Artificial Intelligence Trained by Deep Learning Can Improve Computed Tomography Diagnosis of Nontraumatic Subarachnoid Hemorrhage by Nonspecialists

Toru NISHI,[1] Shigeo YAMASHIRO,[1] Shuichiro OKUMURA,[2] Mizuki TAKEI,[3] Atsushi TACHIBANA,[3] Sadato AKAHORI,[3] Masatomo KAJI,[1] Ken UEKAWA,[1] and Toshihiro AMADATSU[1]

[1]Department of Neurosurgery, Saiseikai Kumamoto Hospital, Stroke Center, Kumamoto, Kumamoto, Japan
[2]Department of Radiology, Saiseikai Kumamoto Hospital, Kumamoto, Kumamoto, Japan
[3]Research & Development Management Headquarters, FUJIFILM Corporation, Tokyo, Japan

Abstract

Subarachnoid hemorrhage (SAH) is a serious cerebrovascular disease with a high mortality rate and is known as a disease that is hard to diagnose because it may be overlooked by noncontrast computed tomography (NCCT) examinations that are most frequently used for diagnosis. To create a system preventing this oversight of SAH, we trained artificial intelligence (AI) with NCCT images obtained from 419 patients with nontraumatic SAH and 338 healthy subjects and created an AI system capable of diagnosing the presence and location of SAH. Then, we conducted experiments in which five neurosurgery specialists, five nonspecialists, and the AI system interpreted NCCT images obtained from 135 patients with SAH and 196 normal subjects. The AI system was capable of performing a diagnosis of SAH with equal accuracy to that of five neurosurgery specialists, and the accuracy was higher than that of nonspecialists. Furthermore, the diagnostic accuracy of four out of five nonspecialists improved by interpreting NCCT images using the diagnostic results of the AI system as a reference, and the number of oversight cases was significantly reduced by the support of the AI system. This is the first report demonstrating that an AI system improved the diagnostic accuracy of SAH by nonspecialists.

Keywords: subarachnoid hemorrhage, misdiagnosis, diagnosis, deep learning, artificial intelligence

Introduction

Cerebral stroke is a cerebrovascular disease that holds a top-ranking position as a cause of death even today. Among stroke subtypes, the mortality rate of patients with subarachnoid hemorrhage (SAH) is particularly high. SAH is a disease condition caused by the rupture of cerebral aneurysms in many cases and is treated by preventing rerupture and managing cerebrovascular spasms with direct surgery or intravascular surgery. Although both of them need sophisticated treatment techniques, the techniques and instruments are improving on a daily basis. It has been reported, however, that the clinical outcome of SAH patients is influenced by their disease grade at the beginning of the treatment.[1,2] Furthermore, it is known that the diagnosis of SAH itself before treatment is not easy.

This difficulty is caused by a wide variety of symptoms at the onset of SAH. The diagnosis of SAH is relatively easy in patients presenting with disturbance of consciousness or sudden severe headache with typical vomiting, but it is not easy in some patients presenting with clear consciousness and mild headache.[2–5]

The first medical examination for patients with suspected SAH is noncontrast computed tomography (NCCT).[6] It has been reported that the sensitivity of NCCT is very high in patients with suspected SAH,[7–9] and experienced diagnostic radiologists or neurosurgeons can achieve 100% diagnostic accuracy if they diagnose the patients within 6 hours after onset.[8,10–12]

In most medical institutions, an emergency physician or a general practitioner first examines patients with suspected SAH in an emergency outpatient unit.[3,13] Many previous studies have reported that a considerable proportion of SAHs are overlooked at the first visit.[3,4,13–15] Oversight of SAH at the first visit results in a delayed start of the treatment and rerupture of the aneurysm before the start of the treatment, leading to worsening of the disease grade at the start of the treatment and poor clinical outcome.[2,4,13,14,16–18]

Therefore, artificial intelligence (AI) systems that automatically analyze the findings of NCCT and detect the presence and location of intracranial lesions have been developed.[19–25] However, there has been no report that demonstrates actual improvement of diagnostic accuracy using these AI systems.

In the present study, we developed an SAH-specific AI system capable of detecting the presence and location of bleeding in every slice of NCCT and examined whether the system supported the improvement of the detection accuracy of nonspecialists.

Materials and Methods

Data set

Since the present study was a single-center study, all image data were retrospectively collected from Saiseikai Kumamoto Hospital. All the data used were obtained from patients who gave us informed consent in advance, and this study was approved by the Ethics Committee, Saiseikai Kumamoto Hospital (IRB no. 546).

The objects of this study were NCCT images taken in Saiseikai Kumamoto Hospital, and the images obtained from patients with nontraumatic SAH and from healthy subjects were used for algorithm development or experiments on image interpretation.

NCCT images obtained from 757 subjects were used for the development of the algorithm. Two neurosurgery specialists judged the presence or absence of bleeding in each slice and created a mask to the bleeding area. Among 419 cases with SAH, 392 randomly extracted cases were used for deep learning and 27 cases were used for validation. Among 338 healthy cases, 316 cases were used for deep learning and 22 cases were used for validation.

NCCT images obtained from 331 cases were used for the experiment on image interpretation; they randomly included images from 135 cases with SAH and those from 196 healthy cases. The CT images of 419 cases were used for deep learning and 135 cases were used for the interpretation experiment, all of which were finally diagnosed with SAH by clinical symptoms and interpretation by two radiologists. Furthermore, the presence of aneurysms was confirmed by 3D-CTA in all 135 cases used in the interpretation experiment. Two neurosurgery specialists participating in the experiment who specialized in cerebrovascular diseases created correct answers by judging the presence of bleeding two months after the experiment on image interpretation. First, each of two neurosurgery specialists blindly judged the presence of bleeding in each slice of NCCT images based on clinical information and image data. Slices for which the judgment of two specialists agreed were regarded as correctly diagnosed. If the judgments of the two specialists were different, the correct diagnosis was decided by discussion between them.

NCCT scanning

The NCCT data were acquired with an Aquilion 16 or Aquilion CXL scanner (Canon Medical Systems Corporation, Otawara, Tochigi, Japan) using nonhelical scanning. The scan parameters were as follows: tube voltage of 120 kV, tube current of 300 mA, rotation time of 1.0 seconds, reconstructed slice thickness for the upper brain of 8.0 mm (scanning slice thickness of 4.0 mm), reconstructed slice thickness for the lower brain of 4.0 mm (scanning slice thickness of 2.0 mm), slice gap of 0.0 mm, field of view size of 240.0 mm, and pixel matrix of 512×512.

Deep learning algorithm

We utilized the deep neural network architecture for segmentation tasks to develop an AI system for SAH detection. Segmentation is effective for diagnosing SAH because it is important to know not only the presence of bleeding but also the detailed region. Convolutional neural networks (CNNs) are widely used to solve segmentation tasks. Recent work has shown that this technique can be successfully applied to the multi-organ segmentation in 3D CT images.[26] In our method, we developed a CNN that segments the bleeding region by using mask data created by neurosurgery specialists as training data.

Supplementary Fig. 1 shows a schematic drawing of the architecture (All Supplementary files are available Online). For our algorithm, a 3D U-net, which is an expanded version of a 2D U-net, was

used. The 2D U-net was proposed as a method for region extraction of medical images.[27] Since hematoma often spreads into multiple slices in SAH—the target of this study—utilization of the information on the upper and lower slices in addition to the target slice is useful for region segmentation. In this study, we used 2D U-net as a base software and applied 3D convolution to a part of the network. The input was 3D CT image and the output was the probability of bleeding in each voxel. In the preprocessing steps, the Hounsfield unit of the input 3D CT images was clipped to the [−100.0, 200.0] range and then normalized to be in the [0.0, 1.0] interval. After that, input 3D CT images were rescaled to 1.0 mm voxels at only the x and y axes.

For each training iteration, data augmentation was applied. In particular, an affine transformation was applied consisting of a random rotation around the z axis between −3.5 and +3.5 degrees and random scaling in the x and y directions between −6.25% and +6.25%. The horizontal flip was applied randomly to images. For training our network, dice loss[28] based on the dice similarity coefficient (DSC) was used as the loss function. The DSC is defined as follows:

$$DSC = \frac{2 \cdot TP}{2 \cdot TP + FP + FN}$$

where true positive (TP) and false positive (FP) denote the number of voxels correctly and incorrectly segmented as lesions, respectively, and false negative (FN) denotes the number of voxels incorrectly segmented as nonlesions.

Image interpretation by algorithm

The accuracy of our algorithm for detecting SAH was evaluated in each slice of NCCT images. The presence of SAH in each slice was judged by measuring if there were any voxels that had a higher bleeding probability than the threshold.

Experiment on image interpretation by neurosurgery specialists

We comparatively evaluated the detection accuracy of SAH in each slice of NCCT images by the algorithm developed by us and by neurosurgery specialists in this study. Five neurosurgery specialists and our algorithm judged the presence of SAH in each slice of NCCT images obtained from 331 cases. The time to interpret an image was set as the time that was spent for each interpretation in daily clinical practice, and experiment interpretations were carried out individually.

Experiment on image interpretation by nonspecialists

We examined whether our algorithm improved the ability of nonspecialists (general physicians who were not specialized for cerebral diseases) to interpret NCCT images obtained from cases with SAH. First, five nonspecialists judged the presence of SAH on each slice of NCCT images without the help of the algorithm, and then, they finally judged it again after presenting the results of the image extraction by the algorithm. The results of nonspecialists before and after presenting the extraction results by the algorithm were separately collected.

Statistical analysis

TP, true negative (TN), FP, and FN were calculated as the evaluation indices, and sensitivity, accuracy, precision, F-measure (F), and Matthews correlation coefficient (MCC) were also calculated. Then, receiver operating characteristic (ROC) analysis was performed by changing the threshold for the probability of bleeding, the algorithm output, in 10 steps.

To evaluate the degree of answer concordance among neurosurgery specialists, nonspecialists, algorithms, and correct answers, the Cohen's kappa coefficient was calculated for each slice and each case.

Paired t-tests were performed on each of the evaluation indices (sensitivity, specificity, accuracy, F, and MCC) to examine whether the algorithm developed by us contributed to the improvement of the interpreting ability of nonspecialists.

Results

ROC analysis

The AI system developed by us demonstrated the capability of detecting SAH with equal or slightly less accuracy compared to that of neurosurgery specialists in both the patient-based analysis (area under the curve [AUC] = 0.99, operating point: sensitivity = 0.99, specificity = 0.92) (Fig. 1A) and the slice-based analysis (AUC = 0.94, operating point: sensitivity = 0.89, specificity = 0.98) (Fig. 1B). In addition, our AI system was able to diagnose SAH with higher accuracy than three out of five nonspecialists in the patient-based analysis and all of nonspecialists in the slice-based analysis.

Index score

The AI system gave equal or slightly fewer results in all indices evaluated compared to neurosurgery specialists in the patient-based analysis (Table 1). In the analysis, the specificity, accuracy, precision, F-measure, and MCC of the AI system were lower
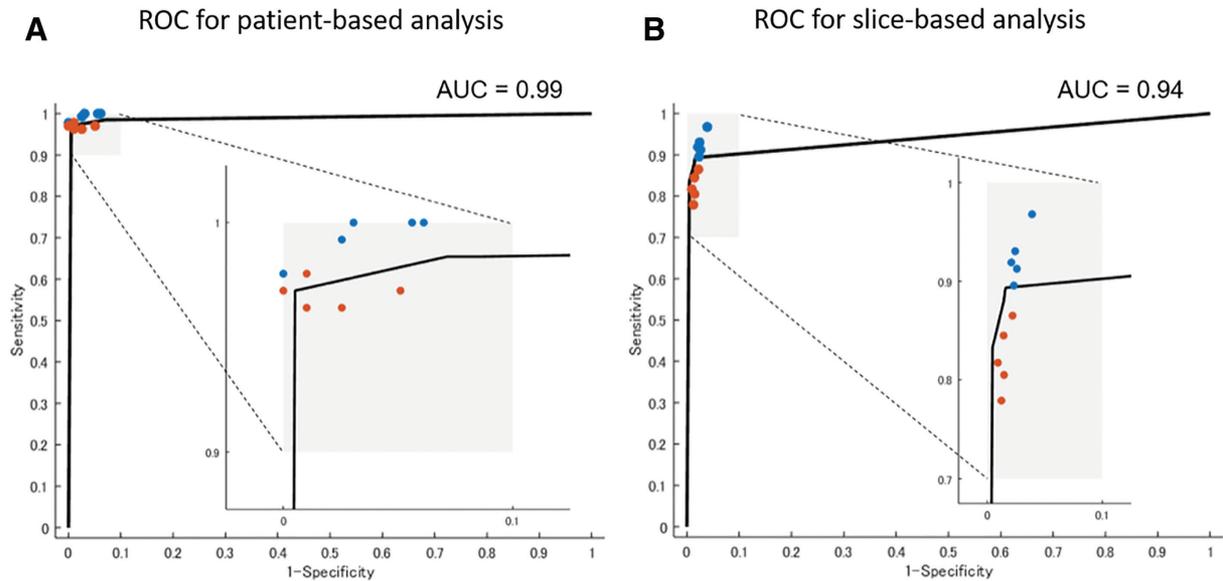
Fig. 1 ROC curve for the performance of AI system (black lines), five neurosurgery specialists, and five nonspecialists (colored circles) to predict 135 SAH and 196 healthy subjects from the CT examinations. The AI achieved AUCs of 0.99 in patient-based analysis (A) and 0.94 in slice-based analysis (B) referenced to the GT (consensus interpretation of two neurosurgery specialists). The AI performance exceeded than three out of five nonspecialists in patient-based analysis and all of the nonspecialists in slice-based analysis. The blue-colored circles and red-colored circles correspond to neurosurgery specialists and nonspecialists, respectively. AI: artificial intelligence, AUCs: areas under the curve, GT: ground truth, ROC: receiver operating characteristic.

Table 1 Index for patient-based analysis of five neurosurgery specialists, five nonspecialists, and AI

	Rater 1	Rater 2	Rater 3	Rater 4	Rater 5	Rater 6	Rater 7	Rater 8	Rater 9	Rater 10	AI
TP	135	135	135	132	134	130	130	132	131	131	133
FP	11	12	6	0	5	5	2	2	0	10	15
FN	0	0	0	3	1	5	5	3	4	4	2
TN	185	184	190	196	191	191	194	194	196	186	181
Sensitivity	1.00	1.00	1.00	0.98	0.99	0.96	0.96	0.98	0.97	0.97	0.99
Specificity	0.94	0.94	0.97	1.00	0.97	0.97	0.99	0.99	1.00	0.95	0.92
Accuracy	0.97	0.96	0.98	0.99	0.98	0.97	0.98	0.98	0.99	0.96	0.95
Precision	0.92	0.92	0.96	1.00	0.96	0.96	0.98	0.99	1.00	0.93	0.90
F	0.96	0.96	0.98	0.99	0.98	0.96	0.97	0.98	0.98	0.95	0.94
MCC	0.93	0.93	0.96	0.98	0.96	0.94	0.96	0.97	0.98	0.91	0.90

Raters 1 to 5 are neurosurgery specialists and Raters 6 to 10 are nonspecialists.
AI: artificial intelligence, F: F-measure, FN: false negative, FP: false positive, MCC: Matthews correlation coefficient, TN: true negative, TP: true positive.

than those of nonspecialists in some cases. A similar tendency was found between the comparison of neurosurgery specialists and nonspecialists. In the slice-based analysis, the sensitivity, specificity, accuracy, F-measure, and MCC of the AI system were equal to or slightly less than those of neurosurgery specialists (Supplementary Table 1). Specificity and precision were superior in the AI system

or nonspecialists compared to neurosurgery specialists in some slices.

Agreement using Cohen's kappa coefficients between specialists, nonspecialists, and AI

The AI system, neurosurgery specialists, and nonspecialists showed an excellent degree of agreement in detecting SAH, and the Cohen's kappa

Table 2 Patient-based agreement using Cohen's kappa coefficient between five neurosurgery specialists, five nonspecialists, AI, and GT

	Rater 2	Rater 3	Rater 4	Rater 5	Rater 6	Rater 7	Rater 8	Rater 9	Rater 10	AI	GT
Rater 1	0.88	0.90	0.91	0.91	0.88	0.89	0.91	0.91	0.86	0.87	0.93
Rater 2		0.89	0.91	0.91	0.90	0.89	0.91	0.90	0.88	0.85	0.93
Rater 3			0.94	0.93	0.90	0.92	0.94	0.94	0.88	0.86	0.96
Rater 4				0.94	0.94	0.95	0.96	0.98	0.91	0.89	0.98
Rater 5					0.94	0.94	0.96	0.94	0.93	0.87	0.96
Rater 6						0.94	0.93	0.96	0.90	0.87	0.94
Rater 7							0.95	0.96	0.89	0.88	0.96
Rater 8								0.96	0.92	0.88	0.97
Rater 9									0.91	0.89	0.97
Rater 10										0.83	0.91
AI											0.90

Raters 1 to 5 are neurosurgery specialists and Raters 6 to 10 are nonspecialists.
AI: artificial intelligence, GT: ground truth.

Table 3 Index for patient-based analysis without AI and with AI of five nonspecialists

	Without AI					With AI				
	Rater 6	Rater 7	Rater 8	Rater 9	Rater 10	Rater 6	Rater 7	Rater 8	Rater 9	Rater 10
TP	130	130	132	131	131	130	132	133	132	132
FP	5	2	2	0	10	5	2	1	0	7
FN	5	5	3	4	4	5	3	2	3	3
TN	191	194	194	196	186	191	194	195	196	189
Sensitivity*	0.96	0.96	0.98	0.97	0.97	0.96	0.98	0.99	0.98	0.98
Specificity	0.97	0.99	0.99	1.00	0.95	0.97	0.99	0.99	1.00	0.96
Accuracy	0.97	0.98	0.98	0.99	0.96	0.97	0.98	0.99	0.99	0.97
Precision	0.96	0.98	0.99	1.00	0.93	0.96	0.99	0.99	1.00	0.95
F**	0.96	0.97	0.98	0.98	0.95	0.96	0.98	0.99	0.99	0.96
MCC	0.94	0.96	0.97	0.98	0.91	0.94	0.97	0.98	0.98	0.94

Raters 6 to 10 are nonspecialists.
AI: artificial intelligence, F: F-measure, FN: false negative, FP: false positive, MCC: Matthews correlation coefficient, TN: true negative, TP: true positive.
*p-value = 0.034, **p-value = 0.049.

coefficients of agreement between the AI system and nonspecialists or between the AI system and neurosurgery specialists were approximately 0.8 in both the patient-based (Table 2) and the slice-based (Supplementary Table 2) analyses.

Results of experiments on image interpretation supported by the AI system

Nonspecialists interpreted NCCT images using the presence or absence of SAH in each slice indicated by the AI system as a reference. Based on the patient-based analysis, the number of FP patients decreased by one to three in the interpretation of two nonspecialists and that of FN patients decreased by one to two in the interpretation of three nonspecialists with the support of the AI system (Table 3). Without the support of the AI system, the accuracy of judgment by three out of five nonspecialists was lower than that of the AI system (Fig. 2A); with the support of the AI system, it was higher in four out of five nonspecialists than that of the AI system (Fig. 2B). Paired t-tests on each of the evaluation indices obtained before and after AI support revealed significant improvements in sensitivity (p = 0.034)

A ROC for patient-based analysis
Non-specialist without CAD

B ROC for patient-based analysis
Non-specialist with CAD

AUC = 0.99

AUC = 0.99
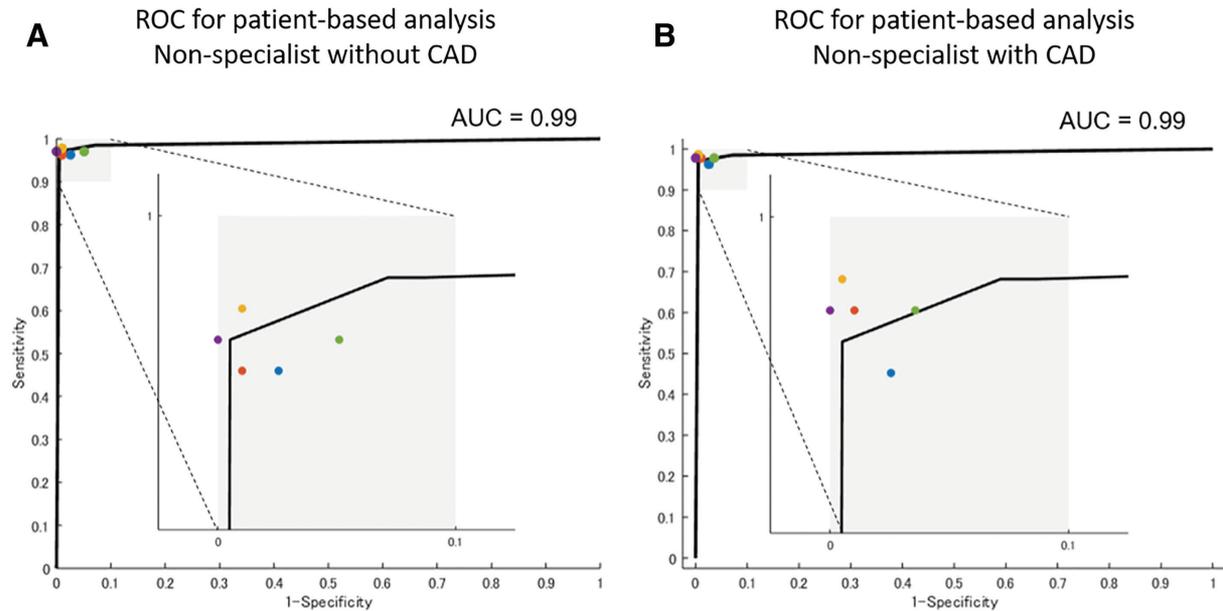
Fig. 2 ROC curve for the performance of AI system (black lines) and five nonspecialists (colored circles) to predict 135 SAH and 196 healthy subjects from the CT examinations in patient-based analysis. The performance of prediction without AI (A) and with AI (B) of five nonspecialists is shown. The diagnostic accuracy of four out of five nonspecialists was improved after referring to the results of AI. Blue, red, yellow, purple, and green circles correspond to attending Rater 6, Rater 7, Rater 8, Rater 9, and Rater 10, respectively. AI: artificial intelligence, ROC: receiver operating characteristic, SAH: subarachnoid hemorrhage.

and F-measure (p = 0.049) with the support of the AI system in the patient-based analysis (Table 3). In the slice-based analysis, the precision of the judgment was significantly improved (p = 0.049) with the support of the AI system (Supplementary Table 3).

A representative case for which support by the AI system was helpful for the judgment of nonspecialists is shown in Supplementary Fig. 2. The bleeding areas judged by the AI system (blue areas) on the image from Supplementary Fig. 2A are shown in Supplementary Fig. 2B. The AI system presented the presence and location of the bleeding area as rectangles surrounded by blue lines to study subjects (nonspecialists) (Supplementary Fig. 2C).

Optimization of algorithm

The Dice score on the validation set was 0.513 in this study. In addition, the sensitivity and specificity in each case were 96.3% and 90.9%, respectively, in the validation.

Discussion

The AI system we created was able to diagnose SAH with high accuracy approaching to that of neurosurgery specialists. This study also demonstrated that the AI system significantly improved the

diagnostic accuracy of nonspecialists in neurosurgery (general physicians) by presenting the area where the AI system itself judged the bleeding area as a rectangular region on each slice. However, even considering how accurately an AI system may diagnose SAH, by law, only physicians can make clinical judgments for the treatment of SAH in the current social structure. Therefore, we think that it is important whether an AI system can support physicians' judgment and improve their diagnostic ability. In a previous study, an AI system made a superior diagnosis compared with that of diagnostic radiologists on intracranial lesions,[21] but the authors did not reveal whether the AI system improved the diagnostic ability of nonspecialists. Our present report was the first to demonstrate the improvement of the diagnostic ability of nonspecialists by the AI system developed by us. It is considered that the AI system enables accurate diagnosis of SAH even in frontline hospitals where neurosurgery specialists or diagnostic radiologists are not working 24 hours a day.

The most important finding in the present study may be that obtained by the patient-based analysis on the experiment on the interpretation of NCCT images by nonspecialists. In image interpretation by four out of five nonspecialists, the number of FN patients significantly decreased by one to two.

In addition, in image interpretation by two out of five nonspecialists, the number of FP patients decreased by one to three, although the change was not significant. The mean sensitivity of nonspecialist judgment was also significantly improved. These results proved that the number of patients with SAH whose SAH was overlooked was significantly decreased by using the AI system developed by us, indicating that such patients may be treated in their good condition before rerupture of aneurysms and may have favorable outcomes. Therefore, the development of AI systems may contribute to the improvement of the treatment outcome of SAH as a whole. In addition, it enables us to avoid the risk of potential lawsuits due to misdiagnosis. Furthermore, FP patients will be treated and examined unnecessarily for SAH. Therefore, the reduction in the number of FP patients also reduces the possibility of examinations that have the risk of exposure to X-ray and allergy caused by contrast-CT.

In the experiment on image interpretation, statistical analyses revealed that the specificity and precision of nonspecialist judgment were higher than those of neurosurgery specialists or the AI system in some cases. In the diagnosis of SAH, however, it is most important that the number of FN patients is small in the patient-based analysis, meaning that the number of oversights is small. Among five neurosurgery specialists, three of them with long experience had no FN patients, and the sensitivity of their judgment was 1.00; the other two neurosurgery specialists had one and two FN patients. Conversely, all five nonspecialists had three to five FN patients. In that sense, it should not be overlooked that there were FN cases in the interpretation by AI. The accuracy of AI is not yet considered to be comparable to that of neurosurgeon specialists.

There are several reasons for the uncertainty of diagnosis for SAH. The first is that symptoms are too mild to proceed to further examinations, including NCCT, in some patients.[5,13] It has been reported that there are certain proportions of patients who complain of mild or no headache. However, it is possible to choose appropriate examinations and to reach a correct diagnosis by paying attention to various sudden symptoms without the presence of a headache[13,29]; these procedures are not difficult for experienced neurosurgeons. To achieve this goal with an AI system, it will be necessary to create an algorithm that can suggest the necessity of NCCT examination from the results of a medical history taken that includes formulaic questions and to develop an AI system that participates in the medical history recording process.

The second is situations where an examiner is not able to interpret the finding of NCCT images as SAH.[5,13,30] There are two situations: the first is a case where an examiner notices an abnormal finding but misdiagnoses it as another intracranial lesion, and the second is a case where an examiner does not notice an abnormality and judges it as normal.[31,32] The first case may lead to medical treatments by neurosurgeons because the examiner has still recognized an abnormality. However, the second case is serious because it means delay of treatment for SAH, which should be treated immediately. Therefore, it indicates the necessity of an AI system that prevents oversight of SAH by nonspecialists.

No diagnostic radiologists participated in the creation of the AI system or the NCCT image interpretation experiment in the present study. Diagnostic radiologists have engaged in the creation of correct answers in many previous studies on diagnostic imaging.[10,21] In Japan, however, neurosurgeons or neurologists have made a final diagnosis for diagnostic imaging of intracranial lesions in most cases[13] because the number of diagnostic radiologists who can respond 24 hours a day is limited at frontline emergency hospitals in Japan. Therefore, five experienced neurosurgeons specializing in cerebrovascular diseases participated in the experiment on image interpretation and the creation of correct answers. Since neurosurgery specialists who created correct answers among them were certified as technical advisory specialists by the Japanese Society of Surgery for Cerebral Stroke, their diagnostic ability is considered to be reliable.

The AI system created by us in the present study was specialized for the diagnosis of SAH. Intracranial diseases include cerebrovascular diseases such as cerebral infarction, cerebral hemorrhage, and traumatic lesions such as brain contusion, traumatic SAH, acute subdural hematoma, acute epidural hematoma, and chronic subdural hematoma, all of which are serious diseases. Therefore, the existence of an AI system that can diagnose these diseases may be extremely useful. AI systems capable of diagnosing various intracranial hemorrhages have been reported thus far, but the systems developed at the beginning were not appropriate for the detection of SAH.[33–37] An AI system that can delineate bleeding areas was reported recently,[21] but another study stated that the diagnosis of SAH was the most difficult.[20]

Among the diseases listed above, SAH is considered to be the most critical disease that brings serious disadvantages to patients because it is most likely overlooked on NCCT images by a

nonspecialist, which causes a delay in treatment. Therefore, the development of AI systems specialized for the diagnosis of SAH is considered to be significant.

Conclusion

The use of our AI support system for image interpretation will allow us to improve the accuracy of SAH diagnosis when a neurosurgery specialist is not available or when a diagnostic radiologist cannot interpret images in a timely manner in the emergency outpatient unit and to perform timely therapeutic interventions. Therefore, we expect that our AI system contributes to the improvement of the treatment outcome of SAH.

Limitations

The present study is a single-center retrospective study; thus, multicenter prospective studies using images taken by NCCT equipment may be required. Images with difficult-to-read motion artifacts or strong metal artifacts are excluded from learning and interpretation experiments.

Acknowledgments

Conflicts of Interest Disclosure

References

1) Kassell NF, Torner JC, Haley EC, Jane JA, Adams HP, Kongable GL: The International Cooperative Study on the timing of aneurysm surgery. Part 1: Overall management results. *J Neurosurg* 73: 18–36, 1990
2) Mayer PL, Awad IA, Todor R, et al.: Misdiagnosis of symptomatic cerebral aneurysm. Prevalence and correlation with outcome at four institutions. *Stroke* 27: 1558–1563, 1996
3) Vermeulen MJ, Schull MJ: Missed diagnosis of subarachnoid hemorrhage in the emergency department. *Stroke* 38: 1216–1221, 2007
4) Kowalski RG, Claassen J, Kreiter KT, et al.: Initial misdiagnosis and outcome after subarachnoid hemorrhage. *JAMA* 291: 866–869, 2004
5) Oh SY, Lim YC, Shim YS, et al.: Initial misdiagnosis of aneurysmal subarachnoid hemorrhage: associating factors and its prognosis. *Acta Neurochir (Wien)* 160: 1105–1113, 2018
6) Rao VM, Levin DC, Parker L, Frangos AJ, Sunshine JH: Trends in utilization rates of the various imaging modalities in emergency departments: nationwide Medicare data from 2000 to 2008. *J Am Coll Radiol* 8: 706–709, 2011
7) Boesiger BM, Shiber JR: Subarachnoid hemorrhage diagnosis by computed tomography and lumbar puncture: are fifth generation CT scanners better at identifying subarachnoid hemorrhage? *J Emerg Med* 29: 23–27, 2005
8) Cortnum S, Sørensen P, Jørgensen J: Determining the sensitivity of computed tomography scanning in early detection of subarachnoid hemorrhage. *Neurosurgery* 66: 900–902, discussion 903, 2010
9) Westerlaan HE, van Dijk JM, van Dijk MJ, et al.: Intracranial aneurysms in patients with subarachnoid hemorrhage: CT angiography as a primary examination tool for diagnosis–systematic review and meta-analysis. *Radiology* 258: 134–145, 2011
10) Perry JJ, Stiell IG, Sivilotti ML, et al.: Sensitivity of computed tomography performed within six hours of onset of headache for diagnosis of subarachnoid haemorrhage: prospective cohort study. *BMJ* 343: d4277, 2011
11) Backes D, Rinkel GJ, Kemperman H, Linn FH, Vergouwen MD: Time-dependent test characteristics of head computed tomography in patients suspected of nontraumatic subarachnoid hemorrhage. *Stroke* 43: 2115–2119, 2012
12) Blok KM, Rinkel GJ, Majoie CB, et al.: CT within 6 hours of headache onset to rule out subarachnoid hemorrhage in nonacademic hospitals. *Neurology* 84: 1927–1932, 2015
13) Takagi Y, Hadeishi H, Mineharu Y, et al.: Initially missed or delayed diagnosis of subarachnoid hemorrhage: a nationwide survey of contributing factors and outcomes in Japan. *J Stroke Cerebrovasc Dis* 27: 871–877, 2018
14) Ois A, Vivas E, Figueras-Aguirre G, et al.: Misdiagnosis worsens prognosis in subarachnoid hemorrhage with good Hunt and Hess score. *Stroke* 50: 3072–3076, 2019
15) Yamada T, Natori Y: Evaluation of misdiagnosed cases of subarachnoid hemorrhage and causal factors for misdiagnosis. *J Stroke Cerebrovasc Dis* 22: 430–436, 2013
16) Edlow JA, Caplan LR: Avoiding pitfalls in the diagnosis of subarachnoid hemorrhage. *N Engl J Med* 342: 29–36, 2000
17) Bederson JB, Connolly ES, Batjer HH, et al.: Guidelines for the management of aneurysmal subarachnoid

hemorrhage: a statement for healthcare professionals from a special writing group of the *Stroke* Council. *American Heart Association. Stroke* 40: 994–1025, 2009

18) van Lieshout JH, Bruland I, Fischer I, et al.: Increased mortality of patients with aneurysmatic subarachnoid hemorrhage caused by prolonged transport time to a high-volume neurosurgical unit. *Am J Emerg Med* 35: 45–50, 2017

19) Ginat DT: Analysis of head CT scans flagged by deep learning software for acute intracranial hemorrhage. *Neuroradiology* 62: 335–340, 2020

20) Ye H, Gao F, Yin Y, et al.: Precise diagnosis of intracranial hemorrhage and subtypes using a three-dimensional joint convolutional and recurrent neural network. *Eur Radiol* 29: 6191–6201, 2019

21) Kuo W, Häne C, Mukherjee P, Malik J, Yuh EL: Expert-level detection of acute intracranial hemorrhage on head computed tomography using deep learning. *Proc Natl Acad Sci U S A* 116: 22737–22745, 2019

22) Prevedello LM, Erdal BS, Ryu JL, et al.: Automated critical test findings identification and online notification system using artificial intelligence in imaging. *Radiology* 285: 923–931, 2017

23) Chang PD, Kuoy E, Grinband J, et al.: Hybrid 3D/2D convolutional neural network for hemorrhage evaluation on head CT. *AJNR Am J Neuroradiol* 39: 1609–1616, 2018

24) Chilamkurthy S, Ghosh R, Tanamala S, et al.: Deep learning algorithms for detection of critical findings in head CT scans: a retrospective study. *Lancet* 392: 2388–2396, 2018

25) Titano JJ, Badgeley M, Schefflein J, et al.: Automated deep-neural-network surveillance of cranial images for acute neurologic events. *Nat Med* 24: 1337–1341, 2018

26) Roth HR, Oda H, Zhou X, et al.: An application of cascaded 3d fully convolutional networks for medical image segmentation. *Comp Med Imag Graph* 66: 90–99, 2018

27) Ronneberger OFP, Brox T: U-net: Convolutional networks for biomedical image segmentation. In: Nassir Navab JH, William Wells WM, Frangi A, eds. 18th International Conference on Medical Image Computing and Computer-Assisted Intervention. MICCAI. Munich, Germany: Springer International Publishing, 2015, pp 234–241

28) Milletari F, Navab N, Ahmadi SA: V-Net: fully convolutional neural networks for volumetric medical image segmentation. *Fourth International Conference on 3D Vision,* 2016, pp 565–571

29) Malatt C, Zawaideh M, Chao C, Hesselink JR, Lee RR, Chen JY: Head computed tomography in the emergency department: a collection of easily missed findings that are life-threatening or life-changing. *J Emerg Med* 47: 646–659, 2014

30) Wysoki MG, Nassar CJ, Koenigsberg RA, Novelline RA, Faro SH, Faerber EN: Head trauma: CT scan interpretation by radiology residents versus staff radiologists. *Radiology* 208: 125–128, 1998

31) Renfrew DL, Franken EA, Berbaum KS, Weigelt FH, Abu-Yousef MM: Error in radiology: classification and lessons in 182 cases presented at a problem case conference. *Radiology* 183: 145–150, 1992

32) Bahrami S, Yim CM: Quality initiatives: blind spots at brain imaging. *Radiographics* 29: 1877–1896, 2009

33) Baldy RE, Brindley GS, Ewusi-Mensah I, et al.: A fully-automated computer-assisted method of CT brain scan analysis for the measurement of cerebrospinal fluid spaces and brain absorption density. *Neuroradiology* 28: 109–117, 1986

34) Loncaric S, Dhawan AP, Broderick J, Brott T: 3-D image analysis of intra-cerebral brain hemorrhage from digitized CT films. *Comput Methods Programs Biomed* 46: 207–216, 1995

35) Chan T: Computer aided detection of small acute intracranial hemorrhage on computer tomography of brain. *Comput Med Imaging Graph* 31: 285–298, 2007

36) Yuh EL, Gean AD, Manley GT, Callen AL, Wintermark M: Computer-aided assessment of head computed tomography (CT) studies in patients with suspected traumatic brain injury. *J Neurotrauma* 25: 1163–1172, 2008

37) Bardera A, Boada I, Feixas M, et al.: Semi-automated method for brain hematoma and edema quantification using computed tomography. *Comput Med Imaging Graph* 33: 304–311, 2009

Corresponding author: Toru Nishi, MD, PhD
Department of Neurosurgery, Sakurajyuji Yatsushiro Rehabilitation Hospital, 2-4-31 Honmachi, Yatsushiro, Kumamoto 866-0861, Japan.
e-mail: tnishi2@me.com