

ORIGINAL RESEARCH

Performance improvement in computerized detection of cerebral aneurysms by retraining classifier using feedback data collected in routine reading environment

Yukihiro Nomura¹, Yoshitaka Masutani², Soichiro Miki¹, Mitsutaka Nemoto¹, Shouhei Hanaoka³, Takeharu Yoshikawa¹, Naoto Hayashi¹, Kuni Ohtomo³

1. Department of Computational Diagnostic Radiology and Preventive Medicine, The University of Tokyo Hospital, Tokyo, Japan. 2. Graduate School of Information Sciences, Hiroshima City University, Hiroshima, Japan. 3. Department of Radiology, The University of Tokyo Hospital, Tokyo, Japan.

Correspondence: Yukihiro Nomura. Address: Department of Computational Diagnostic Radiology and Preventive Medicine, The University of Tokyo Hospital, 7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan. E-mail: nomuray-tky@umin.ac.jp

Received: September 29, 2014

Accepted: October 18, 2014

Online Published: October 27, 2014

DOI: 10.5430/jbgc.v4n4p12

URL: <http://dx.doi.org/10.5430/jbgc.v4n4p12>

Abstract

Introduction: The performance of computer-assisted detection (CAD) software depends on the quality and quantity of the dataset used for supervised learning. To realize the continuous clinical use and performance improvement of CAD software, it is necessary to continuously collect data for supervised learning in practical use and to improve CAD software by retraining with the collected data. In this study, we investigated the performance improvement of cerebral aneurysm detection software based on retraining the classifier through a simulation-based study.

Methods: We collected data for retraining during the practical use of our cerebral aneurysm detection software and retrained the classifier for false positive (FP) reduction using the collected data. The effect on improving the performance was compared by changing the number of training cases and the training algorithms.

Results: The performance was improved significantly ($p < .05$) by retraining using additional training cases. In contrast, there were no statistical differences in the performance upon retraining among the four training algorithms for boosting. The sensitivity at 3 FPs/case was improved from 81.5% to 89.5% by retraining with additional training cases.

Conclusions: The performance of the software was effectively improved by adding training cases rather than by changing the training algorithm.

Key words

Computer-assisted detection, Cerebral aneurysm, Magnetic resonance angiography, Retraining, Boosting

1 Introduction

Computer-assisted detection/diagnosis (CAD) software has been developed by many research groups, and commercial CAD software has also been released^[1-3]. However, most radiologists have not yet utilized CAD software in daily practice since the performance of current CAD software has certain limitations^[3].

The performance of CAD software depends on the quality and quantity of the dataset for supervised learning. If the characteristics of data used in development and in practical use are different, the performance of CAD software will be degraded. The data characteristics may be changed by updating the imaging hardware or imaging protocols for the software, even in a single institute. Therefore, it is desirable to use the same data characteristics for development and practical use. However, it is difficult to collect a dataset of significant quantity and quality in the initial development of CAD software. Hence, it is necessary to continuously collect data for supervised learning in practical use and to update CAD software by retraining it with the collected data.

To realize the continuous clinical use and performance improvement of CAD software, we defined a roadmap for the development, clinical use, and refinement of CAD software as follows:

- Phase I: Development and implementation of the algorithm and initial training classifier using the dataset for the initial development.
- Phase II: Clinical use and evaluation of the CAD software implemented in Phase I (including data collection for refinement).
- Phase III: Algorithm refinement including the reselection of the classifier and the retraining of the classifier using the data collected in Phase II, followed by further clinical use.

To realize Phase II in a routine reading environment, we previously developed a web-based CAD software processing and evaluation platform (clinical infrastructure for radiologic computation of united solutions clinical server, CIRCUS CS), which provides the on-line processing of CAD software and interfaces to evaluate the results obtained from CAD software (clinical feedback) ^[4,5]. Two kinds of CAD software, which are for cerebral aneurysm detection in magnetic resonance angiography (MRA) images and lung nodule detection in computed tomography (CT) images of the chest ^[6], have been used and evaluated in the screening section of our institute since 2009. The collected feedback data enable us to improve the performance of CAD software by retraining the classifier. From our preliminary study on cerebral aneurysm detection, the performance of CAD software was improved by retraining the classifier by adding positive cases ^[5]. However, the number of positive cases used for retraining in the previous study was at most 110 cases. To clarify the performance improvement of CAD software by retraining the classifier, further investigation of the strategies for retraining, such as the use of even more training cases and the reselection of the classifier, is required.

As a related study, several research groups have developed CAD software based on a life cycle including the design, development, and evaluation of CAD software ^[7,8]. However, the refinement of CAD software after clinical use (Phase III described above) was not realized.

In this study, we investigated a strategy to improve the performance of cerebral aneurysm detection software based on retraining using feedback data collected in the routine reading environment.

2 Materials and methods

2.1 Algorithm of cerebral aneurysm detection

Figure 1 shows a flowchart of our detection method. The algorithm consists of three steps: preprocessing, lesion candidate extraction, and false positive (FP) reduction.

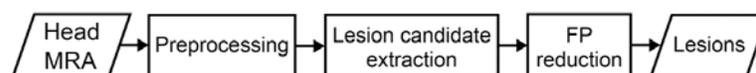


Figure 1. Flowchart of our detection method

2.1.1 Preprocessing

First, MRA images are resampled using trilinear interpolation to obtain the isotropic volume. The resampled voxel size is equal to the pixel size in the axial slice. After that, vessel segmentation is carried out. The steps in the processing procedure for vessel segmentation are described as follows:

- 1) The mean and standard deviation of brain voxels, denoted μ_{brain} and σ_{brain} , are calculated from the center quarter region of the two central axial slices of the volume data.
- 2) The initial mask for the brain is extracted as the set of connected voxels with voxel value $\geq \mu_{brain} - 2\sigma_{brain}$. After that, closing with a spherical kernel of 5.9 voxel radius and erosion with a spherical kernel of 9.9 voxel radius are applied to obtain the final brain mask \mathbf{M}_{brain} .
- 3) The initial vessel region within \mathbf{M}_{brain} is extracted as the set of connected voxels with voxel value $\geq \mu_{brain} + 2.5\sigma_{brain}$. The number of extracted voxels is defined as V_{init} .
- 4) Connected component analysis is performed, and then at most two components with the following conditions are selected.
 - The number of connected voxels is greater than or equal to 2.7% of V_{init} .
 - The x -coordinates of the center of gravity is between 0.25 and 0.75 of the width of \mathbf{M}_{brain} .
 - The y -coordinates of the center of gravity is between 0.32 and 0.72 of the height of \mathbf{M}_{brain} .
- 5) Dilation with a spherical kernel of 1.9 voxel radius is carried out. The result is defined as \mathbf{R}_{vessel} .

2.1.2 Lesion candidate extraction

The shape index $S(\mathbf{x}, \sigma)$ [9, 10] is calculated at the voxels of \mathbf{R}_{vessel} using the original voxel values, where \mathbf{x} is the 3D-coordinates for the position of the voxel and σ (voxel) is the standard deviation of the Gaussian filter. In this study, we adopted multiscale integration based on the Hessian [11, 12] eigenvalues [6]. Let the eigenvalues of Hessian be $\lambda_1(\mathbf{x}, \sigma)$, $\lambda_2(\mathbf{x}, \sigma)$, and $\lambda_3(\mathbf{x}, \sigma)$ ($\lambda_1(\mathbf{x}, \sigma) \leq \lambda_2(\mathbf{x}, \sigma) \leq \lambda_3(\mathbf{x}, \sigma)$). In our multiscale integration, Hessian eigenvalues are calculated for several values of σ , and $\sigma_{opt}(\mathbf{x})$ is determined as follows:

$$\sigma_{opt}(\mathbf{x}) = \arg \min_{\sigma} (\sigma^2 \lambda_1(\mathbf{x}, \sigma)) \quad (1)$$

The range of σ is set to $\{1, 2, 3, 4, 5\}$ voxels in experiments. After that, $S_{opt}(\mathbf{x})$ is calculated as follows:

$$S_{opt}(\mathbf{x}) = S(\mathbf{x}, \sigma_{opt}(\mathbf{x})) \quad (2)$$

After calculating S_{opt} , a voxel-based differentiation process based on a classifier using the Mahalanobis distance ratio [13] is employed at voxels with $S_{opt}(\mathbf{x}) \geq 0.75$. The feature values of voxel-based differentiation are as follows:

- voxel value ($I(\mathbf{x})$)
- principal curvatures ($k_1, k_2, k_1/k_2$) [9, 10]
- $S_{opt}(\mathbf{x})$
- magnitude of first derivatives of $I(\mathbf{x})$ ($|\nabla I(\mathbf{x})|$)
- $S_{blob}, S_{line}, S_{bif}$

S_{blob} , S_{line} , and S_{bif} , which enhances the blob, line, and bifurcation derived from the Hessian eigenvalues, respectively, are defined as follows:

$$S_{blob}(\mathbf{x}, \sigma) = \begin{cases} \frac{|\lambda_3(\mathbf{x}, \sigma)|}{|\lambda_1(\mathbf{x}, \sigma)|} & \lambda_1(\mathbf{x}, \sigma) \leq \lambda_2(\mathbf{x}, \sigma) \leq \lambda_3(\mathbf{x}, \sigma) < 0 \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

$$S_{line}(\mathbf{x}, \sigma) = \begin{cases} \frac{|\lambda_2(\mathbf{x}, \sigma)| - |\lambda_3(\mathbf{x}, \sigma)|}{|\lambda_1(\mathbf{x}, \sigma)|} & \lambda_1(\mathbf{x}, \sigma) \leq \lambda_2(\mathbf{x}, \sigma) < 0 \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

$$S_{bif}(\mathbf{x}, \sigma) = s(\mathbf{x}, \sigma)h(\mathbf{x}) \quad (5)$$

$$s(\mathbf{x}, \sigma) = \begin{cases} \frac{|\lambda_1(\mathbf{x}, \sigma) - \lambda_2(\mathbf{x}, \sigma)| |\lambda_2(\mathbf{x}, \sigma) - \lambda_3(\mathbf{x}, \sigma)|}{|\lambda_1(\mathbf{x}, \sigma)| |\lambda_2(\mathbf{x}, \sigma)|} & \lambda_1(\mathbf{x}, \sigma) \leq \lambda_2(\mathbf{x}, \sigma) < 0 \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

$$h(\mathbf{x}) = \exp\left(-\frac{|\nabla I(\mathbf{x})|^2}{2\alpha^2}\right) \quad (7)$$

α is set to 30 in experiments. The voxels with a Mahalanobis distance ratio of greater than or equal to 2.0 are extracted. After that, conditional closing within \mathbf{R}_{vessel} with a spherical kernel of 1.9 voxel radius and connected component analysis are performed. All components fewer than 10 voxels are discarded.

2.1.3 FP reduction

A classifier ensemble trained by the boosting algorithm ^[14] is employed to classify lesion candidates into true positives (TPs) and FPs based on 63 feature values of the candidates. Each of the weak classifiers is formed from a feature measured at a region and its threshold. The feature values are as follows:

- Numbers of voxels (V) and surface voxels
- statistics of voxel values (minimum, maximum, mean, second moment, standard deviation, skewness, kurtosis, and entropy)
- contrast measures (*Contrast1* and *Contrast2* in ^[15])
- sphericity
- ratio of V to the number of voxels in the bounding box
- similarity of sphere (S_s)
- statistics of the distance between the center of the candidate and its boundary (minimum, maximum, mean, second moment, standard deviation, skewness, kurtosis, and minimum/maximum)
- statistics of k_1 , k_2 , k_1/k_2 , shape index, curvedness ^[9, 10], σ_{opt} , $|\nabla I|$, S_{blob} , S_{line} , and S_{bif} (minimum, maximum, mean, and standard deviation)

The similarity of sphere in the i th lesion candidate is defined as follows:

$$S_s(i) = \frac{V_s(i)}{V(i)} \quad (8)$$

where $V_s(i)$ is the number of intersecting voxels between the i th lesion candidate and the volume of the equivalent sphere whose center is defined as the center of gravity of the candidate.

2.2 CIRCUS CS

Our cerebral aneurysm detection software described previous subsection was installed as a plug-in for CIRCUS CS [4, 5]. Figure 2 shows an example of the result of cerebral aneurysm detection displayed in CIRCUS CS. In Figure 2, the top three lesion candidates with the highest likelihood are displayed. A radiologist labeled each lesion candidate as follows:

- known TP: cerebral aneurysm ≥ 2 mm in diameter recognized before CAD reference
- missed TP: cerebral aneurysm ≥ 2 mm in diameter overlooked before CAD reference
- FP: false positive of the CAD software
- pending: difficult to label as TP or FP

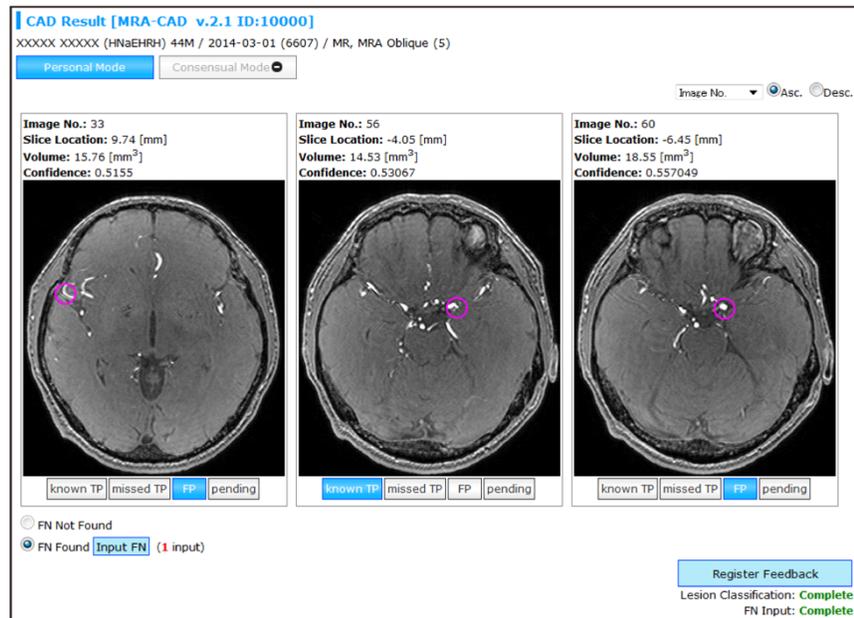


Figure 2. Example of CAD result displayed in CIRCUS CS. The top three lesion candidates are displayed, each of which has toggle buttons to indicate “known TP”, “missed TP”, “FP”, and “pending”.

If the lesion detected by the radiologist was not included in the three candidates detected by the CAD software, the radiologist manually entered the location of the aneurysm by clicking on the 2D image (see Figure 3).

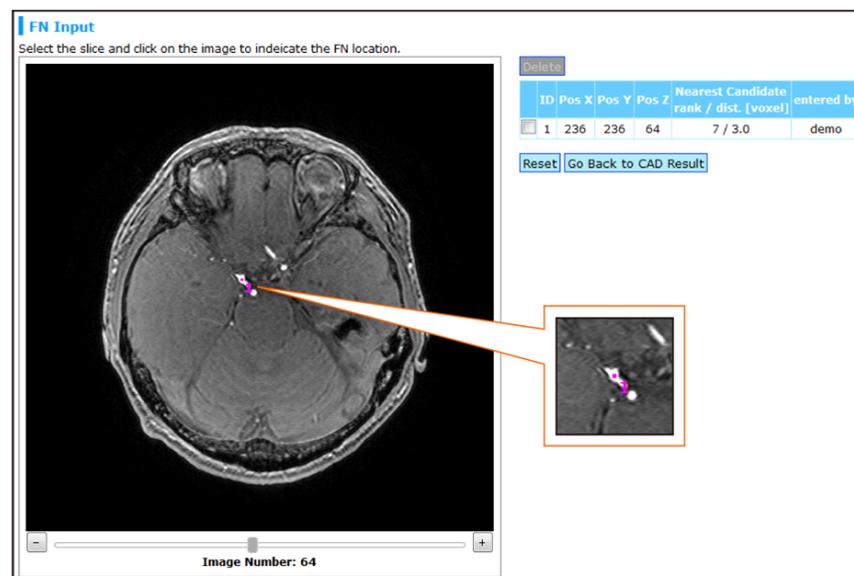


Figure 3. Interface for entering false negative locations. A radiologist enters the locations of correct aneurysms, if not indicated by CIRCUS CS, by clicking on the 2D image.

In the screening section of our institute, the double-reading method combined with CIRCUS CS was used with the procedure. First, two radiologists independently interpreted MRA images without seeing the CAD results. After that, they independently reviewed the CAD results and independently registered their personal feedback into CIRCUS CS. Finally, the two radiologists discussed and determined the final clinical diagnosis and the consensual feedback was registered.

2.3 Simulation of retraining using data collected in routine reading environment

We investigated the performance improvement of our cerebral aneurysm detection software by simulation for the case of retraining the classifier using feedback data collected in the routine reading environment. In this study, we focused on the classifier for FP reduction. In contrast, we did not retrain the voxel-based classifier for lesion candidate extraction. This is because there was little improvement in performance from the result of a pilot experiment in which the voxel-based classifier in the cerebral aneurysm detection software was retrained by adding 15 positive cases including false negatives (FNs).

2.3.1 Datasets

This study was approved by the ethical review board of our institute. Three-dimensional time-of-flight unenhanced MRA was performed as a part of brain screening using three 3T MR scanners (two Signa HDxt and one Discovery MR750, GE Healthcare, Waukesha, WI, USA). The acquisition parameters were as follows: field-of-view, 240 mm; matrix size, 512×512; pixel spacing, 0.469 mm; slice thickness, 1.2 mm; slice interval, 0.6 mm; repetition time, 25 msec; echo time, 2.7 msec for Signa HDxt and 2.9 msec for Discovery MR750; flip angle, 15 degree. In this study, we utilized the following three datasets.

Dataset for initial development: This dataset contains image files and clinical properties including areas of lesions defined by pixel-by-pixel painting. These data were registered in our image database system, named CIRCUS DB^[4].

Dataset for additional data: This dataset includes 309 positive cases collected between August 2009 and February 2012. The dataset is divided into two groups. The first group, called “group 1”, includes 155 positive cases collected between August 2009 and March 2011. The second group, called “group 2”, includes remaining 154 positive cases.

Dataset for evaluation: This dataset includes the cases of 1,779 first-time visitors to the screening section of our institute between March 2012 and December 2013.

The datasets for additional data and evaluation contain the results obtained from CAD software and consensual feedback data. Table 1 shows the numbers of cases in the datasets. Each positive case includes at least one aneurysm of 2 mm or more in diameter, which was determined by consensual reading by two experienced radiologists. Table 2 shows the size distribution of aneurysms for each dataset.

Table 1. Number of cases for each dataset

Dataset for	Number of cases	
	Total cases	Positive cases
initial development	181	53
additional data (group 1)	155	155
additional data (group 2)	154	154
evaluation	1,779	110

Table 2. Size distribution of aneurysms for each dataset

Size (mm)	Dataset for			
	Initial development	Additional data		Evaluation
		Group 1	Group 2	
2	41	58	71	51
3	24	75	67	38
4	17	30	23	25
5	8	8	8	4
≥ 6	10	9	5	6
Total	100	180	174	124

2.3.2 Simulation conditions

Figure 4 shows the relationship between the training data and test data in the simulation. Retraining and evaluation were performed by adding each group of additional data (total: 2 steps). We also compared four types of training algorithm for boosting, which were AdaBoost^[14], MadaBoost^[16], cost-sensitive^[17] AdaBoost, and cost-sensitive MadaBoost. MadaBoost has higher tolerance to outliers of feature vectors than AdaBoost. A cost-sensitive algorithm is an effective means of classification when training data is imbalanced. In this study, we adopted the *AdaC2* algorithm^[17]. The number of weak classifiers used for boosting was set to 300 in the experiments. The area under the receiver operating characteristic (ROC) curve (AUC) was utilized as the evaluation criterion. We calculated the AUC using the R pROC package (version 1.7.1)^[18] and also carried out a bootstrap test to compare the effects of differences in the number of training cases and the training algorithm, respectively. The number of bootstrap replicates was set to 100.

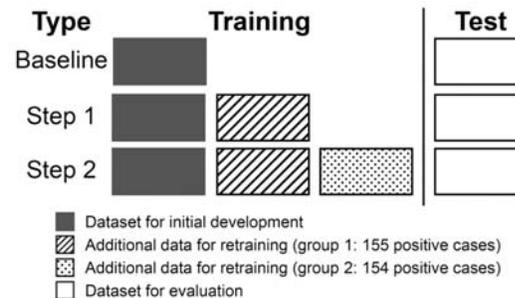


Figure 4. Relationship between training data and test data for simulation.

3 Results

Figure 5 shows learning curves obtained by simulation. From the results of the bootstrap test, the difference between the baseline and step 1 was significant for all the training algorithms ($p < .05$). Also, the difference between the baseline and step 2 was significant for all the training algorithms ($p < .05$). Consequently, the AUC was improved by increasing the number of training cases for all the training algorithms.

Regarding the difference between the training algorithms, MadaBoost was inferior to the other algorithms at all steps. From the results of the bootstrap test, the difference between MadaBoost and AdaBoost or cost-sensitive AdaBoost was significant at the baseline ($p < .05$). However, there were no statistical differences among the AUCs for the four training algorithms at step 1 and step 2.

Figure 6 shows the free-response ROC (FROC) curve for each step, and Table 3 shows the sensitivities at 3, 5, and 9 FPs/case derived from the FROC curves shown in Figure 6. The training algorithm for each step was selected on the basis of the maximization of the AUC. From these results, the sensitivity was improved by increasing the number of training cases, particularly at small FPs/case. As shown in Table 3, the sensitivity at 3 FPs/case was improved from 81.5% to 89.5% by retraining.

Figure 5. Learning curves for retraining of classifier

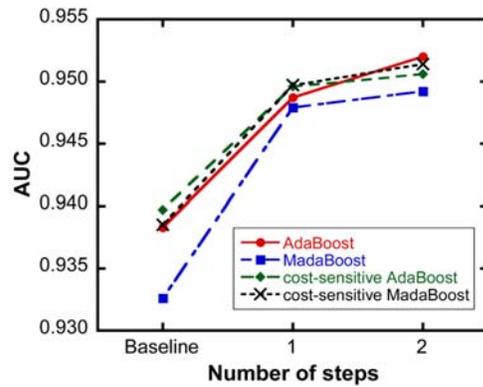


Figure 6. FROC curve for each step. The training algorithm for each step was selected on the basis of the maximization of the AUC

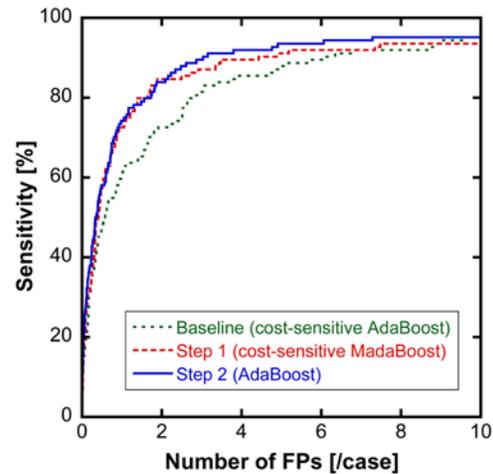


Table 3. Sensitivity of our cerebral aneurysm detection method derived from Figure 6 (Unit: %)

Step	FPs/case		
	3	5	9
Baseline	81.5	87.0	94.4
Step 1	87.1	91.1	93.5
Step 2	89.5	93.5	95.2

4 Discussion

The results of the simulation-based study show that the performance of the cerebral aneurysm detection software was improved by retraining with data collected in the routine reading environment. According to the results shown in Figure 5, the performance of the software was improved by increasing the number of training cases for all the training algorithms. However, there were no statistical differences in the AUCs between step 1 and step 2. This is because the effect on improving the performance tends to decrease with increasing number of training cases^[19]. If the performance becomes saturated in spite of adding more training cases, there is scope for algorithm refinement including the refinement of feature values using the data collected during practical use. Regarding the training algorithms, the AUC of MadaBoost was inferior to that of the other algorithms at all steps. However, there were no statistical differences in the AUCs among the training algorithms at step 1 and step 2. In this study, the performance of our software was effectively improved by adding training cases rather than by changing the training algorithm. The effects of our software on radiologists’ detection performance in the routine reading environment including inter-observer and intra-observer variability are under way.

A number of research groups have reported a variety of computerized detection methods for cerebral aneurysms^[20-25]. Yang’s method^[23] which was validated using the largest number of cases among these studies, achieved sensitivities of

80% at 3 FPs/case and 95% at 5 FPs/case. From the results shown in Table 3, the baseline in our method was equivalent in performance to Yang's method, and our method after retraining was superior to Yang's method.

At our institute, the initial version of the cerebral aneurysm detection software, in which the classifier for FP reduction was trained by the AdaBoost algorithm using the dataset for the initial development, was utilized from August 2009 to March 2012. Since April 2012, a refined version of the software, in which the classifier for FP reduction was trained by the cost-sensitive MadaBoost algorithm using the dataset for the initial development with additional data, has been utilized.

The limitations of this study are discussed as follows. First, the voxel-based classifier for lesion candidate extraction was not retrained on the basis of the results of the pilot experiment by adding a small dataset. To train the voxel-based classifier using more cases, it is necessary to define the areas of lesions by pixel-by-pixel painting. However, it is time-consuming to paint lesions on images. To reduce the time required to paint a lesion in the routine reading environment, the development of a painting interface with semi-automatic/interactive image segmentation functions is required. Such an interface is under development as part of the refinement of CIRCUS DB.

The second limitation is that the datasets used for retraining and evaluation were collected from only three 3T MR scanners in our institute. The quality of MRA images depends on the scanner, particularly the strength of the static magnetic field^[26]. For example, artificial signal loss of the vessel is observed due to magnetic susceptibility artifacts caused by adjacent bone or air. In general, the extent of the artifact depends on the strength of the static magnetic field^[27]. Hence, it is necessary to investigate a strategy that can be used to improve the performance of CAD software for each scanner. To effectively collect data for retraining from various scanners in multiple institutions, we have implemented CIRCUS CS in a teleradiology environment, and a multicenter trial has been underway since September 2011^[28].

5 Conclusions

We have investigated the performance improvement of our cerebral aneurysm detection software based on retraining the classifier for FP reduction using feedback data collected in the routine reading environment. According to the results, the performance of the software was effectively improved by adding training cases rather than by changing the training algorithm.

Acknowledgments

The Department of Computational Radiology and Preventive Medicine, The University of Tokyo Hospital is sponsored by High Technology Medical Complex Co., LTD., Siemens Japan K.K., and GE Healthcare Japan. This work was supported by JSPS Grant-in-Aid for Young Scientists (B) Grant Number 25750182.

References

- [1] Giger ML, Chan HP, Boone J. Anniversary paper: History and status of CAD and quantitative image analysis: the role of Medical Physics and AAPM. *Med Phys.* 2008 Dec; 35(12): 5799-820. PMID: 19175137. <http://dx.doi.org/10.1118/1.3013555>
- [2] Boyer B, Balleyguier C, Granat O, Pharaboz C. CAD in questions/answers: Review of the literature. *Eur. J. Radiol.* 2009 Jan; 69(1): 24-33. PMID: 18977103. <http://dx.doi.org/10.1016/j.ejrad.2008.07.042>
- [3] van Ginneken B, Schaefer-Prokop CM, Prokop M. Computer-aided diagnosis: how to move from the laboratory to the clinic. *Radiology.* 2011 Dec; 261(3): 719-32. PMID: 22095995. <http://dx.doi.org/10.1148/radiol.11091710>
- [4] Nomura Y, Hayashi N, Masutani Y, Yoshikawa T, Nemoto M, Hanaoka S, *et al.* CIRCUS: an MDA platform for clinical image analysis in hospitals. *T. Mass-Data Analysis of Images and Signals.* 2010 Sep; 2(1): 112-27.
- [5] Masutani Y, Nemoto M, Nomura Y, Hayashi N. Clinical machine learning in action: CAD system design, development, tuning, and long-term experience. *Machine Learning in Computer-Aided Diagnosis.* In Suzuki K ed.: *Machine learning in computer-aided diagnosis: Medical imaging intelligence and analysis.* Hershey, PA. IGI Global. 2012; 159-176.
- [6] Nomura Y, Nemoto M, Masutani Y, Hanaoka S, Yoshikawa T, Miki S, *et al.* Reduction of false positives at vessel bifurcations in computerized detection of lung nodules. *J. Biomed. Graph. Comput.* 2014 Aug; 4(3): 36-46.

- <http://dx.doi.org/10.5430/jbgc.v4n3p36>
- [7] Piętka E, Kawa J, Spinczyk D, Badura P, Więclawek W, Czajkowska J, *et al.* Role of radiologists in CAD life-cycle. *Eur J Radiol.* 2011 May; 78(2): 225-33. PMID: 19783393. <http://dx.doi.org/10.1016/j.ejrad.2009.08.015>
- [8] Marrocco C, Molinara M, Tortorella F, Rinaldi P, Bonomo L, Ferrarotti A, *et al.* A multidisciplinary approach to the design and development of a CAD system for the detection of clustered microcalcifications. *Proc of the 24th IEEE Symp Comput Based Med Syst.* 2011 Jun; 1-6. <http://dx.doi.org/10.1109/CBMS.2011.5999048>
- [9] Koenderink JJ. *Solid Shape.* Cambridge, MA: MIT Press. 1990.
- [10] Dorai C, Jain AK. COSMOS: a representation scheme for 3D free-form objects. *IEEE Trans Pattern Anal Mach Intell.* 1997 Oct; 19(10): 1115-30. <http://dx.doi.org/10.1109/34.625113>
- [11] Lorenz C, Carlsen IC, Buzug TM, Fassnacht C, Weese J. Multi-scale line segmentation with automatic estimation of width, contrast and tangential direction in 2D and 3D medical images. *Lecture Notes in Computer Science.* 1997 Mar; 1205: 233-42. <http://dx.doi.org/10.1007/BFb0029242>
- [12] Sato Y, Nakajima S, Shiraga N, Atsumi H, Yoshida S, Koller T, *et al.* Three-dimensional multi-scale line filter for segmentation and visualization of curvilinear structures in medical images. *Med Image Anal.* 1998 Jun; 2(2): 143-68. [http://dx.doi.org/10.1016/S1361-8415\(98\)80009-1](http://dx.doi.org/10.1016/S1361-8415(98)80009-1)
- [13] Nemoto M, Shimizu A, Hagihara Y, Kobatake H, Nawano S. Improvement of tumor detection performance in mammograms by feature selection from a large number of features and proposal of fast feature selection method. *Syst Comput Jpn.* 2006 Dec; 37(12): 56-68. <http://dx.doi.org/10.1002/scj.20498> (2006)
- [14] Schapire RE, Freund Y, Bartlett P, Lee WS. Boosting the margin: a new explanation for the effectiveness of voting methods. *Ann Stat.* 1998 Oct; 26(5): 1651-86. <http://dx.doi.org/10.1214/aos/1024691352>
- [15] te Brake GM, Karssemeijer N, Hendriks JH. An automatic method to discriminate malignant masses from normal tissue in digital mammograms. *Phys Med Biol.* 2000 Oct; 45(10): 2843-57. PMID: 11049175. <http://dx.doi.org/10.1088/0031-9155/45/10/308>
- [16] Domingo C, Watanabe O. MadaBoost: A Modification of AdaBoost. *Proc. of the 13th Conf. on Computational Learning Theory.* 2000; 180-9.
- [17] Sun Y, Kamel MS, Wong AKC, Wang Y. Cost-sensitive boosting for classification of imbalanced data. *Pattern Recogn.* 2007 Dec; 40(12): 3358-78. <http://dx.doi.org/10.1016/j.patcog.2007.04.009>
- [18] Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, *et al.* pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics.* 2011 Mar; 12: 77. PMID: 21414208. <http://dx.doi.org/10.1186/1471-2105-12-77>.
- [19] K. Fukunaga. *Introduction to statistical pattern recognition*, 2nd ed. San Diego, CA: Academic Press Professional Inc. 1990.
- [20] Arimura H, Li Q, Korogi Y, Hirai T, Abe H, Yamashita Y, *et al.* Automated computerized scheme for detection of unruptured intracranial aneurysms in three-dimensional magnetic resonance angiography. *Acad Radiol.* 2004 Oct; 11(10): 1093-104. PMID: 15530802. <http://dx.doi.org/10.1016/j.acra.2004.07.011>
- [21] Uchiyama Y, Ando H, Yokoyama R, Hara T, Fujita H, Iwama T. Computer-aided diagnosis scheme for detection of unruptured intracranial aneurysms in MR angiography. *Conf Proc IEEE Eng Med Biol Soc.* 2005 Jan; 3: 3031-4. PMID: 17282882. <http://dx.doi.org/10.1109/IEMBS.2005.1617113>
- [22] Kobayashi S, Kondo K, Hata Y. Computer-aided diagnosis of intracranial aneurysms in MRA images with case-based reasoning. *IEICE Trans Inf & Syst.* 2006 Jan; E89-D(1): 340-50. <http://dx.doi.org/10.1093/ietisy/e89-d.1.340>
- [23] Yang X, Blezek DJ, Cheng LT, Ryan WJ, Kallmes DF, Erickson BJ. Computer-aided detection of intracranial aneurysms in MR angiography. *J Digit Imaging.* 2011 Feb; 24(1): 86-95. PMID: 19937083. <http://dx.doi.org/10.1007/s10278-009-9254-0>
- [24] Hentschke CM, Beuing O, Nickl R, Tonnies KD. Automatic cerebral aneurysm detection in multimodal angiographic images. *Proc of IEEE Nucl Sci Symp and Med Imag Conf.* 2011 Oct; 3116-20. <http://dx.doi.org/10.1109/NSSMIC.2011.6152566>
- [25] Suniaga S, Werner R, Kemmling A, Groth M, Fiehler J, Forkert N. Computer-aided detection of aneurysms in 3D time-of-flight MRA datasets. *Lecture Notes in Computer Science.* 2012 Oct; 7588, 63-9. http://dx.doi.org/10.1007/978-3-642-35428-1_8
- [26] Willinek WA, Born M, Simon B, Tschampa HJ, Krautmacher C, Gieseke J, *et al.* Time-of-flight MR angiography: comparison of 3.0-T imaging and 1.5-T imaging-initial experience. *Radiology.* 2003 Dec; 229(3): 913-20. PMID: 14657322. <http://dx.doi.org/10.1148/radiol.2293020782>
- [27] Matsuura H, Inoue T, Ogasawara K, Sasaki M, Konno H, Kuzu Y, *et al.* Quantitative analysis of magnetic resonance imaging susceptibility artifacts caused by neurosurgical biomaterials: comparison of 0.5, 1.5, and 3.0 Tesla magnetic fields. *Neurol Med Chir (Tokyo).* 2005 Aug; 45: 395-8; discussion 398-9. PMID: 16127256. <http://dx.doi.org/10.2176/nmc.45.395>
- [28] Nomura Y, Masutani Y, Hayashi N, Miki S, Nemoto M, Hanaoka S, *et al.* Additional learning of CAD software based on multicenter trial in teleradiology environment. *Int J CARS.* 2012 June; 8 (Suppl. 1): S355-S357.