る線に直交する線の傾きを求めることで得られる。点 L と P_c を通る線の傾きは式(21)のように表すことができる。

$$\frac{\frac{Z}{yI_c}}{\frac{f}{Z}} = \frac{f}{yI_c} \quad (21)$$

点 L と P_c を通る線と直交する線の傾きを式(22)のように求めることで傾き q を導出することができる。

$$q = -\frac{yI_c}{f} \quad (22)$$

最終的に血管の輝度 E は、得られた Z 、 p 、 q を式(11)に代入することで導出することができる。ここで、血管の C は血管の水平面において血管幅が連続する区間について、 Z のSTDが低い区間から順に C の算出を行う。 C の算出については Smirnov-Grubbs 検定³¹⁾を実施することで外れ値対策を行い、最終的にMEDIANをとることで血管の反射係数 C とする。

2.2.5.3. 腸内壁

血管と腸内壁が連続面を持つと仮定すると、腸内壁の奥行 Z は図12に示す血管と腸内壁の境界線について血管の既知情報を拡張することで導出することができる。ここで、図12の赤色の領域を血管、黄色の領域を腸内壁とする。

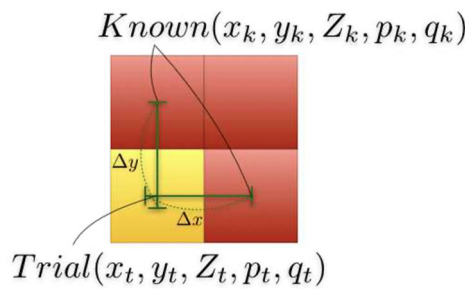


Figure 12. Boundary Surfaces between Blood Vessel and Intestinal-Inner-Wall

血管の領域は既知であるとし、 $Known$ 点の集まりであるとして、血管に隣接する腸内壁の点は $Trial$ 点として式(23)のように近似することができる。

$$Z_t = Z_k + p\Delta X + q\Delta Y \quad (23)$$

ΔX 、 ΔY はそれぞれ、 x 軸、 y 軸に考慮することで、世界座標系のパラメータより、式(24)のように表すことができる。

$$Z_t = Z_k + p\left(x_t \frac{Z_t}{f} - x_k \frac{Z_k}{f}\right) + q\left(y_t \frac{Z_t}{f} - y_k \frac{Z_k}{f}\right) \quad (24)$$

式(24)を Z について解くことで、式(25)を導出することができる。

$$Z_t \left(1 - px_t \frac{1}{f} - py_t \frac{1}{f} \right) = Z_k \left(1 - px_k \frac{1}{f} - py_k \frac{1}{f} \right) \quad (25)$$

最終的に Z_t を Z_{expect} とし、式(25)を Z_{expect} について解くことで、式(26)を導出できる。

$$Z_{expect} = \frac{Z_k(f - px_k - qy_k)}{f - px_t - qy_t} \quad (26)$$

式(26)より腸内壁の奥行 Z を算出し、式(11)を C について解いた式(25)に得られたパラメータを代入することで算出できる。

$$\begin{aligned} f(p, q, a, b, c, d, f, Z) \\ = \frac{1}{\sqrt{p^2 + q^2 + 1}} \left(\frac{-p \left(a - \frac{x}{f} Z \right) - q \left(b - \frac{y}{f} Z \right) + Z}{\left(\left(a - \frac{x}{f} Z \right)^2 + \left(b - \frac{y}{f} Z \right)^2 + Z^2 \right)^{\frac{3}{2}}} \right. \\ \left. + \frac{-p \left(c - \frac{x}{f} Z \right) - q \left(d - \frac{y}{f} Z \right) + Z}{\left(\left(c - \frac{x}{f} Z \right)^2 + \left(d - \frac{y}{f} Z \right)^2 + Z^2 \right)^{\frac{3}{2}}} \right) \end{aligned} \quad (27)$$

また、内視鏡画像の腸内壁における実際の輝度を E_i とし、式(28)を目的関数として C^{opt} を解くことでより妥当な腸内壁の反射係数 C を算出する。

$$C^{opt} = \arg \min_c \sum_{i=1}^n \|E_i - Cf(p, q, a, b, c, d, f, Z)\| \quad (28)$$

3. 実験

3.1. CNN を用いたポリープ分類

3.1.1. 実験条件

実験で使用した学習用とテスト用のデータセットをそれぞれ表 1、表 2 に示す。これらのデータセットに加え、データを増やす目的で、元画像に 3 種類の回転処理を施した画像を追加した。また、学習画像のデータセットのラベルは不均衡であるため、Under-sampling を適用し、モデルの学習を行なった。

Table1 Train Image			Table2 Test Image		
	<i>Malignant</i>	<i>Benign</i>		<i>Malignant</i>	<i>Benign</i>
White Light	188	380	White Light	180	180
Dye	112	408	Dye	180	180
NBI	32	140	NBI	180	180

分類精度の評価として利用した、Sensitivity、Specificity、Accuracy、PPV、NPV の導出式はそれぞれ式(29)から(33)に示すとおりである。

$$Sensitivity = \frac{TP}{TP + FN} \quad (29)$$

$$Specificity = \frac{TN}{TN + FP} \quad (30)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (31)$$

$$PPV = \frac{TP}{TP + FP} \quad (32)$$

$$NPV = \frac{TN}{TN + FN} \quad (33)$$

3.1.2. 評価

CNN と SVM カーネルの組み合わせの評価は、各内視鏡画像の組み合わせから抽出されたドメインに対して行う。合計 6 つの組合せ、AlexNet と線形カーネル、AlexNet と RBF カーネル、AlexNet と多項式カーネル、VGG16 と線形カーネル、VGG16 と RBF カーネル、VGG16 と多項式カーネルに対して評価実験を行い、最も良い CNN と SVM カーネルの組合せを用いて投票処理を実施する。各結果を表 3 から表 8 に示す。ここで、C を SVM のコストパラメーター、 σ を SVM のカーネルパラメーターとした。

Table3. Result of AlexNet with Linear Kernel

	C	Sensitivity	Specificity	Accuracy	PPV	NPV
White Light	51	74.4	75.5	75.0	75.2	74.7
Dye	48	80.5	77.5	79.1	78.3	80.0
NBI	100	91.1	88.8	90.0	89.1	90.9
White Light + Dye	78	83.3	75.0	79.1	76.9	81.8
White Light + NBI	51	87.7	85.5	86.6	85.8	87.5
Dye + NBI	8	89.4	83.8	86.6	84.7	88.8
White Light + Dye + NBI	90	87.7	83.3	85.5	84.0	87.2

Table4. Result of AlexNet with RBF Kernel

	C	σ	Sensitivity	Specificity	Accuracy	PPV	NPV
White Light	37	145	78.3	76.6	77.5	77.5	77.0
Dye	8	170	83.3	83.3	83.3	83.3	83.3
NBI	63	149	90.0	90.5	90.5	90.2	90.5
White Light + Dye	26	236	83.3	78.8	78.8	81.1	79.7
White Light + NBI	9	223	85.0	85.5	85.5	85.2	85.4
Dye + NBI	12	238	88.8	87.7	87.7	88.3	87.9
White Light + Dye + NBI	35	292	87.7	85.0	85.0	86.3	85.4

Table5. Result of AlexNet with Polynomial Kernel

	C	σ	Sensitivity	Specificity	Accuracy	PPV	NPV
White Light	9	142	73.3	76.6	75.0	77.8	74.1
Dye	9	162	78.3	80.0	79.1	79.6	79.1
NBI	40	148	92.7	89.4	91.1	89.7	92.5
White Light + Dye	42	237	80.0	79.4	79.7	79.5	79.8
White Light + NBI	2	224	87.2	85.0	86.1	85.3	86.9
Dye + NBI	78	238	89.4	85.5	87.5	86.0	89.0
White Light + Dye + NBI	3	292	89.4	85.5	87.5	86.0	89.0

Table6. Result of VGG16 with Linear Kernel

	C	Sensitivity	Specificity	Accuracy	PPV	NPV
White Light	7	74.4	73.3	73.8	73.6	74.1
Dye	40	78.3	82.2	80.2	81.5	79.1
NBI	10	87.7	90.0	88.8	89.7	88.0
White Light + Dye	95	84.4	77.2	80.8	78.7	83.2
White Light + NBI	35	87.2	88.8	88.0	88.7	87.4
Dye + NBI	42	88.3	90.0	89.1	89.8	88.5
White Light + Dye + NBI	82	89.4	90.5	90.0	90.4	89.5

Table7. Result of VGG16 with RBF Kernel

	C	σ	Sensitivity	Specificity	Accuracy	PPV	NPV
White Light	5	112	82.2	76.6	79.4	77.8	81.1
Dye	5	135	79.4	81.6	80.5	81.2	79.8
NBI	19	127	88.3	94.4	91.3	94.0	89.0
White Light + Dye	15	192	81.6	77.7	79.7	78.6	80.9
White Light + NBI	4	185	88.8	91.1	90.0	90.9	89.1
Dye + NBI	11	198	88.8	90.0	89.4	89.8	89.0
White Light + Dye + NBI	4	245	87.2	88.3	87.7	88.2	87.3

Table8. Result of VGG16 with Polynomial Kernel

	C	σ	Sensitivity	Specificity	Accuracy	PPV	NPV
White Light	9	115	80.0	78.3	79.1	78.6	79.6
Dye	9	136	78.3	86.1	82.2	84.9	79.8
NBI	40	129	89.4	91.6	90.5	91.4	89.6
White Light + Dye	42	189	81.1	82.1	81.9	82.4	81.4
White Light + NBI	2	187	86.6	89.4	88.0	89.1	87.0
Dye + NBI	78	200	87.7	88.8	88.3	88.7	87.9
White Light + Dye + NBI	3	246	87.2	88.8	88.0	88.7	87.4

Table9. Result of AlexNet + VGG16 with Linear Kernel

	C	Sensitivity	Specificity	Accuracy	PPV	NPV
White Light	75	74.4	73.3	73.8	73.6	74.1
Dye	12	80.5	77.7	79.1	78.3	80.0
NBI	59	91.1	88.8	90.0	89.1	90.9
White Light + Dye	23	83.8	80.0	81.9	80.7	83.2
White Light + NBI	40	88.8	86.1	87.5	86.4	88.5
Dye + NBI	60	89.4	83.8	86.6	84.7	88.8
White Light + Dye + NBI	51	90.5	82.7	86.6	84.0	89.7

Table10. Result of AlexNet + VGG16 with RBF Kernel

	C	σ	Sensitivity	Specificity	Accuracy	PPV	NPV
White Light	5	112	82.2	76.6	79.4	77.8	81.1
Dye	25	173	83.3	83.3	83.3	83.3	83.3
NBI	11	147	90.0	90.5	90.2	90.5	90.0
White Light + Dye	6	214	86.1	81.6	83.8	82.4	85.4
White Light + NBI	21	200	87.2	88.3	87.7	88.2	87.3
Dye + NBI	9	239	88.8	87.7	88.3	87.9	88.7
White Light + Dye + NBI	5	273	89.4	86.1	87.7	86.5	89.0

Table11. Result of AlexNet + VGG16 with Polynomial Kernel

	C	σ	Sensitivity	Specificity	Accuracy	PPV	NPV
White Light	30	110	80.0	78.3	79.1	78.6	79.6
Dye	39	168	78.8	80.0	79.4	79.7	79.1
NBI	67	146	92.7	90.0	91.3	90.2	92.5
White Light + Dye	11	214	83.8	81.1	82.5	81.6	83.4
White Light + NBI	18	204	87.2	87.2	87.2	87.2	87.2
Dye + NBI	3	238	89.4	85.5	87.5	86.0	89.0
White Light + Dye + NBI	18	278	89.4	84.4	86.9	85.1	88.8

先の実験で得られた組合せを用いたポリープの良性・悪性の分類結果を表 12 に示す。実験結果より、最終的に 90%を超える分類精度を実現したことが確認できる。

Table12. Result of AlexNet + VGG16 with Linear Kernel

	Sensitivity	Specificity	Accuracy	PPV	NPV
White Light	82.2	76.6	79.4	77.8	81.1
Dye	83.3	83.3	83.3	83.3	83.3
NBI	92.7	90.0	90.2	90.2	92.5
White Light + Dye	86.1	81.6	83.8	82.4	85.4
White Light + NBI	88.8	86.1	87.5	86.4	88.5
Dye + NBI	89.4	85.5	87.5	86.4	89.0
White Light + Dye + NBI	90.5	82.7	86.6	84.0	89.7
Voting Result	94.4	91.6	93.3	84.0	94.8

3.2. 血管を利用したポリープ形状復元

3.2.1. カメラキャリブレーション

カメラキャリブレーションの結果、得られた焦点距離、主点、放射状歪みの値を表 13 に示す。形状復元はこれら内視鏡の内部パラメータを用いて実験を行う。

Table13. Result of Estimation	
<i>Parameter</i>	<i>Result of Calibration</i>
Focal length (pixels)	[718.7447 $\pm$ 0.8387, 718.3827 $\pm$ 0.8654]
Principal point (pixels)	[879.0439 $\pm$ 0.4669, 533.5813 $\pm$ 0.4240]
Radial distortion	[-0.3913 $\pm$ 0.0010, -0.1178 $\pm$ 0.0008]

また、球画像を使用して推定した内視鏡の光源位置はそれぞれ $S_1 = (a: -5.001, b: 1.996, 0)$ 、 $S_2 = (c: 5.001, d: 1.996, 0)$ となった。

3.2.2. 血管領域抽出

各ピクセルにおける血管確立は、精度向上を目的として複数の予測の平均値とした。平均値はそれぞれ上下左右に 5 ピクセルのストライドを行うことで、各ピクセルにおける値を取得した。評価は、真値として血管領域に手動でマスク画像を作成した内視鏡画像を用いた。図 13 は血管の検出結果を示しており、全体的な予測精度は 0.8432 であった。

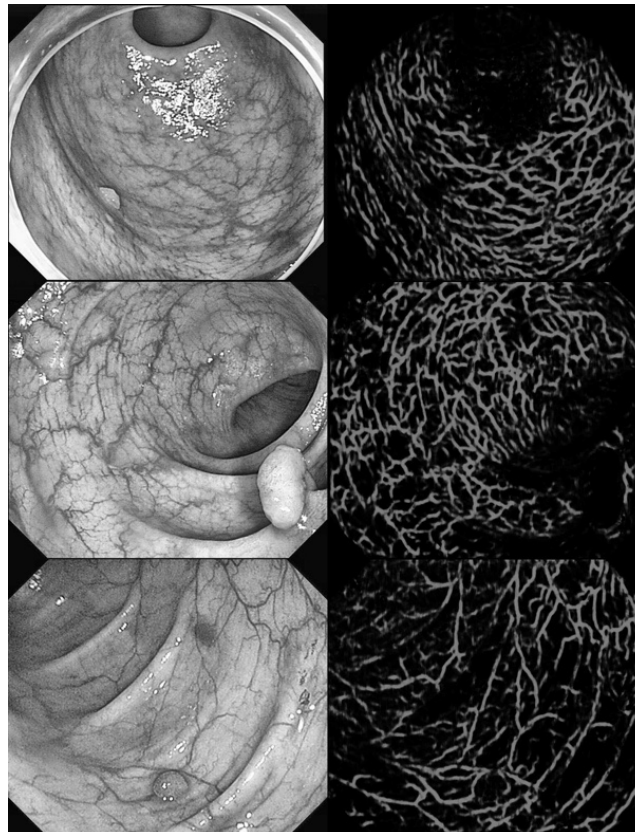


Figure 13. A Part of Prediction Results

図 14 に(Receiver Operating Characteristic)ROC 曲線を示す。実験結果より(Area Under the Curve)AUC は 0.9030 となった。

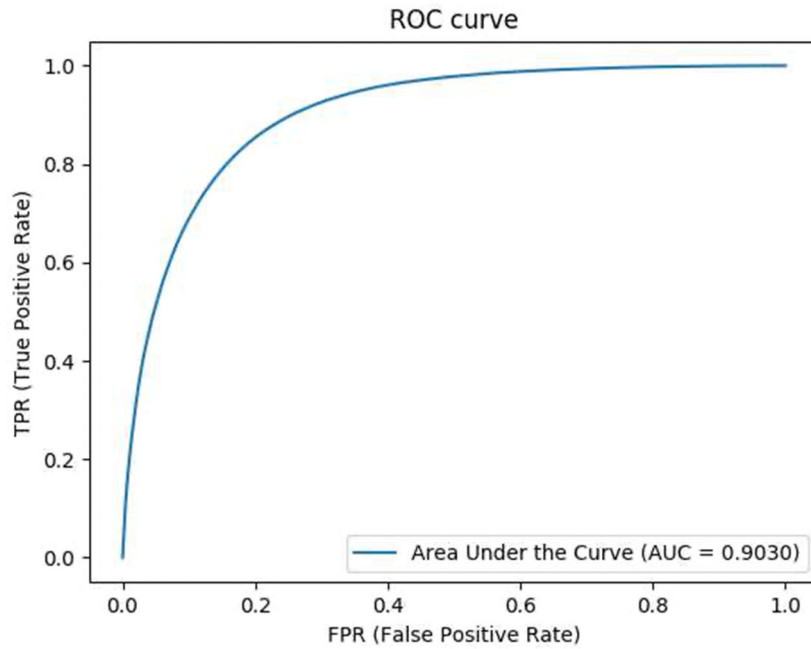


Figure 14. AUC ROC

3.2.3. ポリープの形状復元

血管情報と得られたパラメータを使用し、ポリープ形状復元実験を行う。図 15、19、23 に対象とした元画像を示す。図 16、20、24 に選択した血管の水平面領域を示す。図 17、21、25 に形状復元結果を、図 18、22、26 にそれぞれ計測を行なった抽出後のポリープ領域画像を示す。

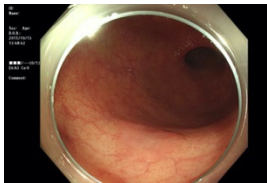


Figure 15. Original Image 1

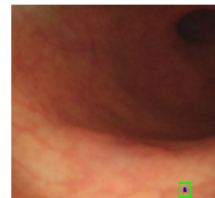


Figure 16. Selected horizontal region for Image 1

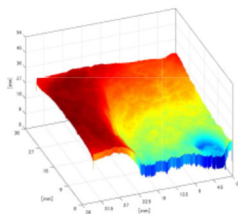


Figure 17. Recovered Shape for Image 1

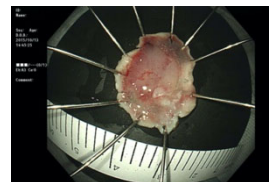


Figure 18. Size Measurement Result
in Image 1 after Extraction

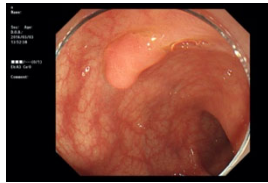


Figure 19. Original Image 2

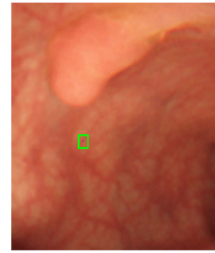


Figure 20. Selected horizontal region for Image 2

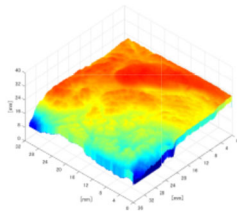


Figure 21. Recovered Shape for Image 2

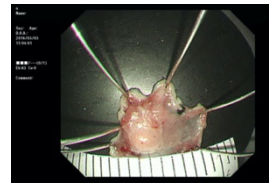


Figure 22. Size Measurement Result
in Image 2 after Extraction

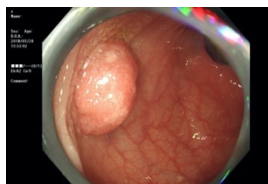


Figure 23. Original Image 3

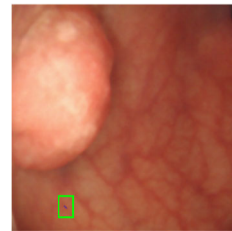


Figure 24. Selected horizontal region for Image 3

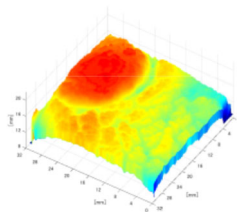


Figure 25. Recovered Shape for Image 3

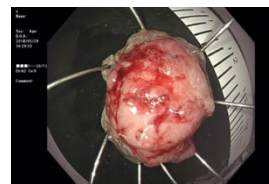


Figure 26. Size Measurement Result
in Image 3 after Extraction

血管領域の反射係数 C は、それぞれ 89217、118153、23074 となった。実験結果より、本手法による形状復元結果は摘出後の測定画像と比較して、おおよそのポリープ形状情報が得られていることが確認できる。

4. おわりに

本研究では、ポリープの良性・悪性分類に関しては、CNN を特徴抽出器として、SVM を分類器として使用することで、一般的な内視鏡で撮影された少量の内視鏡画像データから効果的に良性・悪性ポリープ分類を行う手法を提案した。7 種類の内視鏡画像の組合せに対してそれぞれ、適切な CNN ならびにカーネル関数のより効果的な組合せを検証し、最終的に 90%を超える分類精度を実現した。ポリープの形状復元に関しては、血管の幾何学的構造を活用することで SfS を適用する際の制約緩和を実現し、1枚の内視鏡画像からのポリープの形状復元を実現した。実験よりより少ない制約のもとでポリープの形状を復元できていることが確認できた。特に、一般的な内視鏡

で撮影された内視鏡画像から、反射率一様化処理など特別な前処理なく単一の画像からポリープの形状を復元できたことは、提案手法がより頑健なポリープ形状復元手法であることを示している。

謝辞

本研究は中部大学総合工学研究所 2019(平成31)年度プロジェクト、JSPS 科研費基盤研究(C)17K00252 ならびに研究活動スタート支援 19K24370 の援助を受け遂行されたものであり、ここに謝意を表します。

参考文献

- 1) LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural computation*, 1(4), 541-551.
- 2) Liao, Z., Gao, R., Xu, C., & Li, Z. S. (2010). Indications and detection, completion, and retention rates of small-bowel capsule endoscopy: a systematic review. *Gastrointestinal endoscopy*, 71(2), 280-286.
- 3) Misawa, M., Kudo, S. E., Mori, Y., Nakamura, H., Kataoka, S., Maeda, Y., Kudo, T., Hayashi, T., Wakamura, K., Miyachi, H., & Katagiri, A. (2016). Characterization of colorectal lesions using a computer-aided diagnostic system for narrow-band imaging endocytoscopy. *Gastroenterology*, 150(7), 1531-1532.
- 4) Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009, June). Imagenet: A large-scale hierarchical image database. In 2009 IEEE conference on computer vision and pattern recognition (pp. 248-255).
- 5) Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097-1105).
- 6) Sermanet, P., Eigen, D., Zhang, X., Mathieu, M., Fergus, R., & LeCun, Y. (2013). Overfeat: Integrated recognition, localization and detection using convolutional networks. *arXiv preprint arXiv:1312.6229*.
- 7) Zeiler, M. D., & Fergus, R. (2014, September). Visualizing and understanding convolutional networks. In *European conference on computer vision* (pp. 818-833). Springer, Cham.
- 8) Tischendorf, J. J. W., Wasmuth, H. E., Koch, A., Hecker, H., Trautwein, C., & Winograd, R. (2007). Value of magnifying chromoendoscopy and narrow band imaging (NBI) in classifying colorectal polyps: a prospective controlled study. *Endoscopy*, 39(12), 1092-1096.
- 9) Sharif Razavian, A., Azizpour, H., Sullivan, J., & Carlsson, S. (2014). CNN features off-the-shelf: an astounding baseline for recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops* (pp. 806-813).
- 10) Donahue, J., Jia, Y., Vinyals, O., Hoffman, J., Zhang, N., Tzeng, E., & Darrell, T. (2014, January). Decaf: A deep convolutional activation feature for generic visual recognition. In *International conference on machine learning* (pp. 647-655).
- 11) Bond, J. H. (1993). Polyp guideline: diagnosis, treatment, and surveillance for patients with nonfamilial colorectal polyps. *Annals of internal medicine*, 119(8), 836-843.
- 12) US Preventive Services Task Force, United States. Office of Disease Prevention, & Health Promotion. (1996). *Guide to clinical preventive services: report of the US Preventive Services Task Force*. US Department of Health and Human Services, Office of Public Health and Science, Office of Disease Prevention and Health Promotion.
- 13) Winawer, S. J., Fletcher, R. H., Miller, L., Godlee, F., Stolar, M. H., Mulrow, C. D., ... & Rosen, L. (1997). Colorectal cancer screening: clinical guidelines and rationale. *Gastroenterology*, 112(2), 594-642.
- 14) Byers, T., Levin, B., Rothenberger, D., Dodd, G. D., Smith, R. A., & American Cancer Society Detection And Treatment Advisory Group on Colorectal Cancer). (1997). *American Cancer Society guidelines for screening*

- and surveillance for early detection of colorectal polyps and cancer: update 1997. *CA: a cancer journal for clinicians*, 47(3), 154-160.
- 15) Horn, B. K., & Brooks, M. J. (Eds.). (1989). *Shape from shading*. MIT press.
 - 16) Iwahori, Y., Tatematsu, K., Nakamura, T., Fukui, S., Woodham, R. J., & Kasugai, K. (2015). 3d shape recovery from endoscope image based on both photometric and geometric constraints. In *Knowledge-Based Information Systems in Practice* (pp. 65-80). Springer, Cham.
 - 17) Usami, H., Hanai, Y., Iwahori, Y., & Kasugai, K. (2016, June). 3D shape recovery of polyp using two light sources endoscope. In *2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS)* (pp. 1-6). IEEE.
 - 18) Wu, C., Narasimhan, S. G., & Jaramaz, B. (2010). A multi-image shape-from-shading framework for near-lighting perspective endoscopes. *International Journal of Computer Vision*, 86(2-3), 211-228.
 - 19) Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097-1105).
 - 20) Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
 - 21) Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., ... & Berg, A. C. (2015). Imagenet large scale visual recognition challenge. *International journal of computer vision*, 115(3), 211-252.
 - 22) Heikkila, J., & Silven, O. (1997, June). A four-step camera calibration procedure with implicit image correction. In *Proceedings of IEEE computer society conference on computer vision and pattern recognition* (pp. 1106-1112). IEEE.
 - 23) Zhang, Z. (2000). A flexible new technique for camera calibration. *IEEE Transactions on pattern analysis and machine intelligence*, 22(11), 1330-1334.
 - 24) Shafer, S. A. (1985). Using color to separate reflection components. *Color Research & Application*, 10(4), 210-218.
 - 25) Tchoulack, S., Langlois, J. P., & Cheriet, F. (2008, June). A video stream processor for real-time detection and correction of specular reflections in endoscope images. In *2008 Joint 6th International IEEE Northeast Workshop on Circuits and Systems and TAISA Conference* (pp. 49-52). IEEE.
 - 26) Yang, Q., Wang, S., & Ahuja, N. (2010, September). Real-time specular highlight removal using bilateral filtering. In *European conference on computer vision* (pp. 87-100). Springer, Berlin, Heidelberg.
 - 27) Yang, Q., Tang, J., & Ahuja, N. (2014). Efficient and robust specular highlight removal. *IEEE transactions on pattern analysis and machine intelligence*, 37(6), 1304-1311.
 - 28) Carreira, J., & Sminchisescu, C. (2011). CPMC: Automatic object segmentation using constrained parametric min-cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(7), 1312-1328.
 - 29) Uijlings, J. R., Van De Sande, K. E., Gevers, T., & Smeulders, A. W. (2013). Selective search for object recognition. *International journal of computer vision*, 104(2), 154-171.
 - 30) Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention* (pp. 234-241). Springer, Cham.
 - 31) Grubbs, F. E. (1950). Sample criteria for testing outlying observations. *The Annals of Mathematical Statistics*, 21(1), 27-58.