

CNN特徴とSVM分類を用いた 大腸NBI拡大内視鏡診断支援システムの最適化

岡本 拓巳^{1,a)} 小出 哲士¹ 玉木 徹² Bisser Raytchev² 金田 和文² 吉田 成人³
三重野 寛³ 田中 信治⁴

概要: 本稿では大腸 Narrow Band Imaging (NBI) 拡大内視鏡がん診断支援システム実現のため, Convolutional Neural Network (CNN) 特徴量と Support Vector Machine (SVM) 分類による病理タイプ識別を行う手法の識別精度評価結果について述べる. 提案手法は研究グループで提案している Bag-of-Features (BoF) に基づく診断支援の特徴量抽出処理を, 学習済み AlexNet の結果を用いることで実現する. CNN 中間層データの特徴量として用いた場合の識別精度を評価し, 診断支援システムのリアルタイム処理実現に向けた検証を行う. 識別精度検証の結果, AlexNet における全結合層の 6, 7, 8 層にて非腫瘍・腫瘍の識別において医療現場からの要求性能以上である 95% の識別精度が実現可能であることを確認した.

Optimization for Computer-Aided Diagnosis System with CNN Feature and SVM classifier for NBI Colorectal Endoscopic Images

TAKUMI OKAMOTO^{1,a)} TETSUSHI KOIDE¹ TORU TAMAKI² BISSER RAYTCHEV² KAZUFUMI KANEDA²
SHIGETO YOSHIDA³ HIROSHI MIENO³ SHINJI TANAKA⁴

Abstract: In this paper, we discuss optimization for classifying Narrow-Band Imaging (NBI) colorectal endoscopic images with Convolutional Neural Network (CNN) and Support Vector Machine (SVM) for a Computer-Aided Diagnosis (CAD) system. The proposed method is using the result of pre-learned CNN as a feature extractor on Bag-of-Features framework. Moreover, we use the data of several layers in CNN as the extracted feature for SVM classifier. We estimated accuracy to identify tumor or not, and as an experimental result, we presented that the system accuracy is about 95% using feature of 6th, 7th and 8th fully-connected layers in AlexNet.

1. はじめに

近年, 国内外において大腸ガンの罹患数は年々増加傾向にある. しかし, 大腸ガンは早期ステージで発見, 治療が

できれば完治はほぼ可能である. そのための大腸内視鏡検査方法として, NBI (Narrow Band Imaging) システムを用いた大腸拡大内視鏡による画像検査が行われている. 大腸 NBI 拡大内視鏡診断では, 医師が大腸内壁の血管等の微細模様構造から腫瘍の有無やガン深達度を診断する. この診断を行う医師には専門知識と経験が必要となり, 診断できる医師に限られる.

そこで, コンピュータ画像解析を用いた症状の客観的な判断による診断の正確度向上や, 定量的な指標提示による若手医師に対する教育支援を目指したコンピュータ診断支援 (Computer-Aided Diagnosis: CAD) システムが必要とされている. 我々研究グループでは, 佐野分類 [1], [2], 広島分類 [3], 昭和分類 [4], 慈恵分類 [5] といった日本国内

¹ 広島大学 ナノデバイス・バイオ融合科学研究所
Research Institute for Nanodevice and Bio Systems, Hiroshima University

² 広島大学 工学研究院
Graduate School of Engineering, Hiroshima University

³ JR 広島病院 消化器内科
Department of Gastroenterology Hiroshima General Hospital of West Japan Railway Company

⁴ 広島大学病院 内視鏡診療科
Department of Endoscopy, Hiroshima University Hospital, Japan

a) koide-lab-info@ml.hiroshima-u.ac.jp

で提唱されている NBI 拡大観察所見分類をコンセンサスを経て統合された, JNET (Japan NBI Expert Team) 分類 [6] に基づき, 大腸内視鏡画像を図 1 に示す 4 病理タイプのうち 3 タイプ (Type 1, Type 2A, Type 3) に分類する大腸 NBI 拡大内視鏡画像診断支援システムの実現を目指している。

	Type 1	Type 2A	Type 2B	Type 3
Vessel pattern	● 認識不可	● 口径整 ● 均一な分布 (網目・らせん状)	● 口径不同 ● 不均一な分布	● 疎血管野領域 ● 太い血管の途絶
Surface pattern	● 規則的な黒色または白色点 ● 周囲の正常粘膜と類似	● 整(管状・樹枝状・乳頭状)	● 不整または不明瞭	● 無構造領域
Endoscopic image				

図 1 JNET 分類 [6].

診断支援システムに求められる医療現場からの要求性能として, (I) 高速性と即応性 (スループット 1~5 fps かつ, レイテンシ 1 秒以内), (II) 高い識別精度 (腫瘍部位か, あるいは非腫瘍部位であるかの診断結果が, 医師による診断結果と 90% 以上一致すること) が求められている. これまでに我々研究グループはソフトウェア実装による診断支援システムを開発しており, 大腸 NBI 拡大内視鏡画像中央の 120 × 120 pixel 領域に対し処理速度 14.7 fps を達成し, 正診率約 97% を達成している [7]. また, システムの主要処理部に関して FPGA への実装に向けたハードウェアアルゴリズムの提案, 実装評価を行いリアルタイム処理が可能であることを示している [8], [9], [10].

以下, 2 節にてこれまで研究グループで提案されている, Bag-of-Features (BoF) に基づいた診断支援システムの概要について説明する. 3 節にて Convolutional Neural Network (CNN) の計算結果を特徴量として利用した Support Vector Machine (SVM) による診断支援システムについて説明する. 4 節にて, CNN 特徴と SVM 分類を用いた診断支援システムの組込みシステム化に向けて, CNN 中間層における特徴量抽出と SVM 分類の識別精度に関して評価し, 5 節で結論を述べる.

2. Bag-of-Features に基づく内視鏡画像診断支援システム

本節では, これまでに研究グループで提案されてきた, Bag-of-Features (BoF) に基づいた大腸 NBI 拡大内視鏡画像診断支援システムについて述べる.

システムの概要を図 2 に示す. オフラインで行う学習, 学習結果を基にオンラインで実行する識別の 2 つのフェーズが存在する. システムは (1) 特徴抽出部, (2) 特徴変換部, (3) タイプ識別部の 3 つのモジュールに大きく分ける

ことができる. システムは Bag-of-Features (BoF) と呼ばれる手法に基づいて識別を行っている. BoF とは文書検索を画像に応用した手法で, 画像の局所特徴量によって得られる特徴量ベクトルを 1 つの単語と見做し, その出現頻度によって識別を行う.

学習フェーズでは, 学習用に準備した画像と, 画像に対応したラベル (病理タイプの情報) を用いて学習を行う. 画像から特徴量を抽出し, 各タイプ 256 (= 2⁸) 個のクラスタにクラスタリングを行い, 各クラスタの中心を Visual-Word (VW) として保存しておく. そして, 学習用画像から抽出した特徴量から VW の出現頻度である VW ヒストグラムを作成する. 作成した VW ヒストグラムをタイプ識別部の Support Vector Machine へと入力し, 識別フェーズで必要な Support Vector (SV) を決定する.

識別フェーズでは入力された内視鏡映像のフレームから特徴量を抽出し, 得られた特徴量群を VW と照合して出現頻度によるヒストグラムを作成し, ヒストグラムデータに対してタイプ識別を行い, 結果を表示する.

研究グループによりシステムアルゴリズムの有用性に関する評価を行った結果, 腫瘍・非腫瘍の識別に有用性が確認できている [11]. 図 3 にその評価結果を示す. SVM の学習に Type 1 (非腫瘍) の画像 504 枚, Type 2A と 3 (腫瘍) の画像 1743 枚を使用し, 病理組織診断が得られた 118 症例の内視鏡観察時の画像をテストとして用いた. 識別結果は SVM の出力値 (SVM output value) によって得られ, 現在のシステムカットオフ値である 0.5 を境に腫瘍・非腫瘍の識別が可能であることが確認できている.

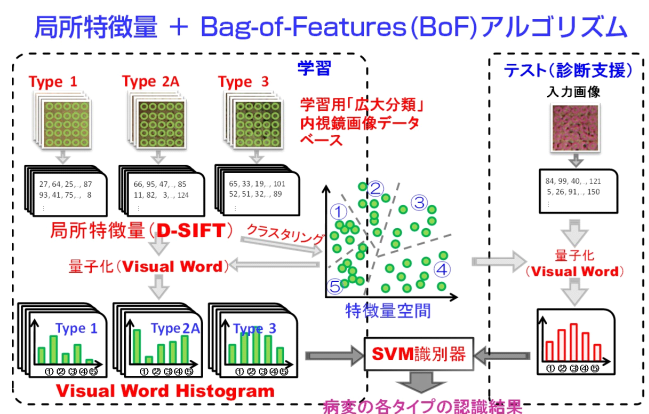


図 2 BoF に基づく大腸 NBI 拡大内視鏡画像がん診断支援システム.

2.1 Support Vector Machine (SVM)

Support Vector Machine (SVM) は, 教師あり学習により 2 クラス分類アルゴリズムの一つである. SVM は学習を行うことで, 図 4 に示すような識別空間において正と負の 2 タイプの距離 (マージン) が最大となるような識別超平面を決定する. その後新たに入力されるデータが識別超

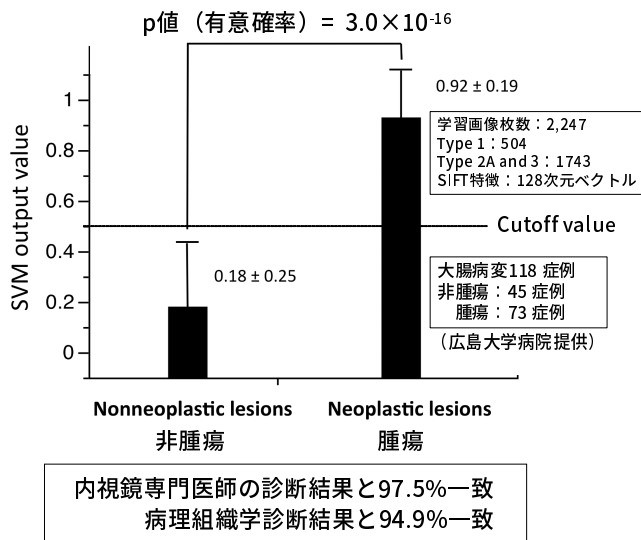


図 3 SVM を用いた大腸腫瘍・非腫瘍の識別有用性評価 [11].

平面を境界として正と負のどちら側に位置するのかを識別関数を計算することにより判断する。また、本システムにおいて識別は Type 1 か $\bar{1}$ (即ち Type 2A か 3) を判断するものとする。本稿において用いる SVM の識別関数は、式 (1) のように示すことができる。式 (1) は入力データ次元数を $dims.$ とした時に Type 1 か $\bar{1}$ かを判断する際の例である。システム実装においては、線形カーネルの識別に対応した LIBLINEAR ライブラリを使用した [12].

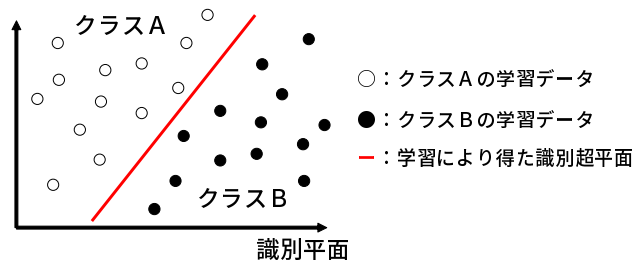


図 4 SVM による 2 クラス分類.

$$f_{1;\bar{1}}(\vec{x}) = \sum_{i=1}^{dims.} coef_i \times x_i \quad (1)$$

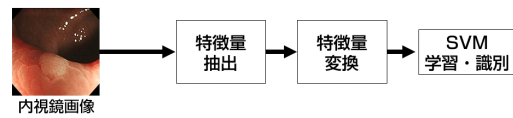
3. CNN 特徴を適用した SVM 分類手法

本節では、これまで研究グループで提案してきた BoF に基づいた分類手法において、内視鏡画像を Convolutional Neural Network (CNN) に入力した時の処理結果を特徴量として利用し、SVM への入力データとして診断支援を行う分類手法について述べる。

図 5 に BoF に基づく大腸がん診断支援システムと、CNN 特徴を適用した SVM 分類手法の概念図を示す。

図 7 に本稿で用いる CNN である、AlexNet のネットワーク構成概要図を示す [13]. 本稿で特徴抽出器として用いる AlexNet は、一般物体認識コンテストの Large Scale Visual

- BoF[†]に基づく大腸がん診断支援システム (これまで研究グループで提案) [†]BoF: Bag-of-Features



- CNN特徴を特徴抽出として利用した大腸がん診断支援システム

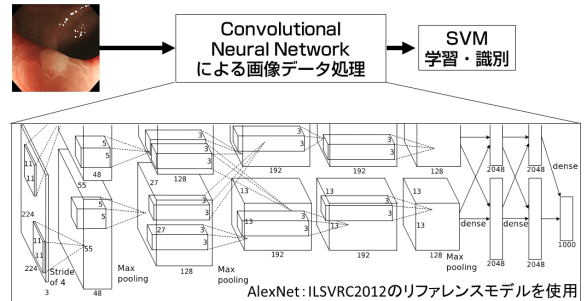


図 5 大腸がん診断支援システム 2 手法の概念図.

(AlexNet のネットワーク図は文献 [13] から引用)

Recognition Challenge 2012(ILSVRC2012) で提供された ImageNet のデータセット 1000 カテゴリを出力とするモデルを用いる [13]. AlexNet は画像サイズ 227×227 pixel を入力とし、畳込み演算による局所特徴量や Pooling による局所領域の値を用いた特徴マップの低解像度化を繰り返し実行し、最終結果として一般物体認識対象の 1000 ラベルそれぞれに属する確率を出力する。

例として、Type 2A の画像を AlexNet へ入力した時に出力される上位 5 つのカテゴリとその確率値を図 6 に示す。リファレンスモデルの AlexNet では内視鏡画像を識別する学習は行っていないため、jellyfish といった学習したカテゴリの中で近いと判断されるものを確率として高く出力している。



図 6 Type 2A 画像を AlexNet へ入力した時の上位 5 ラベル確率.

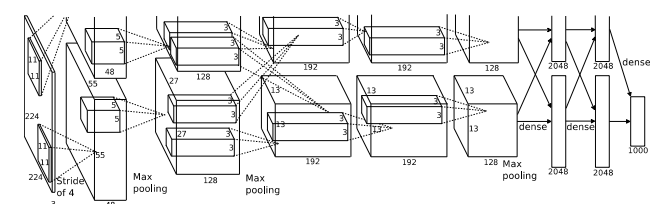


図 7 AlexNet のアーキテクチャ図 [13].

4. 中間層を特徴量として用いた場合の診断支援システムの識別精度シミュレーション

本節では、第3節で特徴量抽出処理に CNN を用いた SVM 分類手法について、その識別精度をシミュレーションにより検証する。

シミュレーション環境としては、Ubuntu 16.04 上で構築したオープンソース Deep Learning Library である、Caffe framework[14] を用いた。用いる CNN としては AlexNet, SVM ライブラリとして LIBLINEAR を用いた。LIBLINEAR は線形カーネルのみに対応し処理高速化を図ったもので、式 (1) に示すようにサポートベクトルの生成を不要とし、各次元の重み係数 $coef_i$ を学習により決定する。

4.1 学習およびテスト用画像データセット

本稿では、広島大学病院にて専門医師により大腸内視鏡診断時に撮影収集された、NBI 拡大内視鏡画像から病理タイプの特徴が明確に確認できる領域を矩形にトリミングした 1260 枚の画像を用いる。この 1260 枚の画像を学習し、右へ 90 度回転させたものを未学習のテスト用の画像として用いる。本稿では分類としては Type 1 かそうでないか (Type 2A と 3) を分類するための学習を行う。

Type 1, 2A, 3 それぞれの枚数内訳と画像サイズ分布、および各タイプの画像の例を図 8 に示す。本稿で用いる AlexNet の入力画像サイズは 227×227 pixel であるため、CNN へ入力する前にリサイズ処理を行う。図 8 に示すようにデータセットの画像サイズとしては $100 \sim 800$ pixel と様々な大きさのものがあ、リサイズ処理による血管構造などの欠落の可能性が考えられるため、注意して今後検証を続ける必要がある。

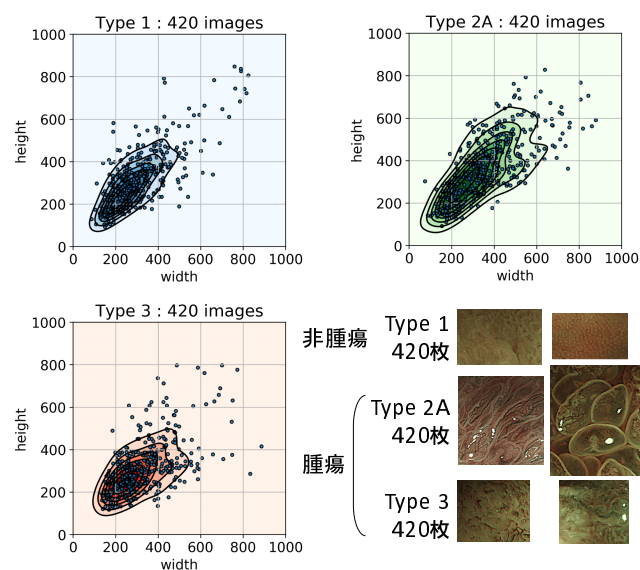


図 8 大腸 NBI 拡大内視鏡画像データセットおよび各タイプの画像サイズ分布 (広島大学病院提供)。

4.2 採用する CNN 中間層の違いによる

リアルタイム処理実現に向けて考慮すべき事柄

提案手法の組み込みシステム化への実現に際し、用いる中間層データの次元数は次段で識別処理を行う SVM のデータ次元数となるため、識別精度と特徴量次元数との間のトレードオフを考慮する必要がある。図 9 に AlexNet において処理の途中結果を特徴量として用いた場合の次元数を示す。最終段である prob は AlexNet の出力値である 1000 カテゴリの確率である。本稿では中間層データとして pool5 (9,216 次元), fc6 (4,000 次元), fc7 (4,000 次元), fc8 (1,000 次元), prob (1,000 次元) を特徴量抽出の候補とし、識別精度に関する検証を行う。

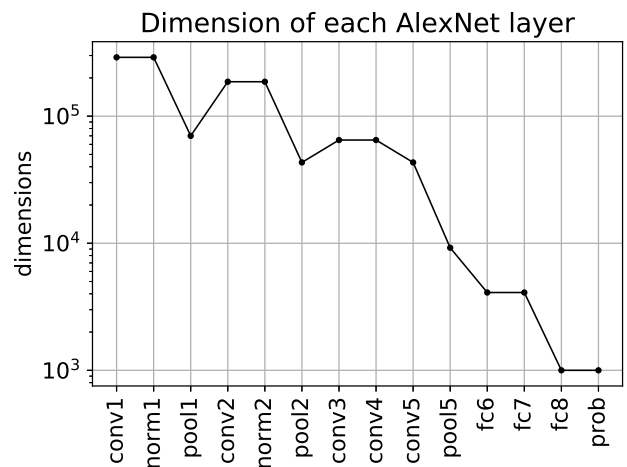


図 9 AlexNet 各中間層での特徴量次元数。

4.3 交差検証 (Cross Validation) による正答率評価

識別精度の検証として、まず交差検証 (Cross Validation: CV) による精度検証を行った。CV ではデータセットの画像 1260 枚を 10 分割し、9 割のものを使って学習、1 割のものを使ってテストを行う。CV による精度検証では、評価指標として正答率 (accuracy) を用いた。accuracy は全テスト枚数に対する正しく腫瘍または非腫瘍と判断された枚数の割合によって算出される。図 10 にその結果を示す。横軸は pool5 から prob までの用いた特徴量を、縦軸は accuracy および特徴量次元数を示している。この結果より、CNN 最終出力である prob を特徴量とした場合よりも、それ以前の中間層データを特徴量として用いた方が正答率が向上していることがわかった。

4.4 True Positive, Precision Rate による

各タイプの識別性能評価

次に、識別の正確さを示す指標である True Positive, 識別の偏りを示す指標である Precision Rate を用いて比較を行う。これにより、正答率では評価できなかった各タイプに関する識別性能を別々に確認することができる。True

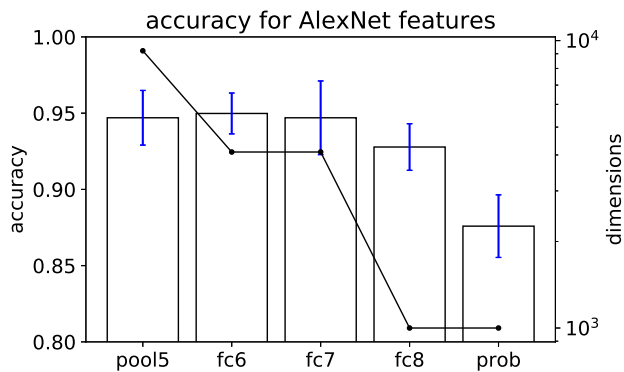


図 10 AlexNet 各中間層での特徴量次元数.

Positive は式 (2), Precision Rate は式 (3) のように示され, 分類する対象毎に評価する. 本稿において, Type 1 と $\bar{1}$ の分類を行うため, 図 11 に示すように, テストとして入力した画像の正しい Type と, SVM が識別した Type とを評価し True Positive と Precision Rate を計算する. また, この識別性能評価に関して, 特徴量として用いる CNN を 2 種類設定した. 1 つは学習済みの AlexNet モデルを特徴量抽出として用いたもの, もう一方は組み込みシステムへ向けて研究グループで提案したハードウェア向けに処理最適化された AlexNet である [15]. 前者を Software, 後者を Hardware として性能評価結果を比較した. なお, 前節図 10 で示した結果は Software 特徴量による評価結果である.

$$True\ Positive(x|1, \bar{1}) = \frac{N_{x:x}}{\sum_{y=1, \bar{1}} N_{x:y}} \quad (2)$$

$$Precision\ Rate(x|1, \bar{1}) = \frac{N_{x:x}}{\sum_{y=1, \bar{1}} N_{y:x}} \quad (3)$$

Type 1 vs $\bar{1}$ ($T = 2A, 3$)		識別された Type		True Positive
		1 (非腫瘍)	$\bar{1}$ (腫瘍)	
入力画像 Type	1 (非腫瘍)	$N_{1:1}$	$N_{1:\bar{1}}$	TP(1)
	$\bar{1}$ (腫瘍)	$N_{\bar{1}:1}$	$N_{\bar{1}:\bar{1}}$	TP($\bar{1}$)
Precision Rate		PR(1)	PR($\bar{1}$)	

図 11 Type 1 と $\bar{1}$ 分類における True Positive (TP), Precision Rate (PR).

4.5 シミュレーション結果

シミュレーション結果として, Type 1 vs $\bar{1}$ 分類器に関して, 図 12 および図 13 に Type 1 に関する True Positive, Precision Rate を示す. 図 14 および図 15 に Type $\bar{1}$ に関する True Positive, Precision Rate を示す.

全体的な傾向としては, CV による accuracy 評価結果と同様の傾向が確認できた. 特に fc6, fc7, fc8 特徴量を用いた場合の評価結果としては True Positive, Precision Rate 共に 95% を超える結果が確認できた. また, Software-Hardware 間の比較としては, 特徴量次元数が 9,216 である pool5 に

て Hardware 特徴量の方が低い結果となった.

5. 結論

本稿では, 大腸 NBI 拡大内視鏡診断支援を実現に向け CNN 特徴と SVM 分類を適用した大腸 NBI 拡大内視鏡画像がん診断支援手法に関する識別性能評価を行った. AlexNet における pooling5, fc6, fc7, fc8, probability 各層のデータを特徴量として用いた場合, fc6, fc7, fc8 特徴量にて識別精度 95% 以上を確認することができた. 今後の課題としては, 組み込みシステムとして診断支援システムのリアルタイム処理を実現するにあたり, SVM 識別部の処理時間および識別精度と用いる特徴量次元数とのトレードオフについてさらに検証を続けていく必要がある.

謝辞

本研究の一部は, JSPS 特別研究員奨励費 (16J06130), JSPS 科研費基盤研究 (B) 26280015, JSPS 科研費基盤研究 (B) 17H01714, 並びに生体医歯工学共同研究拠点の助成を受けたものです.

本研究では, 日本ケイデンス・デザイン・システムズ社の組み込みシステム向けカスタマイズブル DSP プロセッサコア Vision P6 DSP コアと統合開発環境 Xtensa Xplorer を使用しました. また, デモシステムの共同開発に御尽力頂いた菅原崇之様, 戸石浩司様, 辻雅之様, 小田川真之様, 丹場展雄様に感謝致します.

参考文献

- [1] H. Machida, Y. Sano, Y. Hamamoto, M. Muto, T. Kozu, H. Tajiri, and S. Yoshida, "Narrow-Band Imaging in the Diagnosis of Colorectal Mucosal Lesions: a pilot study.," *Endoscopy*, vol.36(12): 1094-1098, 2004.
- [2] H. Ikematsu, T. Matsuda, F. Emura, Y. Saito, T. Uraoka, K.I. Fu, K. Kaneko, A. Ochiai, T. Fujimori, and Y. Sano, "Efficacy of Capillary Pattern type IIIA/IIIB by Magnifying Narrow Band Imaging for Estimating Depth of Invasion of Early Colorectal Neoplasms," *BioMed Central (BMC) Gastroenterology*, 2010.
- [3] H. Kanao, S. Tanaka, S. Oka, M. Hirata, S. Yoshida, and K. Chayama, "Narrow-Band Imaging Magnification Predicts The Histology and Invasion Depth of Colorectal Tumors," *Gastrointestinal Endoscopy*, vol.69, no.3, pp.631-636, mar 2009.
- [4] Y. Wada, S. ei Kudo, H. Kashida, N. Ikehara, H. Inoue, F. Yamamura, K. Ohtsuka, and S. Hamatani, "Diagnosis of Colorectal lesions with the Magnifying Narrow-Band Imaging System," *Gastrointestinal Endoscopy*, vol.70, 2009.
- [5] T. Nikami, S. Saito, H. Tajiri, and M. Ikegami, "The Evaluation of Histological Atypia and Depth of Invasion of Colorectal Lesions using Magnified Endoscopy with Narrow-Band Imaging," *Gastrointestinal Endoscopy*, vol.51, 2009.
- [6] Y. Sano *et al.*, "Narrow-Band Imaging (NBI) Magni-

ifying Endoscopic Classification of Colorectal Tumors Proposed by the Japan NBI Expert Team,” Digestive Endoscopy, vol.28, 2016.

- [7] T. Tamaki, J. Yoshimuta, M. Kawakami, B. Raytchev, K. Kaneda, S. Yoshida, Y. Takemura, K. Onji, R. Miyaki, and S. Tanaka, “Computer-Aided Colorectal Tumor Classification in NBI Endoscopy using Local Features,” Medical Image Analysis, vol.17, no.1, pp.78–100, jan 2013.
- [8] T. Okamoto, T. Koide, A.T. Hoang, T. Shimizu, K. Sugi, T. Tamaki, T. Hirakawa, B. Raytchev, K. Kaneda, S. Yoshida, H. Mieno, and S. Tanaka, “An FPGA Implementation of SVM for Type Identification with Colorectal Endoscopic Images,” Proceedings of the 20th Workshop on Synthesis And System Integration of Mixed Information technologies (SASIMI 2016), pp.81–86, 2016.
- [9] T. Koide, T. Okamoto, T. Shimizu, K. Sugi, A.T. Hoang, T. Tamaki, B. Raytchev, K. Kaneda, S.Y. and Hiroshi Mieno, and S. Tanaka, “Compact and High-Speed Hardware Feature Extraction Accelerator for Dense Scale-Invariant Feature Transform,” Proceedings of the 31st International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC2016), pp.387–390, 2016.
- [10] T. Koide, T. Okamoto, K. Sugi, T. Shimizu, A.T. Hoang, T. Tamaki, B. Raytchev, K. Kaneda, S.Y. and Hiroshi Mieno, and S. Tanaka, “A Hardware Accelerator for Bag-of-Features based Visual Word Transformation in Computer Aided Diagnosis for Colorectal Endoscopic Images,” Proceedings of the 31st International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC2016), pp.233–236, 2016.
- [11] Y. Kominami, S. Yoshida, S. Tanaka, Y. Sanomura, T. Hirakawa, B. Raytchev, T. Tamaki, T. Koide, K. Kaneda, and K. Chayama, “Computer-Aided Diagnosis of Colorectal Polyp Histology by using a Real-Time Image Recognition System and Narrow-Band Imaging Magnifying Colonoscopy,” Gastrointestinal Endoscopy, 2016.
- [12] R.E. Fan, K.W. Chang, C.J. Hsieh, X.R. Wang, and C.J. Lin, “LIBLINEAR: A Library for Large Linear Classification,” Journal of Machine Learning Research, vol.9, pp.1871–1874, 2008.
- [13] A. Krizhevsky, I. Sutskever, and G.E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” in Advances in Neural Information Processing Systems 25, ed. F. Pereira, C.J.C. Burges, L. Bottou, and K.Q. Weinberger, pp.1097–1105, Curran Associates, Inc., 2012.
- [14] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, “Caffe: Convolutional Architecture for Fast Feature Embedding,” arXiv preprint arXiv:1408.5093, 2014.
- [15] 岡本 拓巳, 小出 哲士, 玉木 徹, Bisser Raytchev, 金田和文, 吉田 成人, 三重野 寛, 田中 信治, 戸石 浩司, 菅原 崇之, 辻 雅之, 小田川 真之, and 丹場 展雄, “CNN 特徴と SVM 分類を適用した大腸内視鏡画像がん診断支援システムのカスタマイズブル DSP コアへの実装,” DA シンポジウム 2017, pp.33–38, 2017.

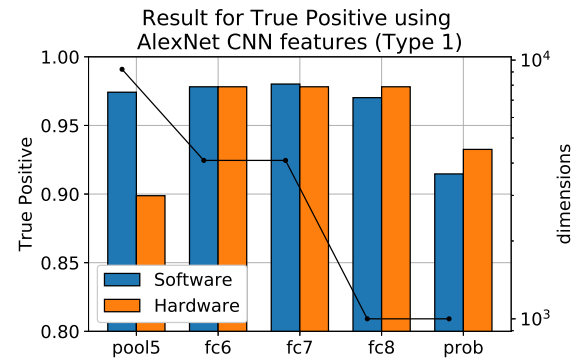


図 12 Type 1 vs. not1 分類における, Type 1 に関する True Positive.

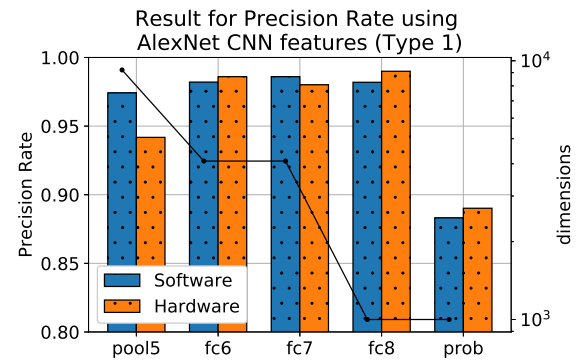


図 13 Type 1 vs. not1 分類における, Type 1 に関する Precision Rate.

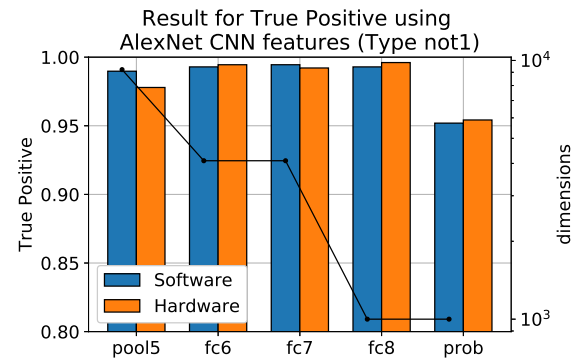


図 14 Type 1 vs. not1 分類における, Type not1 に関する True Positive.

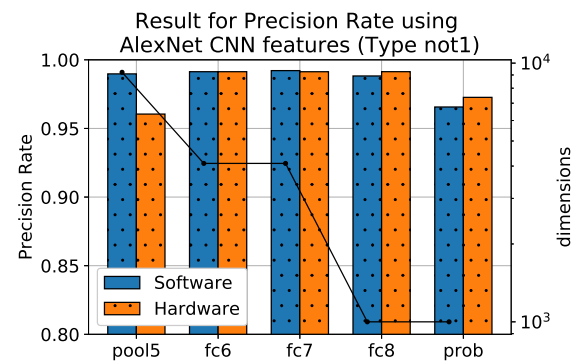


図 15 Type 1 vs. not1 分類における, Type not1 に関する Precision Rate.