

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可能なDSPコアへの実装

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 戸石 浩司⁵ 菅原 崇之⁵ 辻 雅之⁵ 小田川 真之⁵ 丹場 展雄⁵

概要: 本稿は、Convolutional Neural Network (CNN) の処理結果を特徴量として使用し、Support Vector Machine (SVM) による病理タイプ分類を行う、大腸 Narrow Band Imaging (NBI) 拡大内視鏡がん診断支援システムのリアルタイム処理実現を目的とする。CNN 特徴と SVM 分類を適用した診断支援システムをカスタマイズ可能な Digital Signal Processing (DSP) コアである、Cadence 社の Cadence Tensilica Vision P6 DSP コアに実装し評価を行った。その結果、コアアーキテクチャに適したアルゴリズム改良を行うことで CNN や SVM の主要処理が効率的に実行でき、改良前と比較して処理サイクル数が 1/30 に削減されリアルタイム処理 (30 fps@100 MHz) が実現可能であることを示した。

Implementation of Computer-Aided Diagnosis System for Colorectal Endoscopic Images with CNN features and SVM on Customizable DSP Core

TAKUMI OKAMOTO^{1,a)} TETSUSHI KOIDE¹ TORU TAMAKI² BISSER RAYTCHEV² KAZUFUMI KANEDA²
SHIGETO YOSHIDA³ HIROSHI MIENO³ SHINJI TANAKA⁴ HIROSHI TOISHI⁵ TAKAYUKI SUGAWARA⁵
MASAYUKI TSUJI⁵ MASAYUKI ODAGAWA⁵ NOBUO TAMBA⁵

Abstract: This paper introduce an implementation of Computer-Aided Diagnosis (CAD) system for colorectal endoscopic images on a customizable Digital Signal Processing (DSP) core, Cadence Tensilica Vision P6 DSP. We estimated the system implemented on Vision P6 DSP core that is using Convolutional Neural Network (CNN) as a feature extraction module for Support Vector Machine. As an estimation for processing cycles of SVM module, we achieved that the SVM module optimized for effective processing onto Vision P6 DSP is about 1/30 faster than before optimized algorithm.

1. はじめに

大腸がん罹患患者数は年々増加傾向にある一方で、早期ス

テージの状態であればほぼ 100%完治可能であることから、早期発見、治療を行うことが肝要となる。早期発見の手段の一つとして、大腸内視鏡検査による大腸内壁の観察、病理有無の確認がある。大腸内視鏡検査の方法としては、NBI(Narrow Band Imaging) システムを用いた大腸拡大内視鏡を用いた画像検査が行われている。大腸 NBI 拡大内視鏡診断では、専門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がん深達度を診断する。この診断を行う医師には高い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といった課題がある。

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

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 広島大学大学院 医歯薬保健学研究所 内視鏡医学
Department of Endoscopy and Medicine Graduate School of Biomedical and Health Science, Hiroshima University
⁵ 日本ケイデンス・デザイン・システムズ社
Cadence Design Systems, Japan
a) koide-lab-info@ml.hiroshima-u.ac.jp

	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].

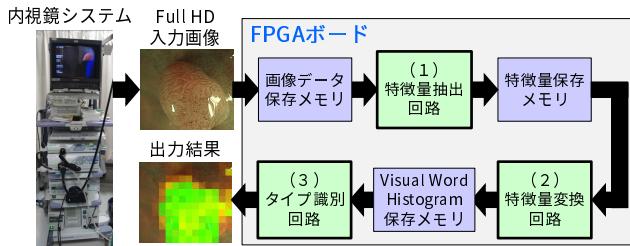


図 2 BoF に基づくリアルタイム診断支援ハードウェアシステム.

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, Japan NBI Expert Team (JNET) 分類 [6] に基づき, 大腸 NBI 拡大内視鏡画像を図 1 に示す 4 タイプのうち, 3 つの病理タイプ (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。

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。これまでの研究成果として, Bag-of-Features (BoF) に基づいたリアルタイム大腸 NBI 拡大内視鏡画像分類システムを提案し, 図 2 に示す 3 つ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 [7], [8], [9].

本稿では, BoF フレームワークにおける特徴量抽出処理として Convolutional Neural Network (CNN) の計算結果を適用し, Support Vector Machine (SVM) 分類を行う手法に関して, カスタマイズ可能な DSP コアへの実装に向けた最適化および統合開発環境上での処理サイクル数の比較を行い, リアルタイム処理が実現可能であることを示す。

2. CNN 特徴を適用した SVM 分類手法

本節では, これまでに研究グループで提案してきた Bag-of-Features に基づく大腸病理タイプ分類手法とは異なるア

プローチとして CNN 特徴を適用した SVM 分類手法について述べる。図 3 に BoF フレームワークの外観図を示し, 図 4 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。

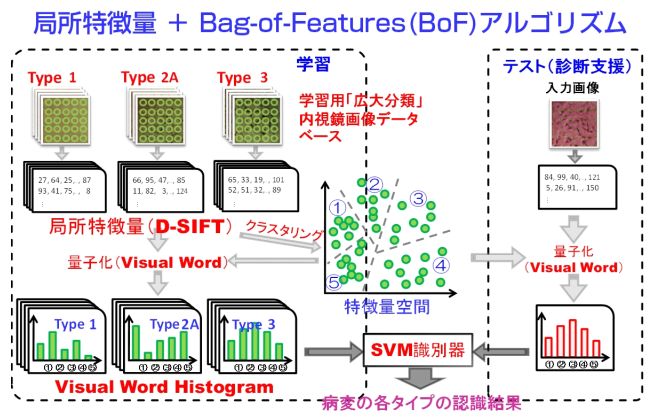
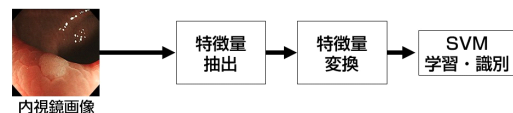


図 3 BoF に基づく大腸 NBI 拡大内視鏡画像がん診断支援システム.

- 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

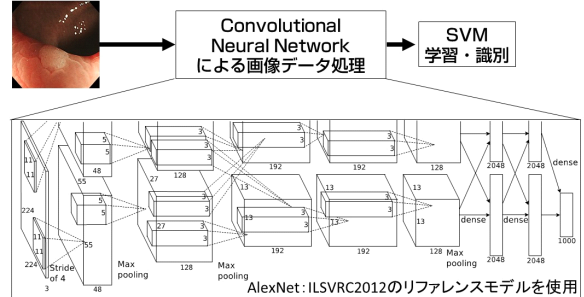


図 4 大腸がん診断支援システム 2 手法の概念図.

(AlexNet のネットワーク図は文献 [10] から引用)

2.1 Convolutional Neural Network (CNN) を用いた特徴量抽出

Convolutional Neural Network は LeCun らによって提案され, 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の繰り返しにより実現され, 一般物体認識の分野で高い認識性能を示していることで知られている [11]. 本稿では, CNN の出力を入力画像の特徴量を表現していると解釈し, 次節に述べる Support Vector Machine における入力データとしてこれを用いる。本稿にて利用する CNN は AlexNet を用いる。具体的には, 一般物体認識コンテストの Large Scale Visual 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 Cadence 社により組み込みシステム向けに処理を最適化したものを用いる [12].

2.2 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 5 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 の識別関数を、式 (1) に示す。式 (1) は入力データである 1000 次元の CNN の出力値が 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 [13]。

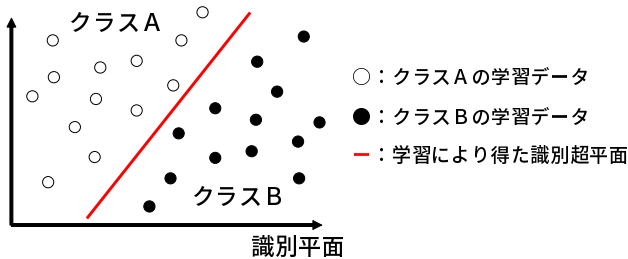


図 5 SVM による 2 クラス分類。

$$f_{1;\bar{1}}(\vec{x}) = \sum_{i=1}^{dims.} coef_i \times x_i \quad (dims. = 1000) \quad (1)$$

ここで、 \vec{x} は入力データであり、CNN の出力である 1000 次元のベクトル値である。coef は学習時に決定された各次元に対する重みパラメータである。これらの積和演算を行うことで入力データがどちらのタイプに属するかを判断する。

3. カスタマイズ可能な DSP コアへの実装に向けたアルゴリズム最適化

本節では第 2 節で述べた CNN 特徴を SVM の入力データとする分類手法に関して、Cadence Design Systems 社が画像処理、CNN 処理を効率よく実行する DSP コアとして提案しているカスタマイズ可能な DSP コアである、Cadence Tensilica[®]*1 Vision P6 DSP コア [14](以降、P6 コア) への実装に向けたアルゴリズム最適化を行う。図 6 に P6 コアのアーキテクチャを示す。P6 コアは以下に示すような機能を有している。

- 5 スロット VLIW/SIMD DSP
 - 16 × 32, 32 × 16, 64 × 8 ビット
 - Xtensa LX7 32 ビットプロセッサ
- 128 ビット命令 RAM/キャッシュ

*1 Tensilica は Cadence Design Systems 社の登録商標です。

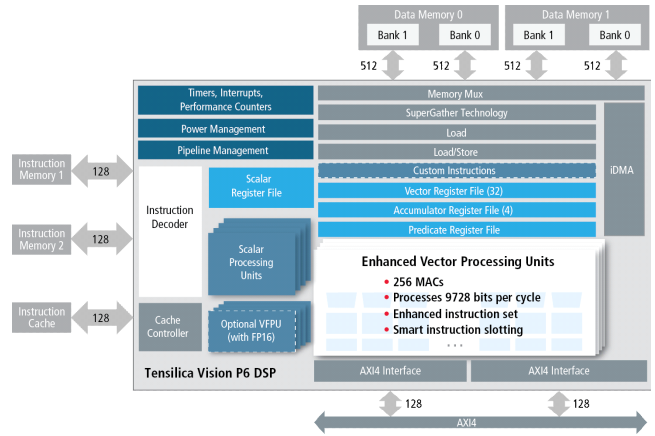


図 6 Cadence Tensilica Vision P6 DSP コア アーキテクチャ。

- 512 ビットデュアルロードストア
- SuperGather
- iDMA
- 256 MAC (8 bit × 8 bit)
- VFPU (16 × 単精度, 32 × 半精度)
- 2x AXI4 マスターインターフェース
- 構成変更及び拡張可能

3.1 SVM 部の並列命令化および float 型演算への変更

アルゴリズムの P6 コアへのポーティングは、P6 コアの統合開発環境である、Xtensa[®]*2 Xplorer [15] の機能の一つである Auto Vectorization 機能を用いる。Auto Vectorization 機能では、P6 コアが持つ並列命令を有効可能であるコードを自動的に並列命令化し、並列化が行えなかったコードに関しても並列化の試行結果をレポートとして一覧表示することができる。また Xtensa Xplorer では処理サイクル数の今回行ったアルゴリズム最適化の手順は以下の通りである。

- (1) Auto Vectorization 機能を使ったコード修正なしでの自動並列命令化
- (2) (1) の結果を踏まえたコード修正
- (3) Xtensa Xplorer の Profile 機能によるサイクル数、パイプライン処理検証
- (4) 更なる並列化の検討

SVM 部の積和演算部に関して、P6 コアへの実装にあわせて並列命令化を実現するため、P6 コアが有する 16 並列実行が可能な単精度浮動小数点 (float) 型への数値計算型の変更を行った。図 7 に修正前のコードの一部を示す。このコードでは Auto Vectorization による並列命令化は成功しておらず、原因としては数値型の未明記、ポインタ型による値読出し時のアライメント不明などが挙げられた。コードを修正し、並列命令化が成功したコード一部を図 15 に示す。このコードでは 1 度の for ループで 8 回の乗算が記

*2 Xtensa は Cadence Design Systems 社の登録商標です。


```
for (; (idx=lx->index) != -1; lx++)
{
    if (idx <= n)
        for (i=0; i < nr_w; i++)
        {
            dec_values[i] += w[idx*nr_w+i]*lx->value;
        }
}
```

図 7 修正前のコード抜粋.

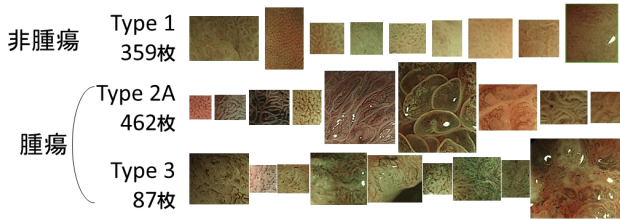


図 8 大腸 NBI 拡大内視鏡画像データセット (広島大学病院提供).

述されているが、並列化によりさらに倍の float 型での 16 並列乗算が有効となった。

また、float 型への修正に際して識別精度の検証を行い、識別精度が同等以上であることを確認し、その結果数値型を変更した計算結果を用いて学習を行うことにより、数値型の変更をカバーすることが可能であることを確認している [16].

3.2 システムにおける SVM 分類器の学習データ

本稿では、システムの SVM の学習には広島大学病院にて専門医師による大腸内視鏡診断時に撮影された、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 908 枚の大腸 NBI 拡大内視鏡画像を用いる。この 908 枚の画像を 227 × 227 pixel にリサイズしたものを CNN へ入力し、画像ごとに出力される 908 個の 1000 次元のベクトルデータを SVM の学習データとして用いる。システムでは非腫瘍・腫瘍の識別を行う、Type 1 か $\bar{1}$ ($\bar{1}$ = Type 2A and 3) かを識別する学習を行った。Type 1, 2A, 3 それぞれの枚数内訳を図 8 に示す。

3.3 システムの動作概観

提案する大腸 NBI 拡大内視鏡画像がん診断支援システムの動作概観図を図 11 に示す。システムは Host コンピュータと、P6 コアが実装されたケイデンス・ラピッド・プロトタイプング・プラットフォーム Protium[®]*3 S1[17] の 2 部から構成される。まず Host 側から画像データが Vision P6 DSP コアに転送され、CNN および SVM の処理が実行された後、SVM の識別結果を Host 側に転送し、それを表示する。

Host-DSP 間のデータ通信は図 11 および図 12 に示すように iPass コネクタによる PCI Express で行われ、プラットフォーム内のデータ通信は AXI バスにより実現し、

*3 Protium は Cadence Design Systems 社の登録商標です。

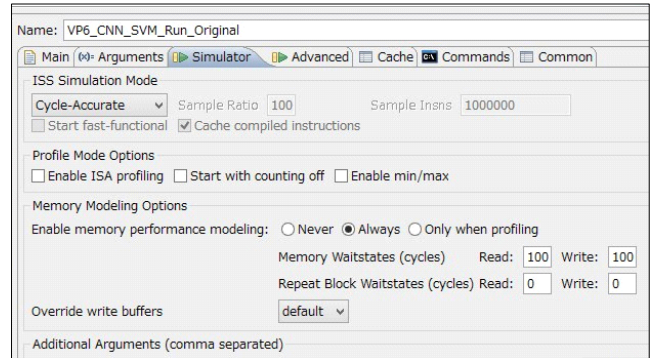


図 9 メモリアクセス最適化に向けたメモリ構成やアクセスサイクル設定画面.

CNN 部の学習データである重み係数データ、SVM 部の学習データであるモデルファイルデータはこれを介して DataRAM や必要に応じて外部 Work RAM に転送される。

システムの動作フローチャートを図 13 に示す。システムが動作開始すると入力画像または動画データからフレームデータとして画像を読み込む。次に CNN への入力データである 227 × 227 pixel の領域を識別領域として切り取り、識別領域を Host-DSP 間で共有しているメモリへ書き込む。その後、DSP 内で CNN および CNN の出力結果を入力とした SVM の識別処理が行われ、識別領域に対する処理結果を得る。その後処理結果を共有メモリへ書き込み、Host 側が読み込み、最後に処理結果を入力画像と共に表示する。以上の処理を終了キーが押下されるまで繰り返し実行する。

システム実装に際して、Host-DSP コア間のデータ通信を含めたハードウェア・ソフトウェア協調設計には Xtensa Xplorer 上での動作シミュレーションによる動作検証を行った。図 9 にシステムのメモリアクセス最適化に向けたシミュレーションを行う際のメモリ構成やアクセスサイクル数の設定画面を示す。図 10 に P6 コア側のパイプライン処理解析画面を示す。

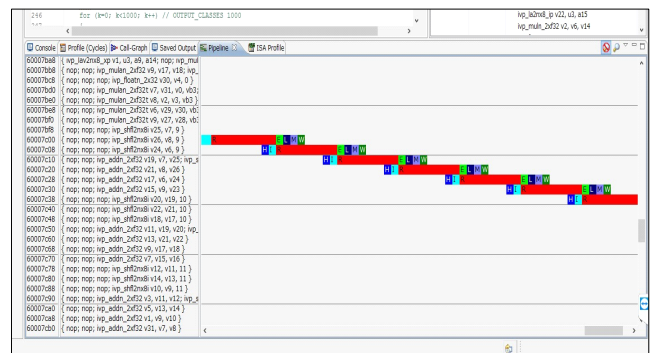


図 10 P6 コアのパイプライン処理解析画面.

3.4 サイクル数評価

本節では、提案システムのサイクル数評価を行う。サイ

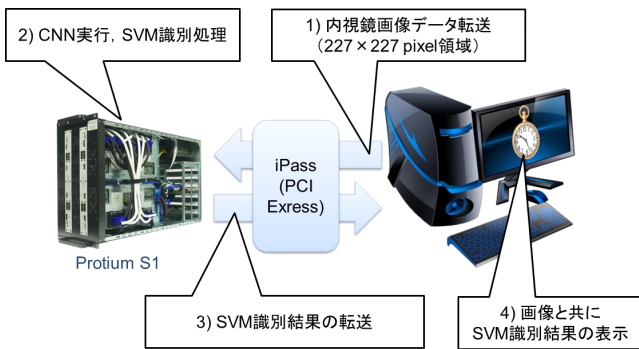


図 11 システムの動作概観.

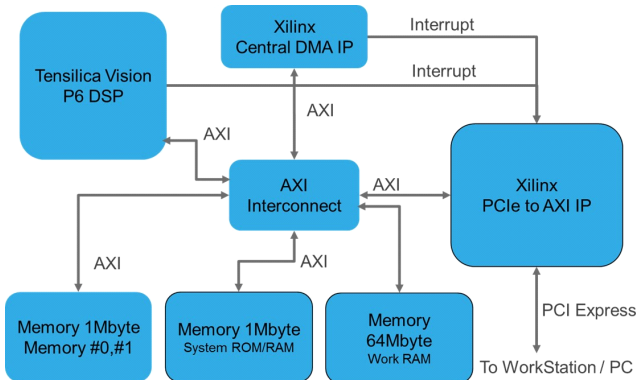


図 12 システムブロックダイアグラム.

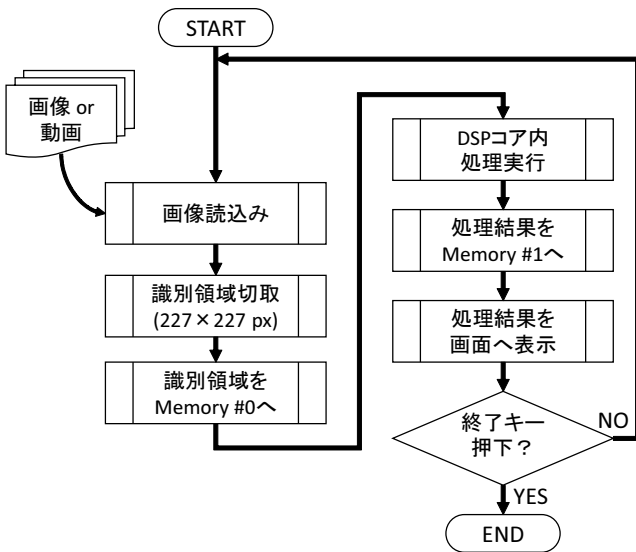


図 13 システム動作フローチャート.

クル数評価には、Vision P6 DSP の統合開発環境である、Xtensa Xplorer の Profile 機能を用いる。Xtensa Xplorer では実行された関数ごとの処理サイクル数をシミュレートすることができ、ソフトウェア・ハードウェア協調設計によるシステム実装の際に課題となるインターフェース部での処理速度試算などが容易に行うことができる。処理サイクル数の比較結果として、図 14 にアルゴリズム最適化前後の、908 枚の内視鏡画像を処理した場合の総サイクル数を示す。また、図 14 最右に示す棒グラフはアルゴリズム最

適化を行い、かつ入力画像データを Type 1, 2A, 3 の 3 タイプへの識別を実現させるため、Type 1 vs ̄1 の識別を行う SVM を 1 つと、Type 2A vs 3 の識別を行う SVM を 1 つの合計 2 つの SVM 識別器を実装した場合のサイクル数を示している。サイクル数評価結果より、DSP コア部の動作クロックを 1 GHz とした場合、SVM 部のサイクル数において最適化前ではフレームレートは約 0.5 fps であるのに対して、最適化後では約 17.5 fps とリアルタイム処理が可能であることが確認できた。また、SVM を 2 つ実装した場合でも約 8.9 fps と、3 タイプ識別を実行するシステム構成であっても要求性能を十分満たすことが確認できた。

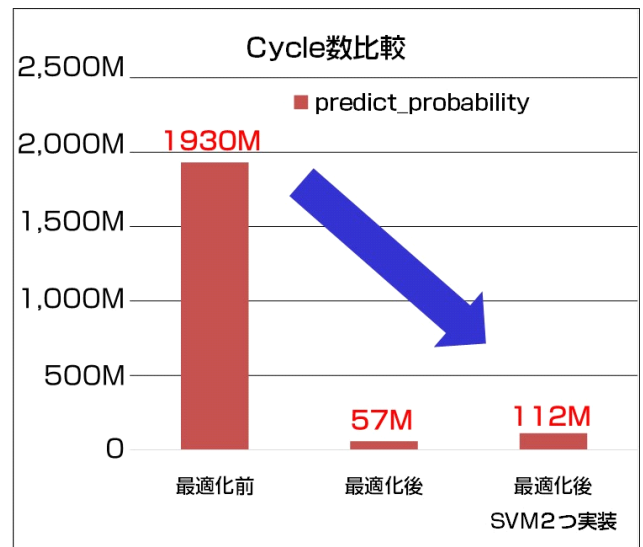


図 14 最適化前後での処理サイクル数比較.

4. 結論

本稿では、CNN の出力結果を特徴量として使用し SVM を用いて病理タイプの識別を行う、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システムのカスタマイズ可能な DSP コアへの実装について述べた。CNN や SVM の主要処理となる積和演算を効率的に実行するため、カスタマイズ可能な DSP コアである Vision P6 DSP コアアーキテクチャに適したアルゴリズムの改良を行い、1/30 の処理サイクル数削減を確認しリアルタイム処理が可能であることを示した。

今後の課題としては、CNN の出力値だけでなく中間層データの特徴量として用いることにより、CNN 部の処理サイクル数削減を図りつつ識別精度が出力値よりも高くなるような条件を模索し、システムのさらなる改良を行うことが挙げられる。

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

```

for(j=0; j<250; j++) // => 1000 / 4
{
    dec_value_tmp0 += (float)w[j + 0] * (float)PROB_SCALAR*(float)ResultbufVEC[j + 0];
    dec_value_tmp1 += (float)w[j + 250] * (float)PROB_SCALAR*(float)ResultbufVEC[j + 250];
    dec_value_tmp2 += (float)w[j + 500] * (float)PROB_SCALAR*(float)ResultbufVEC[j + 500];
    dec_value_tmp3 += (float)w[j + 750] * (float)PROB_SCALAR*(float)ResultbufVEC[j + 750];
}
dec_values[0] = dec_value_tmp0 + dec_value_tmp1 + dec_value_tmp2 + dec_value_tmp3;

```

図 15 修正後のコード抜粋.

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.

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) Magnifying Endoscopic Classification of Colorectal Tumors Proposed by the Japan NBI Expert Team," *Digestive Endoscopy*, vol.28, 2016.
- [7] 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.
- [8] T. Koide, T. Okamoto, T. Shimizu, K. Sugi, A.T. Hoang, T. Tamaki, B. Raytchev, K. Kaneda, S.Y. ands 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.
- [9] T. Koide, T. Okamoto, K. Sugi, T. Shimizu, A.T. Hoang, T. Tamaki, B. Raytchev, K. Kaneda, S.Y. ands 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.
- [10] 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.
- [11] Y. LeCun, B. Boser, J.S. Denker, D. Henderson, R.E. Howard, W. Hubbard, and L.D. Jackel, "Backpropagation Applied to Handwritten Zip Code Recognition," *Neural Computation*, vol.1, no.4, pp.541-551, Dec. 1989.
- [12] G. Efland, S. Parikh, H. Sanghavi, and A. Farooqui, "High Performance DSP for Vision, Imaging and Neural Networks," *IEEE Hot Chips 2016*, 2016.
- [13] 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.
- [14] "Vision DSPs for Imaging and Neural Networks." <https://ip.cadence.com/vision>.
- [15] "Software Development." <https://ip.cadence.com/swdev>.
- [16] 岡本拓巳, 小出哲士, 玉木徹, B. Raytchev, 金田和文, 吉田成人, 三重野寛, and 田中信治, "リアルタイム大腸内視鏡画像診断支援システムのための CNN 特徴と SVM 分類を用いた識別手法," *DA シンポジウム 2017*, 2017.
- [17] "Protium S1 FPGA-Based Prototyping Platform." <http://www.cadence.com/news/protium>.