Graefe's Archive for Clinical and Experimental Ophthalmology
https://doi.org/10.1007/s00417-024-06501-1

ONCOLOGY

Artificial intelligence-based differential diagnosis of orbital MALT lymphoma and IgG4 related ophthalmic disease using hematoxylin–eosin images

Mizuki Tagami[1,3] · Mizuho Nishio[2] · Atsuko Yoshikawa[3] · Norihiko Misawa[1] · Atsushi Sakai[1] · Yusuke Haruna[1] · Mami Tomita[1] · Atsushi Azumi[3] · Shigeru Honda[1]

Received: 9 November 2023 / Revised: 12 April 2024 / Accepted: 28 April 2024

Abstract

Purpose To investigate the possibility of distinguishing between IgG4-related ophthalmic disease (IgG4-ROD) and orbital MALT lymphoma using artificial intelligence (AI) and hematoxylin–eosin (HE) images.

Methods After identifying a total of 127 patients from whom we were able to procure tissue blocks with IgG4-ROD and orbital MALT lymphoma, we performed histological and molecular genetic analyses, such as gene rearrangement. Subsequently, pathological HE images were collected from these patients followed by the cutting out of 10 different image patches from the HE image of each patient. A total of 970 image patches from the 97 patients were used to construct nine different models of deep learning, and the 300 image patches from the remaining 30 patients were used to evaluate the diagnostic performance of the models. Area under the curve (AUC) and accuracy (ACC) were used for the performance evaluation of the deep learning models. In addition, four ophthalmologists performed the binary classification between IgG4-ROD and orbital MALT lymphoma.

Results EVA, which is a vision-centric foundation model to explore the limits of visual representation, was the best deep learning model among the nine models. The results of EVA were ACC = 73.3% and AUC = 0.807. The ACC of the four ophthalmologists ranged from 40 to 60%.

Conclusions It was possible to construct an AI software based on deep learning that was able to distinguish between IgG4-ROD and orbital MALT. This AI model may be useful as an initial screening tool to direct further ancillary investigations.

Key messages

What is known:

- IgG4-related ophthalmic disease (IgG4-ROD) and orbital MALT lymphoma are major histological in orbital tumor.
- In usually, differential diagnosis of two kinds tumors for specialized imuunohistochemistory including IgG4 staining in Professional Facilities.
- It was difficult to differentiate by HE staining, and there are no reports of identification using artificial intelligence.

What is new:

- In this study, To investigate the possibility of distinguishing between IgG4-related-ROD) and orbital MALT lymphoma using artificial intelligence (AI) and hematoxylin-eosin (HE) images. ophthalmic disease (IgG4-ROD) and orbital MALT lymphoma using artificial intelligence (AI) and hematoxylin-eosin (HE) images.
- Our best deep learning models showed the results of ACC = 73.3% and AUC = 0.807.
- It was possible to construct an AI software based on deep learning that was able to distinguish between orbital IgG4-ROD and orbital MALT.

Keywords Artificial intelligence (AI) · IgG4-related ophthalmic disease (IgG4-ROD) · Orbital MALT lymphoma · Hematoxylin–eosin (HE) · Deep learning

Extended author information available on the last page of the article

Abbreviations

AI Artificial intelligence
HE Hematoxylin and eosin
MALT (Extranodal marginal zone lymphoma of mucosa-associated lymphoid tissue type) Mucosa Associated Lymphoid Tissue

Introduction

Orbital tumor has been the most commonly found primary MALT lymphoma of the ocular adnexa and IgG4-related ophthalmic disease (IgG4-ROD), previously referred to as the lymphoproliferative lesion, throughout the world and in Japan [1–6]. Previous reports have also reported that more than half of IgG4-ROD occurs in the orbit other than in the lacrimal gland, with clinical and hematologic findings as well as a pathological diagnosis reported to be necessary for a definitive diagnosis [2]. MALT lymphoma and IgG4-ROD are almost indistinguishable with regard to the macroscopic histopathological findings, and are diagnosed by hematoxylin and eosin (HE) imaging and immunostaining, which includes IgG4 staining [3, 4]. Furthermore, gene rearrangement has been reported to be useful for diagnosing lymphoma not only for distinguishing lymphoma with lymphoproliferative lesion, but in addition, it has been especially utilized in performing a differential diagnosis for IgG4-ROD [5].

When using this procedure, the differentiation of the two groups requires examination by various methods. However, for the RNA-sequence data, etc., there appears to be a molecular biological difference in the gene expression between the two groups.

In addition, Previous study revealed that he characteristic differences in metabolomic profiles between IgG4-ROD and orbital MALT lymphoma [6].

Therefore, we hypothesized that differences should be able to be found even in simple HE pathological images that have not been previously discussed.

In previous our article, we had investigated and reported that two different origins ocular MALT lymphomas could be distinguished, using machine learning methods focus to morphological changes and differences of HE slide [7]. This conclusion was suggested that artificial intelligence (AI) diagnostic imaging may be useful for the morphological differentiation of HE.

In the present study, we rigorously performed an evaluation that examined the difference between the two groups (orbital MALT; IgG4-ROD), which could be separated by simple HE when using the current methodology.

In contrast, in conjunction with the improvement of AI methodologies, computer-aided techniques for determining a morphological diagnosis have steadily advanced [8–12]. In Japan, previous studies have reported the usefulness of deep learning model for the differentiation of lymphoma [13].

The present study investigated whether differences in hematoxylin and eosin (HE)-stained pathological findings could be used for differentiation of ocular diseases when utilizing AI. Deep learning models were applied to determine if these could be developed into an AI software that could be used to distinguish between orbital lymphomas and IgG4-ROD.

Therefore, although the current technology is still somewhat primitive, we thought that it might be possible to use HE images to examine for potential ocular diseases.

Based on our previous findings, we attempted to describe the potential use of simple HE images, the creation of the latest the AI software including deep learning models, along with clarifying the possible diagnostic capabilities that are associated with the hidden morphological HE features.

Materials and methods

Selection of cases and collation of clinicopathological data

This retrospective observational case study was conducted with approval from the Institutional Review Board (Osaka Metropolitan University 2022–064). The described research adhered to the tenets of the Declaration of Helsinki. The right to opt out was guaranteed for patients who had already stopped visiting the clinic. We identified 127 patients between 2008 and 2022 and from whom tissue blocks with orbital MALT lymphomas and IgG4-ROD were able to be procured. Diagnoses of orbital MALT lymphoma were based on clinical characteristics, radiographic findings (computed tomography and magnetic resonance imaging), results from histological and flow cytometric studies, and molecular genetic analyses such as immunoglobulin JH gene rearrangement with Southern blot analysis.

Histopathologic examination of tumor specimens included staining with hematoxylin and eosin (HE) and immunohistochemical analyses. Currently, the following flowcytometry panel was performed for B-cell lymphoma diagnosis including κ and λ light chains. All lymphomas for diagnosis were classified by the 5th edition of the 2017 World Health Organization classification.

Next, all IgG4-ROD cases underwent histopathological analysis, which included both the comprehensive diagnostic criteria for 'definite' or 'probable' IgG4-related disease (IgG4-RD) that was published by Umehara et al. and the diagnostic criteria for IgG4-RD that was published by

Deshpande et al. ($n = 53$) [4, 14]. HE features associated with IgG4-RD were referenced based on previous reports and which included three major histopathological features; i) dense lymphoplasmacytic infiltrate, ii) fibrosis, arranged at least focally in a storiform pattern, and iii) obliterative phlebitis [3, 15].

Image preparation

Histopathological images were prepared using the following procedure (Fig. 1). First, The central areas of the lymphoma and IGG4-ROD were captured at low magnification (× 2 or × 4) using BX53 and DP74 microscopes (Olympus, Tokyo, Japan). Second, the pathological findings of the × 20 magnification images were annotated by manufactory procedure. The total number of annotations was 1270 images (orbital MALT lymphoma, 740; IgG4-ROD, 530). Image patches were captured. In addition, from the periphery of each annotation at × 20 magnification, image patches of 2048 × 2048 pixels were randomly cropped in the tumor areas. If the average RGB value of all pixels were equal and over 200 value, as the image patch was considered to either be too white or to show too little of the specimen. During the testing phase, smaller image patches were sequentially cropped.

Outline of deep learning model

In the present study, the deep-learning-based software was developed and evaluated for the automatic classification of HE image patches between the orbital MALT lymphoma and IgG4-ROD. For this purpose, a total of 1270 image patches from 127 patients were used. An outline of our software is shown in Fig. 2A. The inputs and outputs for our software were the HE image patches and the ground truth label, respectively. By inputting the HE image patches to our software, our software outputted the probability of orbital MALT lymphoma as the results of the binary classification between the orbital MALT lymphoma and IgG4-ROD provided. The 1270 HE image patches were divided into 970 patches for the model development and 300 patches for the performance evaluation, which were based on the patient. Figure 2B illustrates the dataset splitting. For the software development, Python(version3.7: https://www.python.org/), PyTorch (version 1.12.1: https://github.com/huggingface/pytorch-image-models), and PyTorch Image Models (timm packages) (version, 0.8.3.dev0: https://github.com/huggingface/pytorch-image-models) were used. In addition, train.py and inference.py of the timm package were used for developing and evaluating our software, respectively. The nine deep learning models were constructed and

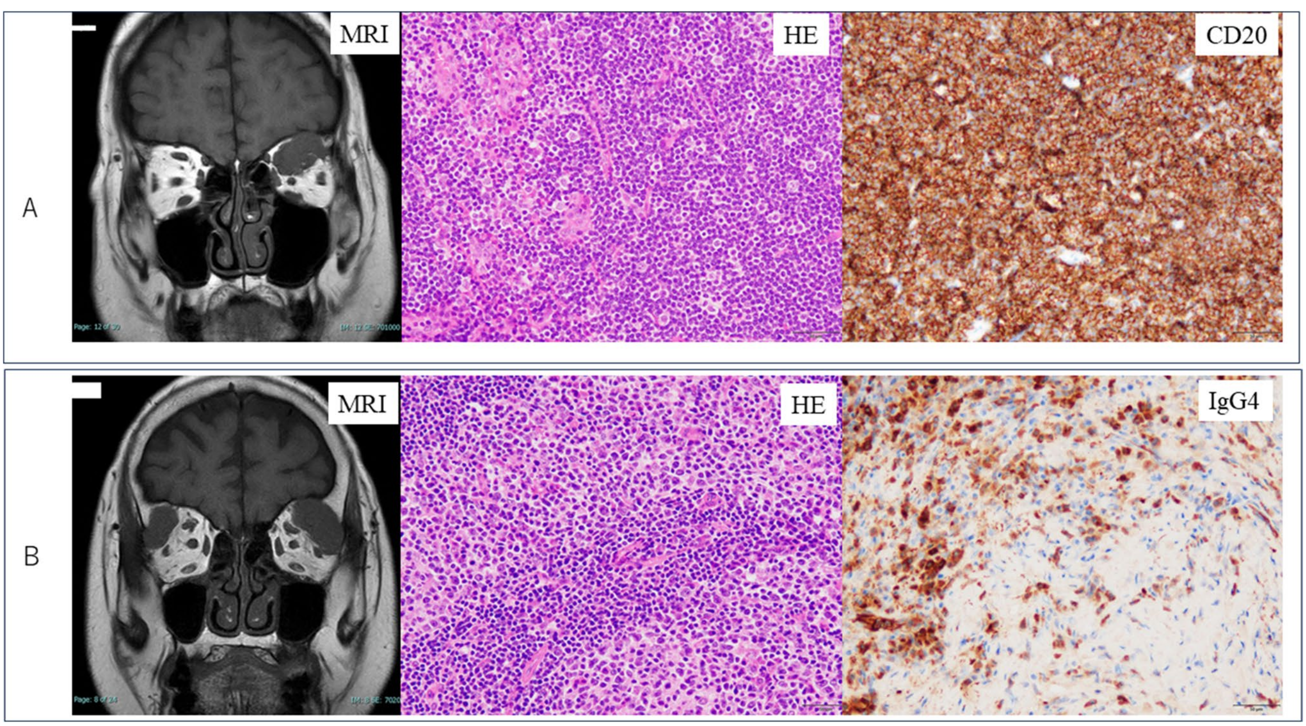

Fig. 1 Representative radiological and histological images of orbital MALT lymphoma and IgG4-related ophthalmic disease **A**: orbital MALT lymphoma, **B**: orbital IgG4-ROD, MRI: magnetic resonance imaging, HE: hematoxylin–eosin, CD20 and IgG4 were immunohistochemical stainings

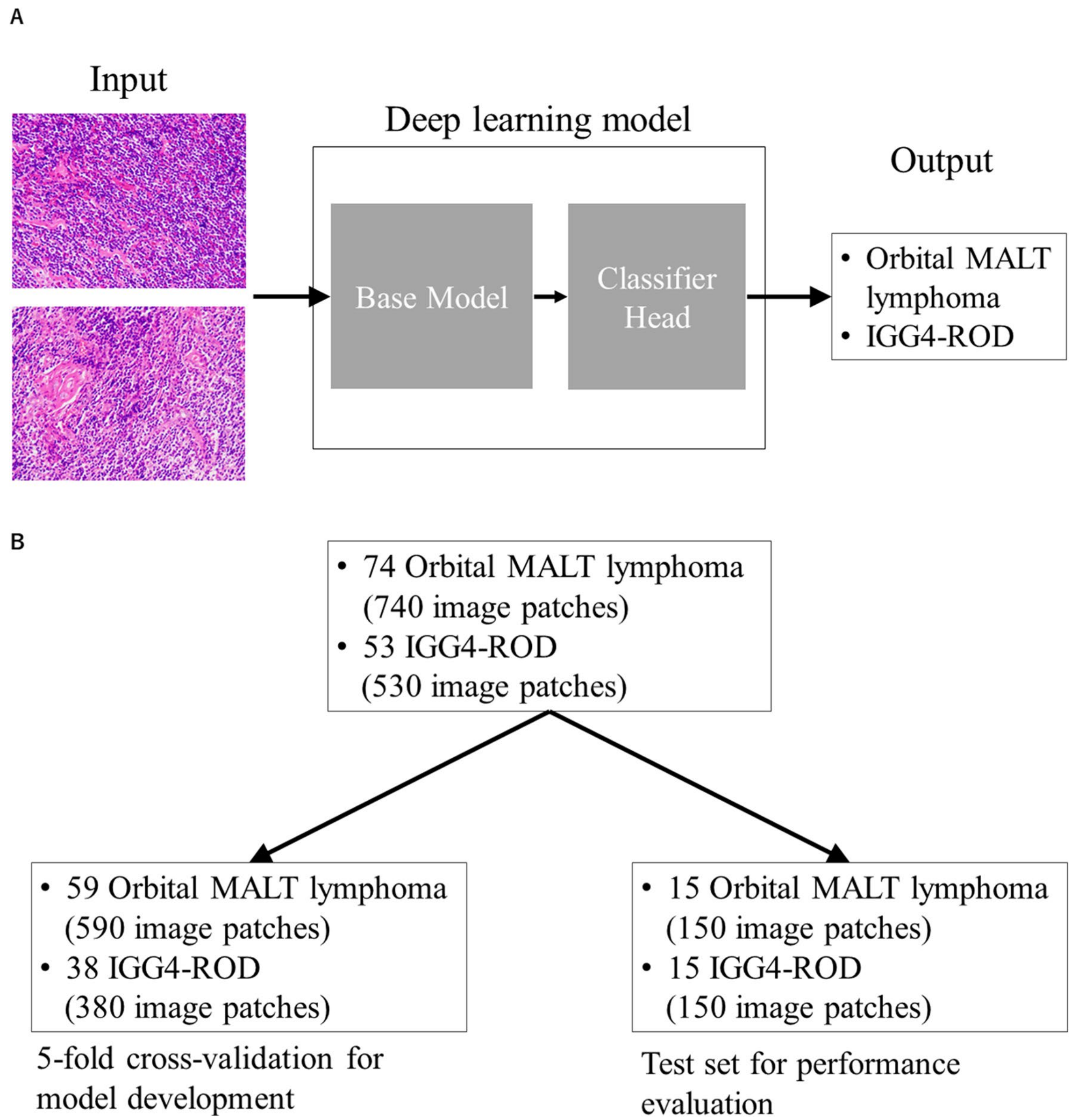

Fig. 2 Schema of deep learning model

evaluated with a workstation that utilized the 12th generation Intel® Core™ i7-12700F, 32 GB main memory, and NVIDIA® GeForce RTX™ 4090.

Details of deep learning model

To develop our deep learning software, pretrained models and transfer learning were employed. As shown in Fig. 2A, our software consisted of classifier heads and base models, with the former specific to the present study, while the latter was used to extract general image features [8]. Using the timm package, nine different pretrained models were used as base model: EVA, which is a vision-centric foundation model used to explore the limits of visual representation, (eva_giant_patch14_336 and eva_giant_patch14_224), [9] Vision Transformer (vit_base_patch16_224 and vit_large_patch14_clip_336), [10] EfficientNet (tf_efficientnet_b3 and tf_efficientnet_b5), [11] Densenet121 (densenet121), [12] Resnet50 (resnet50), [16] and VGG16 (vgg16) [17]. For the development using the pretrained models, patient-based fivefold cross-validation was conducted (Fig. 2C). This involved using the 970 image patches from 97 patients. In the cross-validation, the entire models were trained with the transfer learning, and the pretrained models were fine-tuned. For each fold of the fivefold cross-validation, the trained model with the best

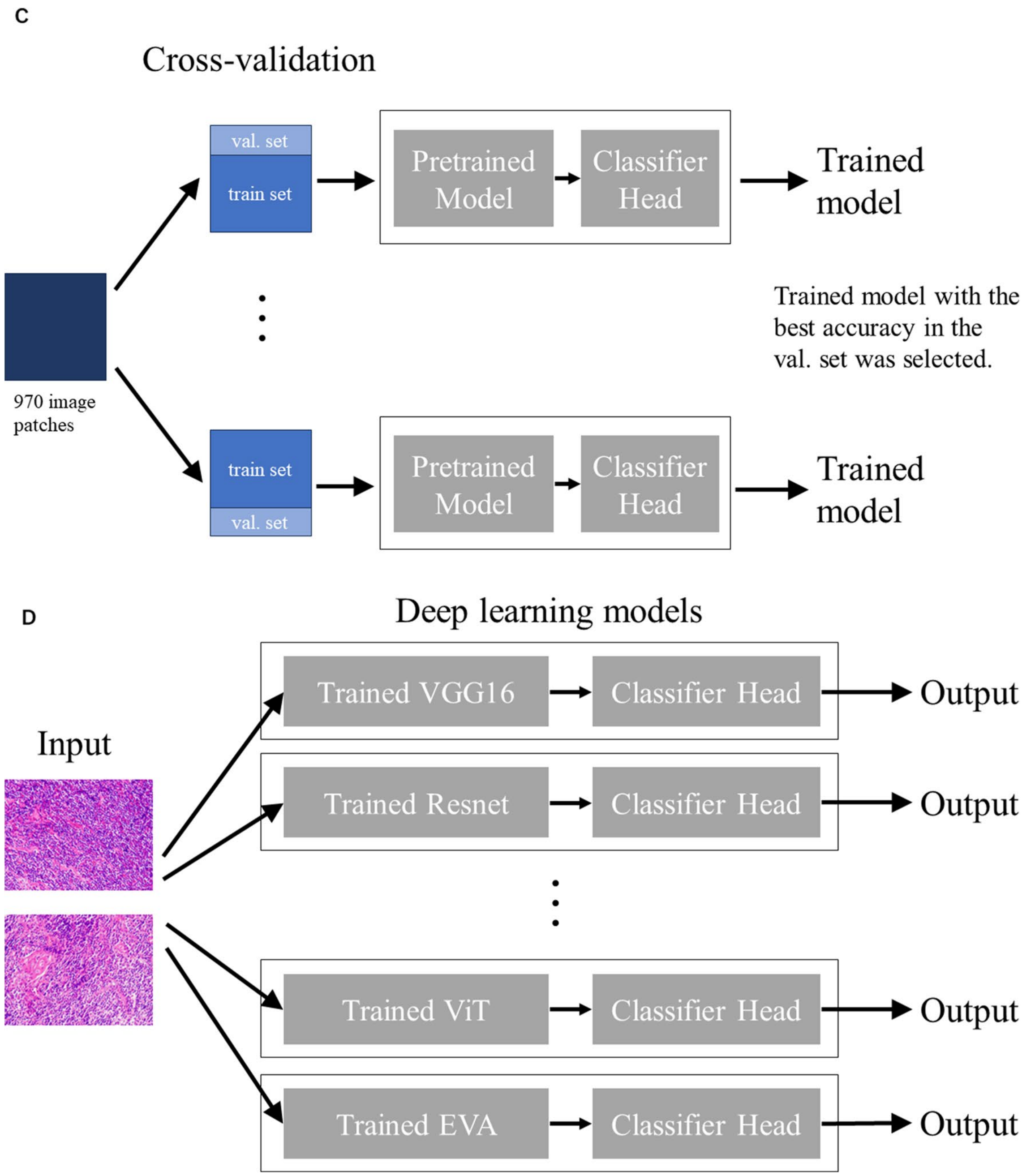

Fig. 2 (continued)

accuracy in the validation set was selected for the evaluating performance on the test set. In the training models, the following hyperparameters were used: number of epochs = 20; batch size = 1; optimizer = stochastic gradient descent; scheduler of learning rate = cosine annealing; and learning rate = 0.0001. Detailed information regarding the deep learning models and their training is available in the Supplementary Material.

Performance evaluation

For performance evaluation, we utilized 300 image patches from the 30 patients (15 orbital MALT lymphoma and 15 IgG4-ROD patients) as the test set. Using the nine different models, we obtained prediction results for the test set. For each of the nine models, an ensemble of 5 trained models obtained through the fivefold cross-validation was utilized to predict the probability of orbital MALT lymphoma from the image patches. A schematic illustration of this process is shown in Fig. 2D. For explainable

AI, saliency maps of the deep learning model were generated for the test set using the timm-vis package (https://github.com/novice03/timm-vis). In addition, the diagnostic performance of four ophthalmologists (Two ophthalmologists hold M.D.s and PH.D.s and have completed specialized residencies in ocular oncology. The remaining two are senior residents with M.D. in ophthalmology). was compared with that of the models for the test set. The four ophthalmologists evaluated and discriminated between the IgG4-ROD or orbital MALT lymphoma utilizing only the HE images and scored the results (0 or 1), with the accuracy rate quantified as the accuracy.

Statistical analyses

Clinical and histopathological characteristics were summarized and assessed using a *t*-test and chi-squared test in Table 1. Statistical analyses were performed using SPSS Statistics software (version 22; IBM Japan, Tokyo, Japan). Values of $P < 0.05$ were considered significant.

For evaluation of our deep learning models, the area under the curve (AUC) of the receiver operating characteristics analysis was calculated between the orbital MALT lymphoma and IgG4-ROD. In addition, sensitivity, specificity, and accuracy were calculated for all of the models. These metrics were also calculated for the results determined by the ophthalmologists. The Youden index was used to determine the optimal threshold in calculating the sensitivity, specificity, and accuracy. To calculate these metrics, Python (version 3.10) and scikit-learn package (version 0.19.3) were used for calculating these metrics.

Results

Clinical findings

Table 1 summarizes the clinical findings for the cohort. All of the 127 patients (100%) were East Asian. The patients included 55 men and 72 women, with a mean age at presentation of 66 ± 14 years. The age for the orbital MALT patients was 60 ± 14 years, while it was 70 ± 13 years for the IgG4-ROD patients.

Next, all IgG4-ROD cases underwent histopathological analysis, which included both the comprehensive diagnostic criteria for 'definite' or 'probable' IgG4-related disease (IgG4-RD) that was published by Umehara et al. and the diagnostic criteria for IgG4-RD that was published by Deshpande et al. ($n = 53$) [4, 15]. Hematological serum IgG4 examination demonstrated there was an average of 528 mg/dL ($n = 49$, min: 101 mg/dL, max: 2630 mg/dL).

In addition, for each of the different diagnoses, all of the cases were confirmed to have a negative JH gene rearrangement. According to Southern blot analysis, all lymphoma cases were positive for the JH gene rearrangement ($n = 74$) and all IgG4-ROD cases were negative for the JH gene rearrangement ($n = 53$) (Table 1). Atypical orbital lymphoma and IgG4-ROD with radiological and histological findings were observed (Fig. 1).

Results of deep learning models

Table 2 shows the results of the performance evaluation for the nine different deep learning models. Table 2 includes sensitivity, specificity, accuracy, and AUC, which all pertain to the binary classification that was performed between the IgG4-ROD and orbital MALT lymphoma. This classification was carried out using the test set that consisted of 300 image patches, and which were collected from a cohort of 30 patients. The values for sensitivity, specificity, accuracy, and the AUC across the nine varied deep learning models ranged from 0.273 to 0.960, 0.0533 to 0.947, 0.507 to 0.733, and 0.435 to 0.807, respectively. Notably, the EVA (eva_giant_patch14_336) model among the nine models stood out with the highest accuracy and AUC (accuracy = 0.733 and AUC = 0.807). Conversely, the Resnet50 (resnet50) model

Table 1 Clinical findings for ocular adnexal MALT lymphoma

Number (%) of patients	ALL	IgG4	Orbital MALT lymphoma	*p* value
• Total	127 (100)	53 (42)	74 (58)	
• Sex				
Male	55 (43)	24 (45)	31 (42)	
Female	72 (57)	29 (55)	43 (58)	N.S
• Age at presentation				
y (average)	66 ± 14	60 ± 14	70 ± 13	$P < 0.01$*
• Blood test				
IgG4 (mg/dL, $n = 49$)		528.8 ± 480.8		

IgG4 = immunoglobulin 4; MALT = mucosa-associated lymphoid tissue; N.S. = No significant changes. Values of $P < 0.05$ were considered significant; * = significant change

Table 2 Diagnostic performance of nine deep learning models in test set

Model name	Sensitivity	Specificity	Accuracy	AUC
EVA (eva_giant_patch14_336)	0.600	0.867	0.733	0.807
EVA (eva_giant_patch14_224)	0.687	0.727	0.707	0.789
ViT (vit_large_patch14_clip_336)	0.413	0.947	0.680	0.736
ViT (vit_base_patch16_224)	0.713	0.627	0.670	0.707
EfficientNet (tf_efficientnet_b5)	0.620	0.413	0.517	0.499
EfficientNet (tf_efficientnet_b3)	0.273	0.767	0.520	0.476
Densenet121 (densenet121)	0.567	0.493	0.530	0.521
Resnet50 (resnet50)	0.960	0.0533	0.507	0.435
Vgg16 (vgg16)	0.467	0.767	0.617	0.628

Abbreviations: area under the curve (AUC), vision transformer (ViT)

registered the lowest accuracy and AUC (accuracy = 0.507 and AUC = 0.435). As seen in Table 2, the diagnostic performances exhibited by convolutional neural network models, such as EfficientNet, Densenet121, Resnet50, and Vgg16, fell short as compared to that demonstrated by the non-convolutional neural network models, specifically the Vision Transformer and EVA models. For explainable AI, saliency maps of the EVA model (eva_giant_patch14_336) with the best values for accuracy and AUC were generated. Figure 3 shows representative images of the HE images and the saliency maps in IgG4-ROD and orbital MALT lymphoma. For the saliency maps, the white and black colors represent the focused and non-focused areas of the model. As seen in Fig. 3, to perform the binary classification between IgG4-ROD and orbital MALT lymphoma, the best model placed the emphasis on the shape and distribution of the stroma in the HE images. Figure 4A shows the curves of the receiver operating characteristics of the nine deep learning models. As seen in Table 2, the EVA (eva_giant_patch14_336) model achieved the best curve.

Comparison between the deep learning model and ophthalmologists

In the following section, we focus on the best model (eva_giant_patch14_336) from among the nine different deep learning models. Figure 4B shows the curve of the receiver operating characteristics of the best model. Furthermore, Fig. 4B presents four marks on the graph, which represent the diagnostic performance determined for the four ophthalmologists. As seen in Fig. 4B, the diagnostic performance of the best deep learning model was higher than that determined for the four ophthalmologists. Figure 5 shows the confusion matrices of the deep learning models and the four ophthalmologists. The accuracy of the best model was 0.733, while those for the four ophthalmologists were 0.520, 0.453, 0.413, and 0.580, respectively. The confusion matrices in Fig. 5 also show that the diagnostic performance of the best deep learning model was higher than that of the four ophthalmologists.

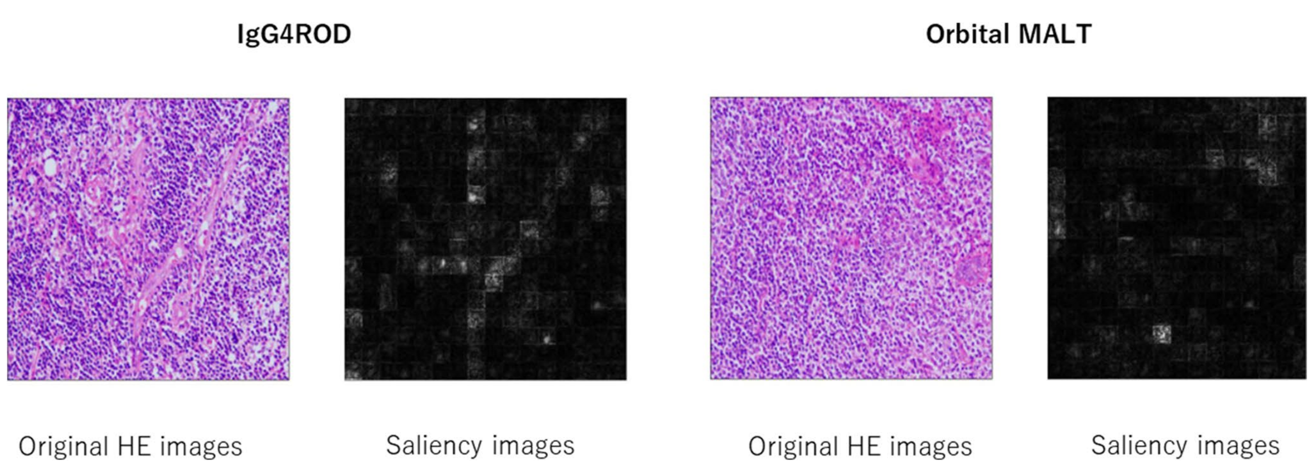

Fig. 3 Original and saliency images Left: IgG4-ROD, Right: orbital MALT lymphoma

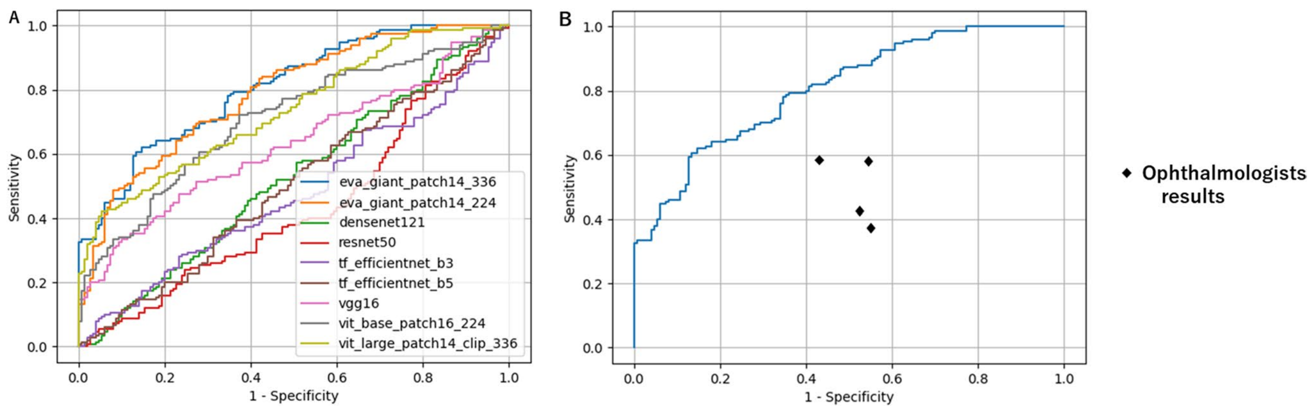

Fig. 4 ROC curve for the nine deep learning models

Discussion

The present results indicate that an accurate AI software can be developed using deep learning for the binary classifications of orbital MALT lymphoma and IGG4-ROD. Our results show that the binary classifications with non-convolutional neural networks were much better than those with convolutional neural networks. In the present study, the accuracy of the best deep learning model was higher than that of each of the four ophthalmologists, and orbital MALT lymphoma and IGG4-ROD were distinguishable with the model. This suggests that our deep-learning-based software could capture the existence of differences in the morphological characteristics of pathological images of HE-stained sections.

We believed that variations in gene expression clusters between different groups could reflect those in the cellular components within tumor masses. There is also an RNA-seq report that describes the difference between orbital MALT and IgG4ROD, and we recently reported that the gene expression of conjunctival MALT and orbital MALT may be significantly different, and that the nature of the tissue stroma in HE is a significant difference between these two groups of lymphomas [7, 18].

However, it could not be mentioned that AI discovered these genetic and biological changes, and it was likely that AI picked up small differences in pattern recognition that were too small for clinicians or pathologists to diagnose.

We also discovered that there can be significant differences and consistency. Based on this, we surmised that MALT and IgG4ROD could be differentiated not only by sophisticated immunohistological studies and examination of genetic expression, however also by simply changing the viewpoint using HE. In order to perfect the objectivity of scientific models, we utilized artificial intelligence (AI).

Consequently, it was probable that discrepancies between orbital MALT lymphoma and IgG4-ROD would also be present in HE-stained samples. Prior studies have already investigated HE images of lymphoma using AI, including deep learning models [19–22]. Additionally, we believed that the current protocol could provide the most objective method. Despite the fact that this research was constrained to data from just one institution, we successfully classified the HE images between orbital MALT lymphoma and IGG4-ROD by AI. Furthermore, the results of AI (accuracy = 0.733) were better than those of the four ophthalmologists (accuracy of about 40–60%). These objective results indicate a likelihood that the two groups studied here exhibit varying morphological characteristics. These morphological distinctions are likely indicative of disparities in cellular components, and they are in alignment with differences in morphological phenotype. In addition, immunostaining and other genetic tests remain important in diseases such as lymphoma, however we are able to look forward to the development of their potential by revisiting HE, as a relatively noise-free two-color image. Furthermore, it has been reported in recent years that Lymphoma occurs from IgG4ROD [23–26]. Although this study did not examine this difference in this report, our previous report suggests that the balance between tumor cells and non-tumor stroma (including fibrous components) may be important in differentiating them.

In the present study, the performance of deep learning models was significantly better than that of ophthalmologists. In general, the differentiation between orbital MALT lymphoma and IGG4-ROD by HE image alone is not often performed. This may be the reason why the accuracy and AUC of the ophthalmologists were lower than the best deep learning model. Generally, deep learning model extracts features from image data for classification and optimizes model parameters. In the present study, the deep learning model

Fig. 5 Confusion matrices of the four ophthalmologists and the best deep learning model in the test set

A Best deep learning model

		Prediction IgG4ROD	Prediction orbital MALT lymphoma
Ground truth	IgG4ROD	130	20
	orbital MALT lymphoma	60	90

B Dr. 1

		Prediction IgG4ROD	Prediction orbital MALT lymphoma
Ground truth	IgG4ROD	69	81
	orbital MALT lymphoma	63	87

C Dr. 2

		Prediction IgG4ROD	Prediction orbital MALT lymphoma
Ground truth	IgG4ROD	72	78
	orbital MALT lymphoma	86	64

may extract features from HE images that ophthalmologists do not understand. The incidence rate of orbital lymphoma is very low [27, 28]. In particular, the amount of tissue that can be biopsied is small, the number of specialized pathological diagnosticians is few, because it is a rare disease, so that we believe that initial screening tool to direct further ancillary investigations, and the simplest form of HE is medically and economically necessary.

As this result, it might be believed that the performance of the deep learning model much better than that of the average ophthalmologists.

Vision Transformer and EVA are non-convolutional neural network models [9]. EVA is a vanilla Vision Transformer

Fig. 5 (continued)

D Dr. 3

		Prediction IgG4ROD	Prediction orbital MALT lymphoma
Ground truth	IgG4ROD	68	82
	orbital MALT lymphoma	94	56

E Dr. 4

		Prediction IgG4ROD	Prediction orbital MALT lymphoma
Ground truth	IgG4ROD	86	64
	orbital MALT lymphoma	62	88

pretrained with a dedicated method. Based on our results of Table 2, it is speculated that the pretraining of EVA was important to achieve the diagnostic performance superior to that of ophthalmologists. In the pretraining of EVA, masked image modeling was used where EVA reconstructed image-text aligned vision features (CLIP features) from the masked image [10]. Using the pretraining, EVA could achieve robustness and generalization capability in various tasks such as image classification and image segmentation for general images. We believe that this generalization capability of EVA contributes to our results for HE images in the present study.

EVA is a new non-convolutional neural network model for image classification. The number of model parameters of EVA and the largest among the nine models presented in the present study, and EVA achieved state-of-the-art performance on general images. Our study shows that EVA was also useful for medical images. To the best of our knowledge, our study is the first to use EVA for medical image classification. From Fig. 4, non-convolutional neural networks performed better than convolutional neural networks. The usefulness of convolutional neural networks in medical image processing has been shown in various studies in the past. We expect that more studies like the present study will show that non-convolutional neural networks, such as Vision Transformer and EVA, can outperform convolutional neural networks.

In the present study, there are several limitations that should be considered when interpreting our results. First, all patients were included from a single institution, and only 1270 image patches were used, representing the limited-size dataset. In the future, a large dataset obtained from multiple centers would allow investigation of the robustness and generalizability of deep learning. In addition, some of the four ophthalmologists who conducted pathological evaluations were in the middle of their residency, which may result in a lower diagnostic rate compared to AI.

In conclusion, the AI software with deep learning was constructed to differentiate between orbital MALT lymphoma and IGG4-ROD. Basically, there are many Japanese and international institutions that can diagnosed these disease, albeit that immunohistochemistry and molecular studies are required, moreover, they require time and socio-economic costs, and it is often difficult to prepare all pathological specimens. and if simple initial screening can be performed using AI of HE image, it may be possible to provide more rapid and appropriate medical care. Of course, this is not an editorial that claims that immunostaining and gene rearrangement tests at the final facility are unnecessary.

Our results suggests that orbital MALT lymphoma and IGG4-ROD may have different morphological characteristics on HE images for initial screening tool to direct further ancillary investigations.

Acknowledgements We gratefully acknowledge the technical assistance of the Clinical Laboratory Department at Kobe Kaisei Hospital.

Authors' contributions MT and MN as equal first authors wrote the main text of the manuscript and prepared the figures. Datasets were prepared by AK, NM, AS, and YH. AA and SH reviewed the manuscript and checked the statistical analyses. All authors read and approved the manuscript in its final form.

Funding This work was supported by JSPS KAKENHI Grant Numbers 23K09013 and 22K07665, 2022 the Charitable Trust Fund for Ophthalmic Research in Commemoration of Santen Pharmaceutical's Founder, and the 2023 Osaka Community Foundation (MT and NM).

Data availability The data that support the findings of this study are available from the corresponding author (MT) upon reasonable request.

Declarations

Ethics declarations All procedures performed in studies involving human participants were in accordance with the ethical standards of the Institutional Review Board(IRB) of Osaka Metropolitan University and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Ethical Approval was provided by Osaka Metropolitan University IRB(approval number: 2022–064).

Informed consent Informed consent was obtained from all individual participants included in the study.

Consent for publication Not applicable.

Competing interests The authors declare no competing interests.

References

1. Goto H, Yamakawa N, Komatsu H, Asakage M, Tsubota K, Ueda SI, Nemoto R, Umazume K, Usui Y, Mori H (2021) Clinico-epidemiological analysis of 1000 cases of orbital tumors. Jpn J Ophthalmol 65:704–723. https://doi.org/10.1007/s10384-021-00857-1
2. Sogabe Y, Ohshima K-i, Azumi A, Takahira M, Kase S, Tsuji H, Yoshikawa H, Nakamura T (2014) Location and frequency of lesions in patients with IgG4-related ophthalmic diseases. Graefes Arch Clin Exp Ophthalmol 252:531–538
3. Deshpande V, Zen Y, Chan JK, Yi EE, Sato Y, Yoshino T, Klöppel G, Heathcote JG, Khosroshahi A, Ferry JA (2012) Consensus statement on the pathology of IgG4-related disease. Mod Pathol 25:1181–1192
4. Andrew NH, Sladden N, Kearney DJ, Selva D (2015) An analysis of IgG4-related disease (IgG4-RD) among idiopathic orbital inflammations and benign lymphoid hyperplasias using two consensus-based diagnostic criteria for IgG4-RD. Br J Ophthalmol 99:376–381
5. Cleary ML, Chao J, Warnke R, Sklar J (1984) Immunoglobulin gene rearrangement as a diagnostic criterion of B-cell lymphoma. Proc Natl Acad Sci 81:593–597
6. Shimizu H, Usui Y, Wakita R, Aita Y, Tomita A, Tsubota K, Asakage M, Nezu N, Komatsu H, Umazume K, Sugimoto M, Goto H (2021) Differential Tissue Metabolic Signatures in IgG4-Related Ophthalmic Disease and Orbital Mucosa-Associated Lymphoid Tissue Lymphoma. Invest Ophthalmol Vis Sci 62:15. https://doi.org/10.1167/iovs.62.1.15
7. Tagami M, Nishio M, Katsuyama-Yoshikawa A, Misawa N, Sakai A, Haruna Y, Azumi A, Honda S (2023) Machine learning model with texture analysis for automatic classification of histopathological images of ocular adnexal mucosa-associated lymphoid tissue lymphoma of two different origins. Curr Eye Res 1–8. https://doi.org/10.1080/02713683.2023.2246696
8. Yamashita R, Nishio M, Do RKG, Togashi K (2018) Convolutional neural networks: an overview and application in radiology. Insights Imaging 9:611–629. https://doi.org/10.1007/s13244-018-0639-9
9. Fang Y, Wang W, Xie B, Sun Q, Wu L, Wang X, Huang T, Wang X, Cao Y (2023) Eva: exploring the limits of masked visual representation learning at scale. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp 19358–19369. https://arxiv.org/abs/2211.07636
10. Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, Dehghani M, Minderer M, Heigold G, Gelly S (2020) An image is worth 16x16 words: transformers for image recognition at scale. https://arxiv.org/abs/2010.11929
11. Tan M, Le Q (2019) EfficientNet: rethinking model scaling for convolutional neural networks. In: Kamalika C, Ruslan S (eds) Proceedings of the 36th International Conference on Machine Learning. PMLR, Proceedings of Machine Learning Research, pp. 6105–6114. https://arxiv.org/abs/1905.11946
12. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ (2017) Densely connected convolutional networks. Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 4700–4708. https://ieeexplore.ieee.org/document/8099726
13. Miyoshi H, Sato K, Kabeya Y, Yonezawa S, Nakano H, Takeuchi Y, Ozawa I, Higo S, Yanagida E, Yamada K (2020) Deep learning shows the capability of high-level computer-aided diagnosis in malignant lymphoma. Lab Invest 100:1300–1310
14. Cheuk W, Yuen HK, Chan JK (2007) Chronic sclerosing dacryoadenitis: part of the spectrum of IgG4-related Sclerosing disease? Am J Surg Pathol 31:643–645
15. Umehara H, Okazaki K, Masaki Y, Kawano M, Yamamoto M, Saeki T, Matsui S, Yoshino T, Nakamura S, Kawa S, Hamano H, Kamisawa T, Shimosegawa T, Shimatsu A, Nakamura S, Ito T, Notohara K, Sumida T, Tanaka Y, Mimori T, Chiba T, Mishima M, Hibi T, Tsubouchi H, Inui K, Ohara H (2012) Comprehensive diagnostic criteria for IgG4-related disease (IgG4-RD), 2011. Mod Rheumatol 22:21–30. https://doi.org/10.1007/s10165-011-0571-z
16. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778. https://arxiv.org/abs/1512.03385
17. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. https://arxiv.org/abs/1409.1556
18. Tagami M, Kasashima H, Kakehashi A, Yoshikawa A, Nishio M, Misawa N, Sakai A, Wanibuchi H, Yashiro M, Azumi A, Honda S (2024) Stromal area differences with epithelial-mesenchymal transition gene changes in conjunctival and orbital mucosa-associated lymphoid tissue lymphoma. Front Oncol 14:1277749. https://doi.org/10.3389/fonc.2024.1277749

19. Swiderska-Chadaj Z, Hebeda KM, van den Brand M, Litjens G (2021) Artificial intelligence to detect MYC translocation in slides of diffuse large B-cell lymphoma. Virchows Arch 479:617–621
20. El Hussein S, Chen P, Medeiros LJ, Wistuba II, Jaffray D, Wu J, Khoury JD (2022) Artificial intelligence strategy integrating morphologic and architectural biomarkers provides robust diagnostic accuracy for disease progression in chronic lymphocytic leukemia. J Pathol 256:4–14
21. El Achi H, Khoury JD (2020) Artificial intelligence and digital microscopy applications in diagnostic hematopathology. Cancers (Basel) 12:797
22. Li D, Bledsoe JR, Zeng Y, Liu W, Hu Y, Bi K, Liang A, Li S (2020) A deep learning diagnostic platform for diffuse large B-cell lymphoma with high accuracy across multiple hospitals. Nat Commun 11:1–9
23. Sato Y, Notohara K, Kojima M, Takata K, Masaki Y, Yoshino T (2010) IgG4-related disease: historical overview and pathology of hematological disorders. Pathol Int 60:247–258
24. Bledsoe JR, Wallace ZS, Deshpande V, Richter JR, Klapman J, Cowan A, Stone JH, Ferry JA (2017) Atypical IgG4+ plasmacytic proliferations and lymphomas: characterization of 11 cases. Am J Clin Pathol 148:215–235
25. Bledsoe JR, Wallace ZS, Stone JH, Deshpande V, Ferry JA (2018) Lymphomas in IgG4-related disease: clinicopathologic features in a Western population. Virchows Arch 472:839–852
26. Li K-M, Xu M-H, Wu X, He W-M (2020) The expression of IgG and IgG4 in orbital MALT lymphoma: the similarities and differences of IgG4-related diseases. Onco Targets Ther 5755–5761. https://doi.org/10.2147/OTT.S242852
27. Jung SK, Lim J, Yang SW, Won YJ (2021) Nationwide trends in the incidence of orbital lymphoma from 1999 to 2016 in South Korea. Br J Ophthalmol 105:1341–1345. https://doi.org/10.1136/bjophthalmol-2020-316796
28. Olsen TG, Heegaard S (2019) Orbital lymphoma. Surv Ophthalmol 64:45–66. https://doi.org/10.1016/j.survophthal.2018.08.002

Authors and Affiliations

Mizuki Tagami[1,3] · Mizuho Nishio[2] · Atsuko Yoshikawa[3] · Norihiko Misawa[1] · Atsushi Sakai[1] · Yusuke Haruna[1] · Mami Tomita[1] · Atsushi Azumi[3] · Shigeru Honda[1]

✉ Mizuki Tagami
mizuki1979feb@yahoo.co.jp

[1] Department of Ophthalmology and Visual Sciences,, Graduate School of Medicine, Osaka Metropolitan University, 1-5-7 Asahimachi, Abeno-Ku, Osaka-Shi, Osaka 545-8586, Japan

[2] Center for Advanced Medical Engineering Research & Development, Kobe University Graduate School of Medicine, Kobe, Hyogo, Japan

[3] Ophthalmology Department and Eye Center, Kobe Kaisei Hospital, Kobe, Hyogo, Japan