Customizable Hardware Architecture of Support Vector Machine in CAD System for Colorectal Endoscopic Images with NBI Magnification

Satoshi Shigemi 1), Tsubasa Mishima 1), Anh-Tuan Hoang 1), Tetsushi Koide 1),
Toru Tamaki 2), Bisser Raytchev 2), Kazufumi Kaneda 2),
Yoko Kominami 3), Rie Miyaki 3), Taiji Matsuo 3), Shigeto Yoshida 3), Shinji Tanaka 3)

1) Research Institute for Nanodevice and Bio Systems, Hiroshima University
2) Graduate School of Engineering, Hiroshima University
3) Graduate School of Biomedical Sciences, Hiroshima University
1-4-2 Kagamiyama, Higashi-Hiroshima, Japan
E-mail: { satoshi-shigemi, mishima-tsubasa, anhtuan, koide } @ hiroshima-u.ac.jp

Abstract - With the increase of colorectal cancer patients in recent years, the needs of quantitative evaluation of colorectal cancer are increased, and the computer-aided diagnosis (CAD) system which supports doctor's diagnosis is essential. In this paper, a hardware design of type identification module in CAD system for colorectal endoscopic images with narrow band imaging (NBI) magnification [1] is proposed for real-time processing of full high definition (Full HD) image (1920 x 1080 pixel). In addition, in order to improve the identification accuracy for type B (TA: tubular adenoma) and type C3 (SM-m cancer), algorithms to realize a 3-class identification, which has high efficiency and high accuracy, is proposed.

I. Introduction

In CAD system for colorectal endoscopic images with NBI magnification [1], identification is performed by Support Vector Machine (SVM) which is a 2-class classifier. A 3-class identification problem can be solved by One-versus-One method [8,9]. In this paper, we propose a customizable multiclass identifier with SVM and verify the trade-off between hardware resources and identification performance.

II. CAD System for Colorectal Endoscopic Images with NBI Magnification

The proposed CAD system classifies colorectal endoscopic images obtained by endoscopic diagnosis into three types (Types A, B, C3) based on the NBI magnification findings [2]. In addition, type B can be divided into types BD, BN and BV. In that case, we have 5-class classification with 5 types of A, BD, BN, BV, and C3. Summary of this system is shown in Fig. 1. It adopts Bag-of-Features (BoF) as a fundamental concept. The overview of processing flow of the system is as follows. First, the feature quantities for each type are extracted from training images for each type by Dense Scale-Invariant Feature Transform (DSIFT) [3]. Then, those feature quantities are clustered and a Visual-Word (VW), which is effective in differentiation of each type image, is elected. In our system, 768-dimension VWs are used for each type. Next, for each type image, a histogram is created by voting all the feature quantities for similar VWs. Then, a learning step by Support Vector Machine (SVM) is performed using the histograms generated above. Finally a histogram is created from an input test image, and is identified by the learned SVM identifiers.

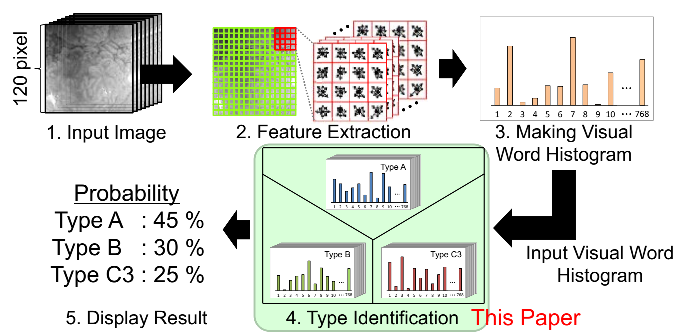


Fig. 1 : Summary of CAD system for colorectal endoscopic images with NBI magnification.

III. Type Identification Based on Support Vector Machine

A. Support Vector Machine (SVM)

Support Vector Machine (SVM) is a technique of binary classification introduced in 1990s by Vapnik [6,7]. Learning of SVM generates an identification hyper-plane that maximizes a margin between a class Y and a class Z. Therefore, the SVM achieves high identification performance to an unlearned image.

The kernel (k) of SVM can be computed by the Radial Basis Function (RBF), the linear, or the χ^2 methods shown in equation (1), (2) and (3), respectively.

$$k_{RBF}(\boldsymbol{x}_1, \boldsymbol{x}_2) = \exp(-\gamma|\boldsymbol{x}_1 - \boldsymbol{x}_2|^2), \tag{1}$$

$$k_{linear}(\boldsymbol{x}_1, \boldsymbol{x}_2) = \boldsymbol{x}_1^{\,T}\boldsymbol{x}_2, \tag{2}$$

$$k_{\chi^2}(\boldsymbol{x}_1, \boldsymbol{x}_2) = \exp(-\frac{\gamma}{2}\sum_i \frac{(x_{1i} - x_{2i})^2}{x_{1i} + x_{2i}}) \tag{3}$$

where γ is a scaling parameter which should be tuned for each problem. The accuracy difference in SVM computation using those different kernels is only 1% [1]. Hence, the simplest one, the linear kernel is selected in our research. The linear kernel is also best suitable for hardware implementation.

Equation (4) shows a decision function of SVM with linear kernel in this research.

$$d_{Y:Z}(\boldsymbol{x}) = \sum_{i=1}^{N_Y+N_Z} coef_i \times \boldsymbol{sv}_i \bullet \boldsymbol{x} + \rho_{Y:Z} \tag{4}$$

Here, x is a test data, $\boldsymbol{sv}_i$ is Support Vector (SV), which is obtained at the learning step, and constituting a hyperplane of identification respectively. $coef_i$ is a coefficient of each SV, and $\rho_{Y:Z}$ is a coefficient of a decision function. If identification function $d_{Y:Z}(x)$ is a positive value, an input data is determined as class Y.

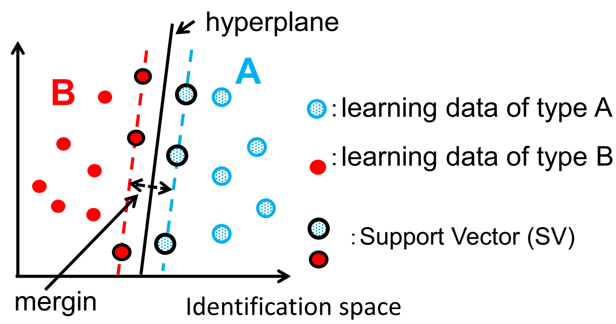


Fig. 2 : Support Vector Machine (SVM).

B. Multiclass Classification

In order to realize 3-class identification, the one-versus-one technique [8, 9] is taken. In our research, *f(x)* is a 2-class identifier with SVM.

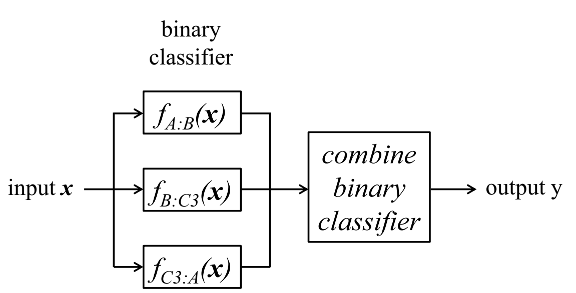


Fig. 3 : 3-class identification with one-versus-one technique.

C. Combine Binary Classifier

There are 2 methods in combine binary classifier. The first method uses majority vote of three binary identifiers and the second method uses probability estimation technique [4]. However, the latter one includes the loop processing, which is not suitable for hardware implementation, and consumes many resources.

In this paper, the final decision is made by majority vote among 2-class identifiers.

IV. Simulation Result for Real Images

A. *Software Implementation*

The Simple Gaussian Filter (GF) which is simplified for hardware implementation in Ref. [5] is used for feature extraction.

For the SVM classifier, we use LIBSVM [4]. In order to implement the SVM identifier by fixed-point number, the calculation of the equation (4), which contains floating-point number arithmetics, is changed into fixed-point number arithmetics based on LIBSVM. The Linear kernel (Inner product) is used as a kernel of SVM, and the number of Visual-Word is set to 768.

B. *Evaluation Method*

1260 real images, which are taken by Hiroshima University Hospital department of endoscopy, are used as the data set. These images are classified based on the NBI magnification findings by at least two professional doctors. Equations (5) and (6) are used for accuracy evaluation. True positive is used to determine whether identification of each type is carried out precisely and the Precision rate is used to determine whether the distinguished type has deflection.

Table 1 : NBI image dataset.

Type	A	B			C3	All
		BD	BN	BV		
Number of images	420	69	270	81	420	1260
		420				

$$True\,\text{Positive}\,(\text{i}) = \frac{\text{Posi\_Num(i)}}{\text{Img\_Num(i)}} \times 100\,[\%] \tag{5}$$

$$\text{Precision Rate}\,(\text{i}) = \frac{\text{Posi\_Num(i)}}{\text{Disc\_Num(i)}} \times 100\,[\%] \tag{6}$$

i : *Type A, Type B, Type C3, All.*

Img_Num(i) : *total number of image data of i.*

Disc_Num(i) : *number of images identified to i (A, B or C3).*

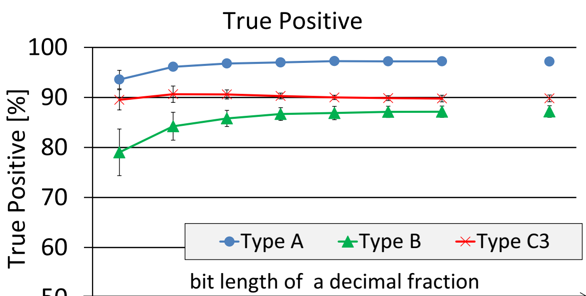
Posi_Num(i) : *number of images identified correctly.*

C. *Results for the 10-fold cross-validation*

Since high dimension feature quantity (768 dimensions) is processed in our medical application, it is necessary to perform the sum-of-product operations in parallel. Therefore, there is a question that how to reduce hardware resources while increasing parallel degree. We implement the SVM identification module by a fixed-point number instead of floating point number. The bit length reduces from 64 bits to around 16 bits, and, so as reduces hardware resources. Then, we suggest an algorithm of 3-class identifier which has efficient processing and high accuracy.

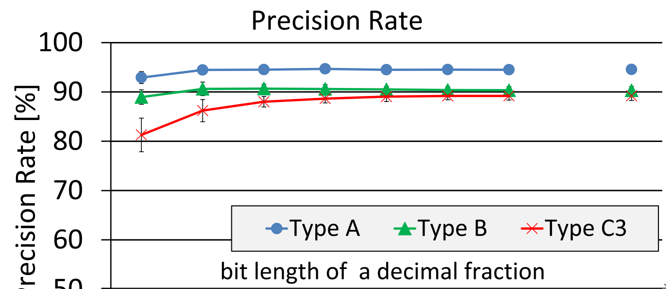
C.I. Evaluation of the bit width of fixed-point arithmetics in 3-class identification

The results of 3-class identification with 10 times 10-fold Cross Validation are shown in Figs. 4 and 5. The result shows that we need only 13 bits decimal fraction data in SVM implementation in hardware to achieve similar identification accuracy with the software implementation, which use 64-bit floating-point number (double precision).



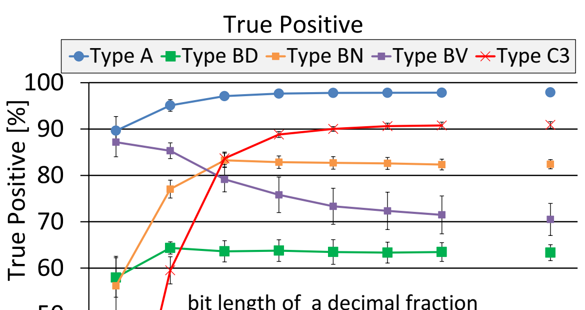
	12	13	14	15	16	17	18		double
Type A	93.6	96.1	96.8	97.0	97.3	97.2	97.2		97.2
Type B	79.0	84.2	85.8	86.7	86.9	87.1	87.1		87.2
Type C3	89.5	90.6	90.6	90.3	90.0	89.9	89.8		89.8

Fig. 4 : True Positive of types A, B, and C3 in 3-class classification.



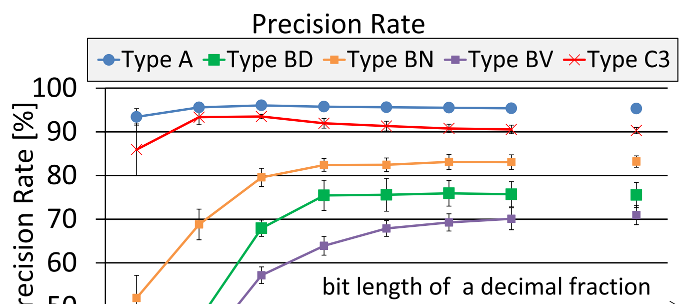
	12	13	14	15	16	17	18		double
Type A	92.9	94.5	94.5	94.7	94.5	94.5	94.5		94.6
Type B	89.0	90.6	90.7	90.6	90.5	90.4	90.3		90.3
Type C3	81.3	86.2	88.0	88.6	89.1	89.2	89.2		89.2

Fig. 5 : Precision Rate of types A, B, and C3 in 3-class classification.



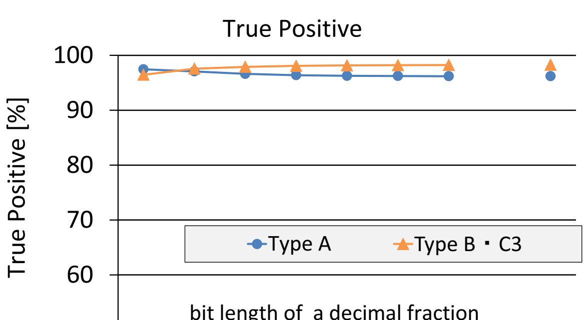
	12	13	14	15	16	17	18		double
Type A	89.6	95.1	97.1	97.6	97.8	97.8	97.8		97.9
Type BD	58.0	64.3	63.6	63.8	63.5	63.3	63.5		63.3
Type BN	56.2	77.0	83.3	82.9	82.7	82.6	82.3		82.4
Type BV	87.2	85.3	79.1	75.8	73.3	72.3	71.5		70.5
Type C3	10.3	59.5	83.7	88.8	90.0	90.6	90.8		90.9

Fig. 6: True Positive of types A, BD, BN, BV, and C3 in 5-class classification.



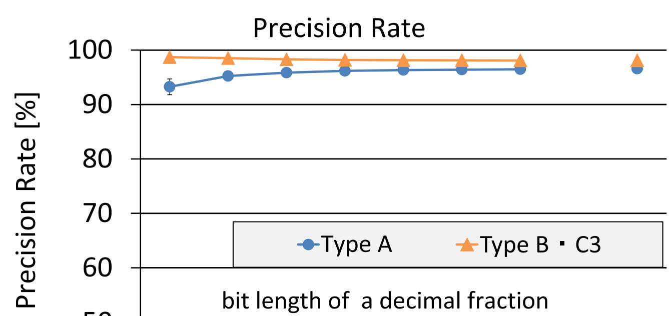
	12	13	14	15	16	17	18		double
Type A	93.4	95.6	96.1	95.8	95.6	95.5	95.4		95.3
Type BD	30.6	46.9	67.9	75.4	75.6	75.9	75.7		75.5
Type BN	51.9	68.8	79.6	82.4	82.5	83.1	83.1		83.2
Type BV	18.7	39.3	57.2	63.9	67.9	69.3	70.1		71.0
Type C3	85.9	93.4	93.5	91.9	91.3	90.8	90.6		90.3

Fig. 7: Precision Rate of types A, BD, BN, BV, and C3 in 5-class classification.



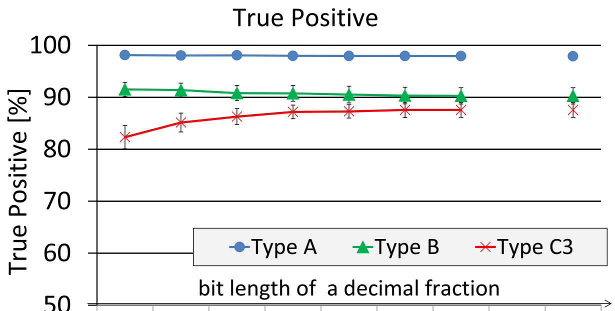
	12	13	14	15	16	17	18		double
Type A	97.5	97.1	96.6	96.4	96.3	96.2	96.2		96.2
Type B・C3	96.5	97.6	97.9	98.1	98.2	98.2	98.2		98.3

Fig. 8: True Positive (2-class, each type).



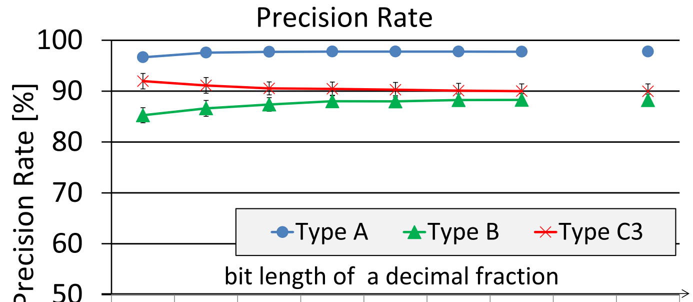
	12	13	14	15	16	17	18		double
Type A	93.3	95.2	95.8	96.2	96.3	96.4	96.5		96.6
Type B・C3	98.7	98.5	98.3	98.2	98.1	98.1	98.1		98.1

Fig. 9: Precision Rate (2-class, each type).



	12	13	14	15	16	17	18	double
Type A	98.1	98.0	98.1	98.0	98.0	98.0	97.9	97.9
Type B	91.5	91.4	90.8	90.8	90.5	90.3	90.3	90.3
Type C3	82.3	85.1	86.3	87.2	87.3	87.6	87.6	87.6

Fig. 10: True Positive (2-step, each type).



	12	13	14	15	16	17	18	double
Type A	96.7	97.6	97.7	97.8	97.8	97.8	97.7	97.8
Type B	85.3	86.6	87.4	88.0	88.0	88.2	88.3	88.2
Type C3	92.0	91.1	90.5	90.4	90.3	90.1	90.0	90.0

Fig. 11: Precision Rate (2-step, each type).

C.II. Evaluation of the bit width of fixed-point arithmetics in 5-class identification

Type B can be divided into BD, BN and BV, and so, 5-class classification is possible. It can improve the classification performance because difference between other classes becomes clear.

The results of 5-class identification with 10 times 10-fold Cross Validation are shown in Fig. 6, 7. The identification performance of type BD and type BV is lower than the others, since learning images of those are fewer than the others.

C.III. Evaluation of the bit width of fixed-point arithmetics in 2-class identification

There is a minimum requirement for correctly discriminate between type A and type B/C3 in order to realize diagnosis support. In this section, we examine a classification performance of type A and the others.

Results are shown in Fig. 8, and 9. As the result, identification accuracy of 97%, which is enough in diagnosis support, is obtained.

C.IV.Evaluation of the bit width of fixed-point arithmetics in 2-step identification for 3-class classification

Since the 2-class classifier described in Sec. C.III exhibit sufficient performance, we consider a configuration in which 3-class identifier can be performed by 2-step identification. The first step occurs with types A and B/C3 identification. The second step occurs by type B and C3 identification as shown in Fig. 12. Table 2 shows the number of SV of each identifier in this chapter. By employing 2-step identifier, it is possible to reduce by approximately 10% SV without losing most of the identification performance.

In addition to identification methods, it may be possible to further improve the classification performance by optimizing and equalizing the number of learning images.

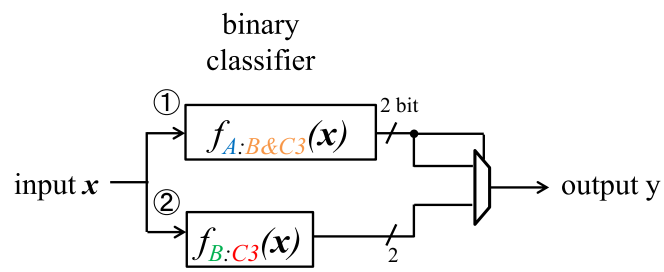


Fig. 12 : Architecture of 3-class identifier realized with two step identifiers.

Table 2 : The number of Support Vector (SV).

composition	Type A	Type BD	Type BN	Type BV	Type C3	All
3-class	68		164		154	386
5-class	74	48	156	62	171	511
2-class	62			72		134
2 step	62		141		142	344

V. Architecture of 3-class identifier using SVM

A. Basic architecture

Fig. 13 shows a block diagram of a 3-class SVM identification module [11]. It consists of three independent 2-class classifiers. The 2-class identifier is implemented using 2 pipeline stage architecture. The 1st stage is used for inner-product computation and the 2nd stage is used for summation. Support Vector (SV) and coefficients obtained by learning step are read from memories.

B. Customizable architecture

Fig. 14 shows a customizable architecture identified by the combination of two-class classifiers. In order to customize the architecture, there is a need to change the configuration of the summation circuits in the subsequent stage by the number of classes. This customization can be implemented with a small area in the circuit. By combining various identification methods, further improvement in the identification accuracy can be expected.

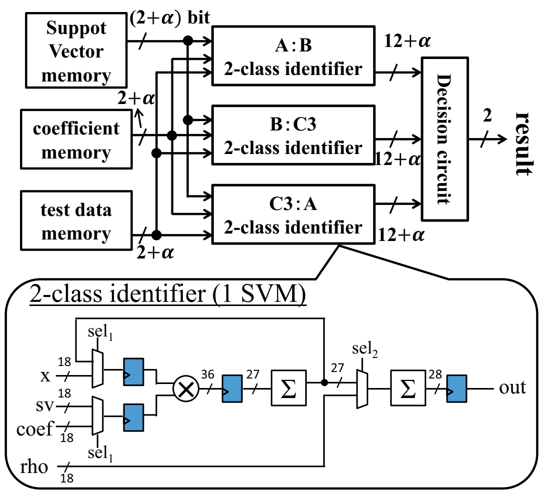


Fig. 13 : A block diagram of a 3-class identifier based on 3 SVMs.

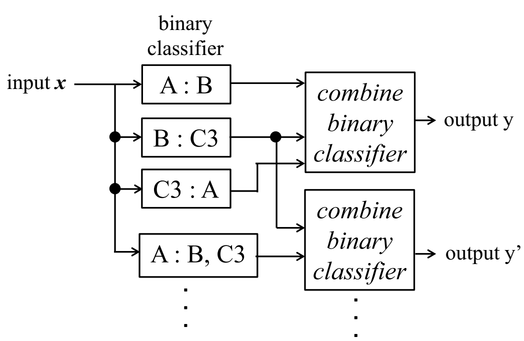


Fig. 14 : Customizable architecture.

C. Improved architecture

When the 2-class identifiers are separately implemented as shown in Fig. 13, the inner-product processing increases 18 % due to the overlap in input data as shown in table 3. The duplication in inner-product computation can be prevented by processing all the inner-products before adding the selected ones for specific 2-class identification. Fig. 15 shows the improvement in implementation of 3-class identifier on Fig. 13. The design includes 2 parts, the inner-products computation and the summation for three 2-class identifiers.

Table 3 : The number of Support Vector (SV) .

Type	# of SV	# of SV used in binary identifier		
		A : B	B : C3	C3 : A
Type A	68	51	-	51
Type B	164	60	132	-
Type C3	154	-	135	45
All	386	111	266	96
		474		

This architecture includes 2 pipeline stages as shown in Fig. 15. In the 1st stage, the inner-products are computed and stored to the corresponding registers.

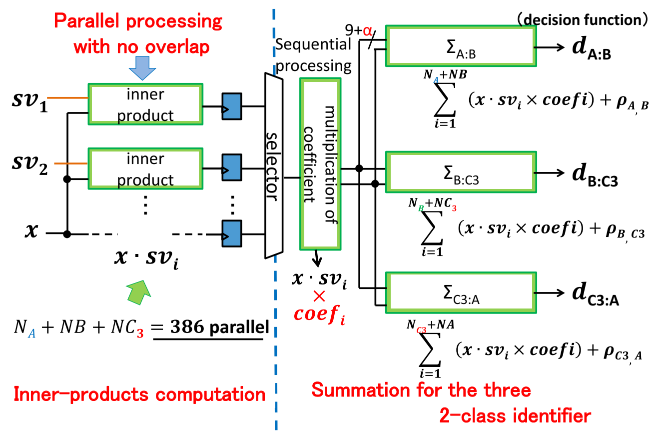


Fig. 15 : Proposed architecture of one-versus-one 3-class identifier.

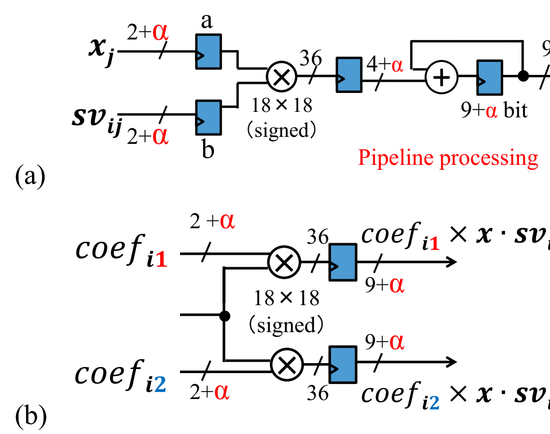


Fig. 16 : (a). A circuit of inner product.
(b). A circuit of multiplication of two coefficients.

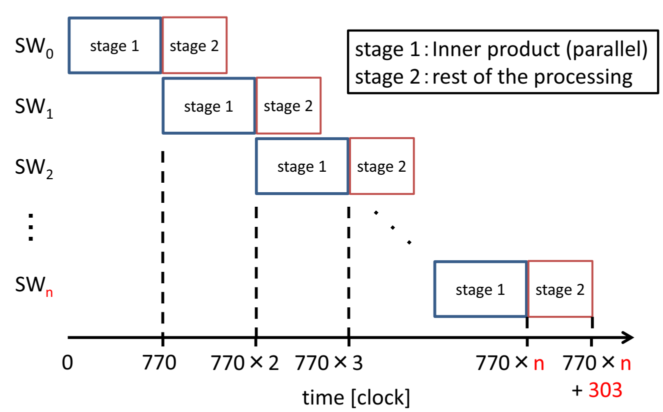


Fig. 17 : Pipeline processing of proposed architecture.

In the 2nd stage, the required inner-products for each 2-class identifier are multiplied with the corresponding coefficient before adding to Σ modules. The inner-product computation and the Σ computation themselves occur in pipeline. This implementation decreases the hardware size and increases the throughput of the whole identifier system as shown in Table 4.

Table 4 : Processing speed of each algorithm.

Version		Standard	Ver.2	Ver.3
Circuit Constitution	parallelizing of innner product	×	○	○
	Pipeline Processing	×	×	○
	multiplier (DSP)	510	289	289
Processing speed	Frequency [MHz]	262	–	–
	Throughput [fps]	3.7	5.2	7.4
	Comparison with software	× 260	× 370	× 520
FPGA implementation		only 2−class	–	–

VI. FPGA implementation

We have implemented the SVM based 2-class classifiers in Fig. 13 into FPGA. The platform is shown in Fig. 18. The PROC Wizard [10] is used for the interface code generation between host (PC) and FPGA board. Table 5 shows the resource utilization. The results show that, it is possible to achieve a throughput of 4.7 fps in the processing of Full HD.

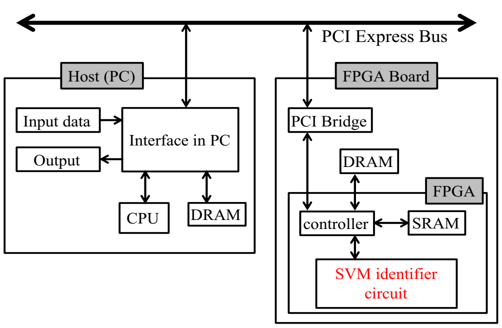


Fig. 18: A block diagram of our platform.

Table 5 : Resource utilization of Stratix II FPGA.

Device	Stratix II
ALUT	2,249 / 143,520 (2%)
Registers	3051
block memory [bit]	36,864 / 9,383,040 (< 1%)

VII. Conclusions

This paper shows the tradeoff between hardware size and identification accuracy of the type identification circuit, which uses fixed-point number with various bit length. Three identification methods of 3-class, 5-class and 2-class classifications are investigated. A pipeline architecture for the identifier is also introduced to increase the throughput of the system. The simulation result shows that it is possible to realize real-time processing of our system.

A customizable architecture, which combines two different identification methods with low penalty in hardware size is introduced to improve the identification accuracy. Investigation on this architecture is our future issue.

Acknowledgements

Part of this work was supported by Grant-in-Aid for Scientific Research(C) JSPS KAKENHI Grant Numbers 23560400 and 2459102, and Adaptable and Seamless Technology Transfer Program through Target-driven R&D, JST.

References

[1] T. Tamaki, et al., “Computer-aided colorectal tumor classification in NBI endoscopy using local features, medical image analysis,” Available online 13 September 2012, ISSN 1361-8415, 10.1016/j.media.2012.08.003. (2012).

[2] H. Kanao, et al., “Narrow–band imaging magnification predicts the histology and invasion depth of colorectal tumors,” Journal of Gastrointestinal Endoscopy, vol. 69, no.3, pp. 631-636, 2009.

[3] http://www.vlfeat.org/

[4] http://www.csie.ntu.edu.tw/~cjlin/libsvm/

[5] T. Mishima, et al., "A Simple and Effective Hardware Oriented Feature Extraction Algorithm for Colorectal Endoscopic Images with NBI Magnification,” Proc. of the 28th International Conference on Circuits / Systems, Computers and Communications (ITC-CSCC2013), pp. 567-570, June 30-July 3, Yeosu, Korea, 2013.

[6] V. Vapnik, “Statistical learning theory,” John Wiley & Sons Inc., New York, 1998.

[7] V. Vapnik, “The nature of statistical learning theory,” Springer Verlag, Berlin, 1995.

[8] E.L. Allwein, et al., “Reducing multiclass to binary : a unifying approach for margin classifier,” Journal of Machine Learning Research, 113-141.

[9] C. Platt. John, et al., “Large margin dags for multiclass classification,” Neural Information Processing Systems 12 (NIPS1999), pp. 547-553.

[10] http://www.gidel.com/procwizard.htm.

[11] S. Shigemi, et al., “An FPGA Implementation of Support Vector Machine Identifier for Colorectal Endoscopic Images with NBI Magnification,” Proc. of the 28th International Conference on Circuits / Systems, Computers and Communications (ITC-CSCC2013), pp. 571-572, June 30-July 3, Yeosu, Korea, 2013.