



TITLE:

Integrating Preprocessing Operations into Deep Learning Model: Case Study of Posttreatment Visual Acuity Prediction

AUTHOR(S):

Otsuki, Ryo; Sugiyama, Osamu; Mori, Yuki; Miyake, Masahiro; Hiragi, Shusuke; Yamamoto, Goshiro; Santos, Luciano; ...
Matsumoto, Shigemi; Tsujikawa, Akitaka; Kuroda, Tomohiro

CITATION:

Otsuki, Ryo ...[et al]. Integrating Preprocessing Operations into Deep Learning Model: Case Study of Posttreatment Visual Acuity Prediction. *Advanced Biomedical Engineering* 2022, 11: 16-24

ISSUE DATE:

2022

URL:

<http://hdl.handle.net/2433/278464>

RIGHT:

Copyright: ©2022 The Author(s).; This is an open access article distributed under the terms of the Creative Commons BY 4.0 International (Attribution) License, which permits the unrestricted distribution, reproduction and use of the article provided the original source and authors are credited.

Integrating Preprocessing Operations into Deep Learning Model: Case Study of Posttreatment Visual Acuity Prediction

Ryo OTSUKI,^{*,#} Osamu SUGIYAMA,^{*} Yuki MORI,^{**} Masahiro MIYAKE,^{**} Shusuke HIRAGI,^{*}
 Goshiro YAMAMOTO,^{*} Luciano SANTOS,^{*} Yuta NAKANISHI,^{**} Yoshikatsu HOSODA,^{**} Hiroshi TAMURA,^{**}
 Shigemi MATSUMOTO,^{**} Akitaka TSUJIKAWA,^{**} Tomohiro KURODA^{*}

Abstract Designing a deep neural network model that integrates clinical images with other electronic medical records entails various preprocessing operations. Preprocessing of clinical images often requires trimming of parts of the lesions shown in the images, whereas preprocessing of other electronic medical records requires vectorization of these records; for example, patient age is often converted into a categorical vector of 10-year intervals. Although these preprocessing operations are critical to the performance of the classification model, there is no guarantee that the preprocessing step chosen is appropriate for model training. The ability to integrate these preprocessing operations into a deep neural network model and to train the model, including the preprocessing operations, can help design a multi-modal medical classification model. This study proposes integration layers of preprocessing, both for clinical images and electronic medical records, in deep neural network models. Preprocessing of clinical images is realized by a vision transformer layer that selectively adopts the parts of the images requiring attention. The preprocessing of other medical electrical records is performed by adopting full-connection layers and normalizing these layers. These proposed preprocessing-integrated layers were verified using a posttreatment visual acuity prediction task in ophthalmology as a case study. This prediction task requires clinical images as well as patient profile data corresponding to each patient's posttreatment logMAR visual acuity. The performance of a heuristically designed prediction model was compared with the performance of the prediction model that includes the proposed preprocessing integration layers. The mean square errors between predicted and correct results were 0.051 for the heuristic model and 0.054 for the proposed model. Experimental results showed that the proposed model utilizing preprocessing integration layers achieved nearly the same performance as the heuristically designed model.

Keywords: deep learning, automation of preprocessing, visual acuity prediction.

Adv Biomed Eng. 11: pp. 16–24, 2022.

1. Introduction

Recent studies have attempted the development of deep learning models for medical diagnoses and lesion detec-

tion; for example, the diagnosis of cervical dysplasia from images and medical test results [1] and detection of COVID-19 from X-ray, ultrasound, and CT scan images [2]. These studies utilized multiple modalities such as clinical images and electronic medical records (EMR) to develop medical diagnosis models.

Various preprocessing operations are required to design deep neural network models based on clinical images and other EMRs. These include resizing and cropping of clinical data, and normalization and conversion of EMRs to categorical vectors for 10-year intervals. Although preprocessing of the input data is critical to the performance of the classification model, this preprocessing may or may not be appropriate for training of the model.

Deep learning models have been proposed to be able to learn the features of the model regardless of the preprocessing content of the input data. For example, use of

This study was presented at the Symposium on Biomedical Engineering 2021, September, 2021.

Received on July 29, 2021; revised on November 7, 2021; accepted on November 24, 2021.

^{*} Graduate School of Informatics, Kyoto University, Kyoto, Japan.

^{**} Graduate School of Medicine, Kyoto University, Kyoto, Japan.

[#] 54 Shogoin-kawahara-cho, Sakyo-ku, Kyoto 606–8507, Japan.

E-mail: rotsuki@kuhp.kyoto-u.ac.jp



Copyright: ©2022 The Author(s). This is an open access article distributed under the terms of the Creative Commons BY 4.0 International (Attribution) License (<https://creativecommons.org/licenses/by/4.0/legalcode>), which permits the unrestricted distribution, reproduction and use of the article provided the original source and authors are credited.

a convolutional neural network (CNN) [3] has led to a proposed model that utilizes magnetic resonance images to predict the probability of developing Alzheimer's disease in patients with mild cognitive impairment [4]. The CNN used in that model can extract both the local and overall features of the images by changing the filter size of the convolution layer and inserting a pooling layer. These features can therefore be learned from only the critical part of the input image. In partial image recognition models that apply CNN, such as Faster R-CNN [5], classification may be possible after cutting out a specific position, thus identifying lesion sites [6, 7]. Based on information about insulin administration and other factors, a sequence-to-sequence model was able to predict rapid increases or decreases in a patient's blood glucose concentration [8]. The sequence-to-sequence model can be trained with data obtained during patient examination and primary data as time-series and interrelated data [9]. In these models, however, input data are preprocessed manually, and it is unclear whether this preprocessing is appropriate for training the model.

To address these problems, we have designed and integrated preprocessing layers for both clinical images and EMRs into multi-media deep neural network models. Preprocessing of clinical images is realized by a vision transformer layer that selectively evaluates the areas of interest in the images, while preprocessing of other EMRs is performed by full-connection layers, before and after normalization.

This study used fundus photographs (FP) and optical coherence tomography (OCT) images as clinical image data, and EMR including gender, age, affected side, and pretreatment decimal visual acuity (VA). Because input data differ markedly in size and range, these properties were accommodated by designing a proposed model that 1, input imaging and EMR data; 2, preprocessed and extracted features from imaging data; 3, preprocessed and extracted features from EMR data; 4, combined the extracted features; and 5, output the predicted VA as a numerical value.

This model does not require preprocessing of the input data, and predicts post-treatment VA with greater accuracy than a manual preprocessing model. Thus, this model may eliminate the need to consider and apply manual preprocessing on a task basis, and prevent loss of information from the input data during preprocessing.

2. Methods

Many deep learning models that receive medical data as input require preprocessing of the input data. Imaging data are frequently preprocessed by resizing and cropping. Resizing refers to the transformation of images to a predefined uniform vertical and horizontal size and is

designed to transform images to a size that the network can receive and align the input data conditions. Cropping refers to cutting out the part of the image evaluated at the time of diagnosis or cutting out the area of the lesion from the image. Cropping is performed to reduce noise in the input data and allow the model focus only on what is needed for evaluation. Standard preprocessing of EMRs includes normalization and categorization. Data are normalized to match the range of input data, whereas categorization assists the model in determining the relationship between one data point and the entirety of the data, as well as the order of the data; for example, by converting into a vector with clusters every 10 years [10].

These preprocessing processes are usually performed manually, but they have a significant impact on the accuracy of the model. Rather than manually preprocessing multiple medical modalities, we have proposed preprocessing integration layers in the neural network. **Figure 1** illustrates the processes of manual preprocessing and how our proposed integration layers replace them. Clinical images were separated into same size patches, and features were extracted using attention and fully connected (FC) layers. Normalization and FC layers were used to preprocess EMRs and to extract features. By training the model incorporating preprocessing, it is assumed that the data can be transformed according to the structure of the input data and the task.

2.1 Clinical image feature extraction replacing preprocessing

Figure 1a shows the method by which the proposed integration layers extract features from clinical images while illustrating the role of preprocessing. Clinical images were preprocessed by a vision transformer (ViT), one of the deep learning models for image recognition [11]. The ViT is a model for image classification that employs a transformer-like architecture rather than a CNN [12].

CNNs have several advantages such as avoiding the need for hand designed visual features; rather, they learn to perform tasks directly from the data. Although CNNs avoid hand crafted feature extraction, the architecture itself is designed for a certain image size. ViT uses transformer architecture originally designed for text-based tasks. An input image is represented as a sequence of image patches, similar to the sequence of word embeddings used when applying transformers to text, with the ViT directly predicting class labels for the image. These features enhance independence from image size, as well as increase the ability of the multi-head attention mechanism to learn the parts of the image to focus on.

Adoption of ViT may allow the development of clinical image feature extraction layers as well as replacing

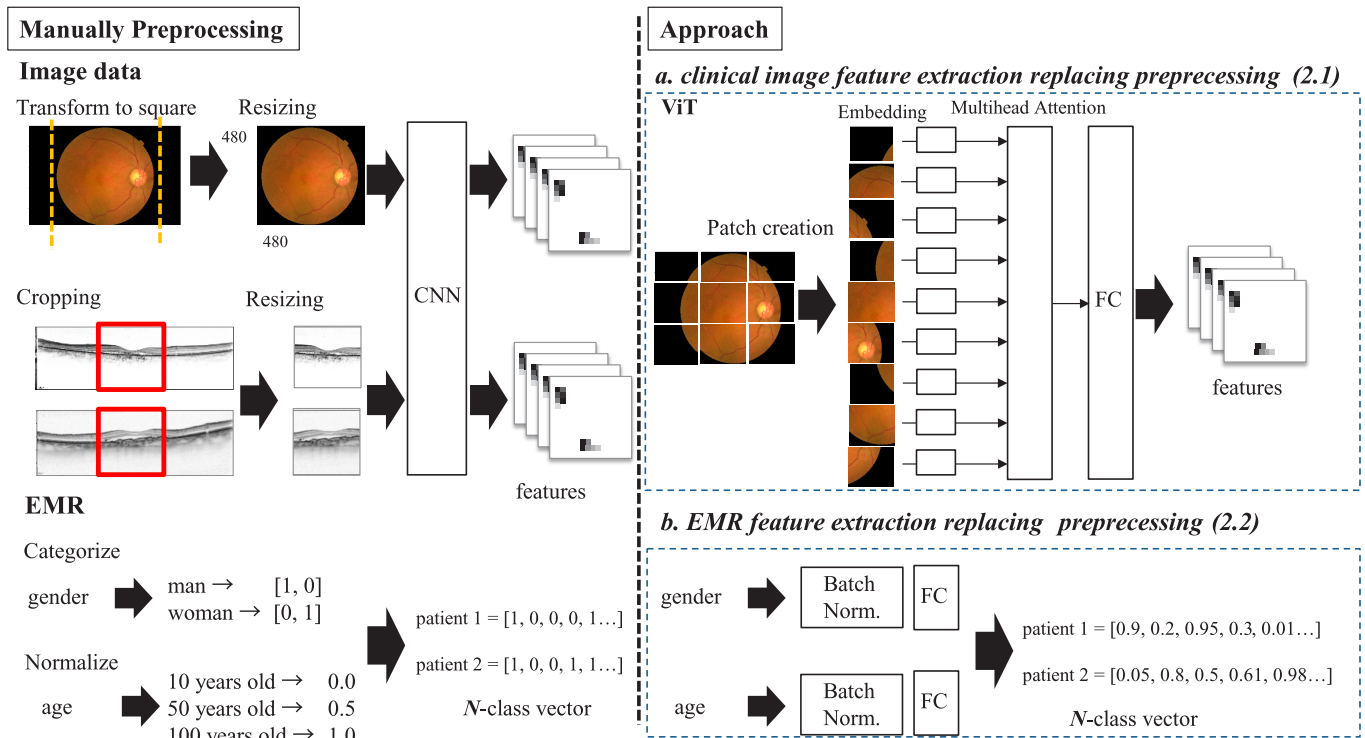


Fig. 1 ViT is used for image data and BN layer and FC layer for EMR to preprocess data and to extract features.

image preprocessing such as resizing and cropping.

2.2 EMR Feature Extraction, replacing Preprocessing

The scalar and discrete values in patient EMRs are essentially multi-dimensional. Manual preprocessing of EMRs include sorting into categorical multi-dimensional vectors based on clinical knowledge, such as whether the interpretation of blood test data does or does not differ in elderly and non-elderly patients.

Figure 1b illustrates the mechanism by which the proposed integration layers extract features of EMRs. Scalar values such as gender and age are input to the layers, which output the feature vectors of EMRs.

EMR preprocessing and feature extraction were performed using a batch normalization (BN) layer and an FC layer. The BN layer transforms the data so that the mean output is close to 0 and its standard deviation close to 1. After normalization, the FC layer extracts and linearly transforms the feature of interest from EMRs. Use of the BN and FC layers results in a structure responsible for preprocessing normalization and feature extraction.

2.3 Feature combination and regression layer

The features extracted from clinical images and EMRs are concatenated and passed into the fully connected layers for classification or regression tasks. To avoid biases, the dimension size of concatenated feature vectors for

each medical modality should be identical.

3. Experiments

Experiments were performed to assess the effectiveness of the preprocessing-integration layers. The experiment consisted of two tasks: evaluating the accuracy of the proposed model and visualizing the results following conversion of the data. Experiments were performed under four conditions, and the differences between predicted and actual outcomes were compared by determining the mean square error (MSE).

3.1 Case study and design of a model including the proposed integration layers

The performance of the proposed integration layers was assessed using a visual acuity prediction task. Specifically, a deep neural network model was developed to predict posttreatment VA in patients with age-related macular degeneration based on medical imaging and patient EMRs.

Figure 2 shows the outline of the proposed model. This model had three layers: (1) a layer that preprocesses and extracts features from imaging data, (2) a layer that preprocesses and extracts features from EMR data, and (3) a layer that combines the features of the first two layers and predicts VA.

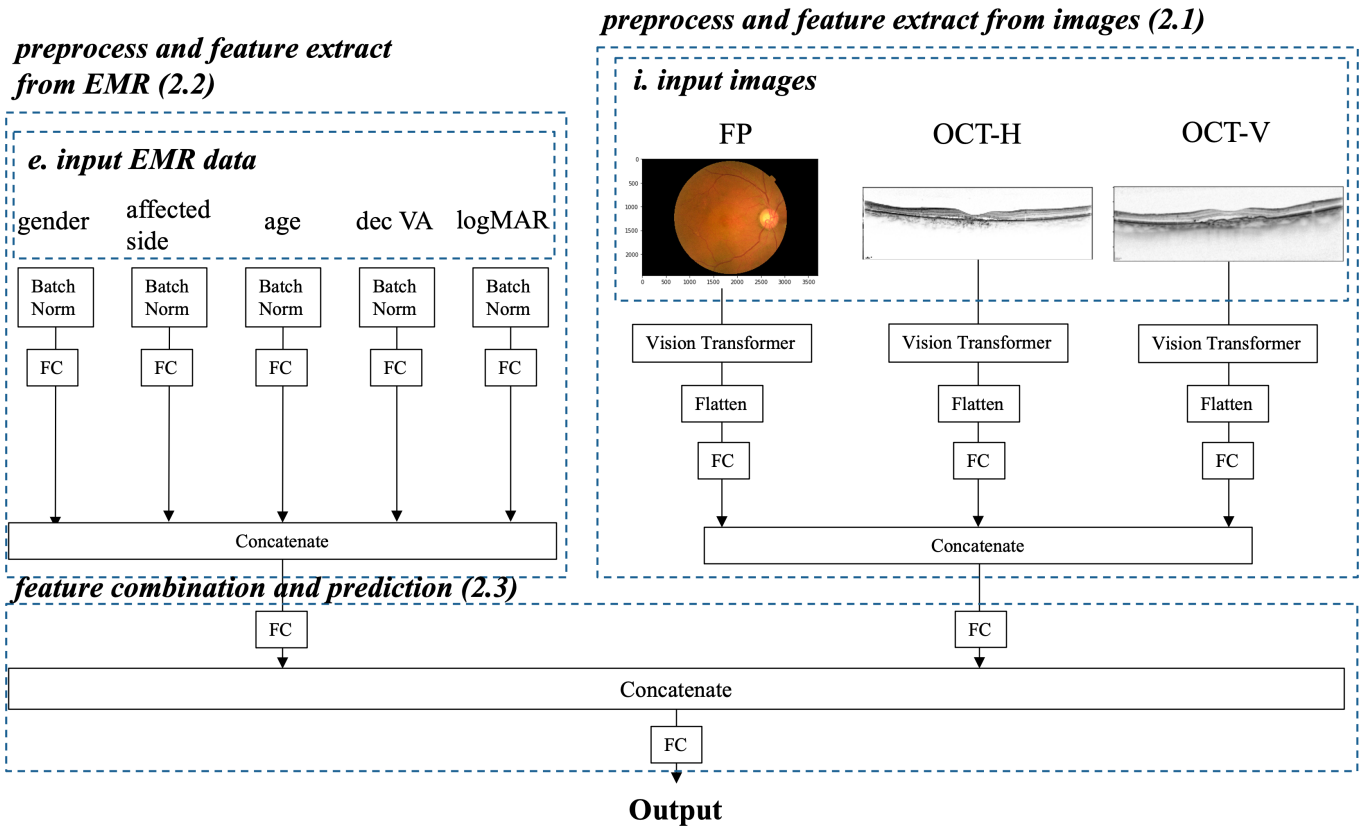


Fig. 2 Our proposed model for case study had three layers. The model receives images and EMR to predict postoperative visual acuity.

3.2 Datasets

Figure 2i shows examples of input imaging data, including FP and OCT images. OCT images have vertical and horizontal cross-sections, called OCT-v and OCT-h, respectively. The patient profile data included gender, age, affected side (right or left), and pretreatment decimal VA [13]. To analyze VA as a continuous variable, decimal VA was converted to logarithm of the minimal angle resolution (logMAR) [14]. Both pretreatment and post-treatment decimal VA were converted to logMAR.

This study was approved by the Ethics Committee of Kyoto University Graduate School and Faculty of Medicine (R2366) and adhered to The Ethical Guidelines for Medical and Health Research Involving Human Subjects in Japan as well as the tenets of the Declaration of Helsinki. Informed consent was obtained from each participant using an opt-out method, as permitted by the Ethics Committee.

Data were obtained from 315 patients who (1) visited the macular clinic at Kyoto University Hospital, (2) were diagnosed with wet age-related macular degeneration (AMD), and (3) completed a fixed regimen of intravitreal injection of aflibercept (IVA) for one year. They were separated into a training set of 252 patients and a validation set of 63 patients for five-fold cross-validation.

Manual preprocessing of the imaging data included resizing and cropping, whereas manual preprocessing of the EMRs included normalization and categorization. FP images were resized to 480×480 pixels and split into three RGB channels. The central square regions of OCT-v and OCT-h were cropped and resized to 480×480 pixels. **Figure 2e** shows the input EMR data. Gender and affected side were categorized (0 or 1). Age was subjected to min-max normalization, as shown in Eq. 1, and normalized age was multiplied by ten to convert to integers. Floating logMAR was converted to an integer by: (1) multiplying decimal VA by ten; and (2) adding the minimum logMAR and multiplying by ten.

$$\text{converted age} = \frac{\text{age} - \text{age}_{\min}}{\text{age}_{\max} - \text{age}_{\min}} \quad (\text{Eq. 1})$$

3.3 Experimental conditions

The predictive accuracy of the manually designed network was compared with that of the network obtained using the proposed preprocessing integration layers. **Table 1a** shows the structures of these models. The model incorporating preprocessing was found to preprocess the data and extract features by the ViT and BN layers, as described in **Fig. 2**. In the manually preprocessed model (“**Baseline**” model), image features were extracted by a

model called VGG16 [15], which is generally used for image classification, and features were extracted from EMR by an FC layer. Meanwhile, for the proposed model with integrated preprocessing layers (“**Proposed**” model), input image was not trimmed and the one-dimensional EMR data was divided into each item. For the “**Baseline**” model, input images and five-dimensional EMR were pre-processed as described in Section 3.2. Because the model incorporating pretreatment was expected to perform pretreatment more effectively for prediction, we hypothesized that MSE would be lower in the model that included preprocessing layers than in the model that included manual preprocessing.

To determine whether the preprocessing layer was more effective for imaging or EMR data, the accuracy of the models was compared in **Table 1b**. When the preprocessing layer was applied only to the image, ViT was applied to the image and the FC layer was applied to the EMR data. The latter was not divided into items but were input in a five-dimensional state. When the preprocessing layer was applied only to EMR, VGG16 was applied to the image and the BN layer was applied to EMR. Thus, EMR, which could only be converted at regular intervals by manual operations, would likely form a cluster more suitable for the prediction task. Therefore, we hypothesized that MSE would be higher when the preprocessing layer was applied only to the image than when it was applied only to EMR.

Finally, we visualized how the EMR was transformed by the preprocessing layer and compared it with the diagnosis by an ophthalmologist. In this experiment, after training the proposed model, validation data were input, and the features before integrating the features in the EMR preprocessing layer were output. Principal component analysis was applied to this feature, with the results presented as a hierarchical clustering image. The cluster was expected to be divided into two parts, with thresholds for age, pretreatment decimal VA, and pretreatment logMAR. We hypothesized that these thresholds would be consistent with those determined by an ophthalmologist.

3.4 Results

Table 2 shows the MSE between predicted and actual logMAR. The effect of the preprocessing layer was determined by comparing the MSE of the **Baseline** model and **Proposed** model. The MSE of the **Proposed** model with the preprocessing layers was nearly the same as the **Baseline** model. A comparison of the MSEs of the **Only image preprocessing** model and **Only EMR preprocessing** model showed that the MSE was larger when the preprocessing layer was applied only to the imaging data.

Figure 3 shows the hierarchical clustering image of the pretreatment patient age output from the preprocessing layer, which was subjected to principal component

Table 1 Conditions of compared models.

model name	layer for images	layer for EMR
<i>a</i> Baseline	VGG16	FC
Proposed	ViT	BN and FC
<i>b</i> Baseline	VGG16	FC
Only image preprocessing	ViT	FC
Only EMR preprocessing	VGG16	BN and FC

Table 2 MSE results.

model name	MSE for training	MSE for validation
<i>a</i> Baseline	0.049	0.051
Proposed	0.039	0.054
<i>b</i> Baseline	0.049	0.051
Only image preprocessing	0.050	0.118
Only EMR preprocessing	0.068	0.061

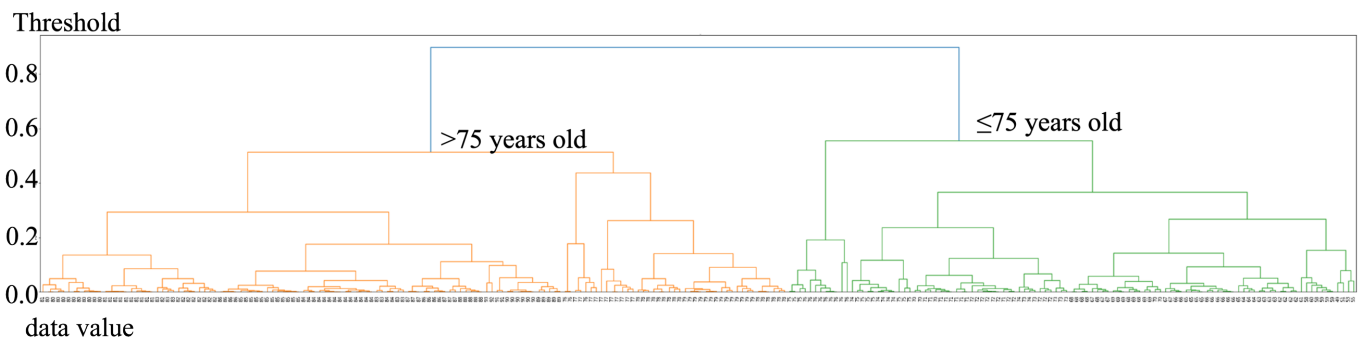


Fig. 3 As a result of clustering, patient pretreatment age was divided into two clusters. The cluster threshold was aged 75 years.

analysis (PCA) to make it two-dimensional. When hierarchical clustering [16] was performed after applying PCA, two clusters appeared, as indicated by the color of the dots in the scatter plot. One cluster contained data from patients aged < 75 years before treatment, and the other cluster contained data from patients aged ≥ 75 years. This age cutoff is essential for predicting pretreatment logMAR.

4. Discussion

This section discusses the significance, performance, limitations, and applicability of the proposed preprocessing layer based on experimental results.

4.1 Performance of the proposed preprocessing layers

In this study, when images were preprocessed by ViT, and EMR data were preprocessed by the BN and FC layers, the MSE between predicted and actual results was 0.054. A logMAR difference of < 0.20 has been reported to indicate the effectiveness of a particular treatment [17, 18]. Because the expected postoperative logMAR error was < 0.20 , the model had sufficient prediction accuracy for practical clinical use.

Moreover, the MSE of the **Baseline** model was 0.051 (< 0.20), and the proposed integration layer had nearly the same performance as the **Baseline** model. Thus, the proposed layer preprocessed data to predict postoperative VA with the same performance as manual preprocessing.

A comparison of each integration layer showed that preprocessing of EMR alone had a smaller MSE than preprocessing of images alone. This finding indicated that optimizing EMR data was more important than optimizing imaging data in predicting VA. Moreover, this finding is consistent with results showing that determination of preoperative VA by an ophthalmologist had the greatest effect on prediction accuracy.

Although the validation MSEs of **Only image preprocessing** and **Only EMR preprocessing** were higher than that of **Baseline**, the total validation MSEs of **Baseline** and our proposed method were almost the same. This fact indicates that the EMR features and image features extracted by our proposed method had coordinated impact compared with the Baseline method.

4.2 Integration of image preprocessing

The main advantage of integrating preprocessing operations into a deep learning model is to reduce the burden for optimizations. Manual preprocessing for image resizing and cropping requires medical knowledge of the lesion site, as well as manual operations for each image. The vision transformer, which has a multi-head attention

mechanism, can automatically detect essential parts of clinical images, resulting in a lower load for designing the model.

In addition, the proposed layer will be accepted regardless of the size or shape of the image. Because a complete image can be input with ViT, one advantage may be the lack of potential information loss, such as cropped images during manual preprocessing.

4.3 Integration of EMR preprocessing

Optimization is enhanced by integrating preprocessing of EMR data. For example, manual optimization of age requires determination of whether age should be dichotomized or processed as a categorical vector of 10-year intervals. Optimal conversion requires determination of the critical age threshold for diagnosis.

The proposed layer dichotomized patients by age into those aged < 75 and ≥ 75 years, resulting in a predictive accuracy greater than that obtained from manual conversion of age to 10-year intervals. The advantage of the proposed layer for EMR is that appropriate preprocessing can be performed without special knowledge of the disease.

4.4 Limitations

This report describes the construction of a layer that performs two types of image preprocessing; resizing and cropping. Although these two are the most commonly used preprocessing steps, other adjustments may include adjustments of rotation, brightness, and color tone. To learn these preprocessing steps, it is necessary to improve the structure of the model. Preprocessing of EMR data frequently consists of quantitation that replaces EMR data with an array of zeros and ones. Improvements of the model are required to achieve explicit quantitation.

5. Conclusion

This study proposed the integration of preprocessing layers for both clinical images and EMRs in deep neural network models. In the proposed model, ViT and BN layers were utilized for preprocessing of and feature extraction from medical images and EMR data, respectively.

To verify the effectiveness of the preprocessing layer, the ability of the model to predict VA was evaluated by comparing the MSE between actual and predicted logMAR in models with and without the preprocessing layers. The MSE was 0.054 with and 0.051 without the preprocessing layers. The experimental results revealed that the regression model with proposed preprocessing layers achieved an accuracy nearly the same as manual pre-processing.

The proposed preprocessing integration layers have several main advantages. By learning preprocessing, the input data could be converted into more efficient feature vectors for the prediction/regression task, and the accuracy of the output result may also be improved. In addition, there is no need to manually preprocess the input data, thus reducing the time and effort required to create training data. Because the proposed layer can handle input data regardless of its shape, it can likely be applied to various clinical decision support systems with multiple modalities, such as clinical images and EMRs.

Conflicts of Interest

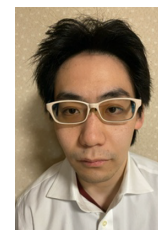
The authors declare they have no conflicts of interest with any companies or commercial organizations per the definition of the Japanese Society for Medical and Biological Engineering.

References

- Xu T, Zhang H, Huang X, Zhang S, Metaxas D N: Multimodal deep learning for cervical dysplasia diagnosis. *Medical Image Computing and Computer-Assisted Intervention -- MICCAI 2016*, 115–123, 2016.
- Liang M, Li Z, Chen T, Zeng J: Integrative data analysis of multi-platform cancer data with a multimodal deep learning approach. *IEEE/ACM Trans Comput Biol Bioinform.* **12**(4), 928–937, 2015.
- Lecun Y, Bottou L, Bengio Y, Haffner P: Gradient-based learning applied to document recognition. *Proc IEEE.* **86**(11), 2278–2324, 1998.
- Lin W, Tong T, Gao Q, Guo D, Du X, Yang Y, Guo G, Xiao M, Du M, Qu X: Alzheimer's Disease Neuroimaging Initiative: Convolutional neural networks-based MRI image analysis for the Alzheimer's disease prediction from mild cognitive impairment. *Front Neurosci.* **12**, 777, 2018.
- Ren S, He K, Girshick R, Sun J: Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Trans Pattern Anal Mach Intell.* **39**(6), 1137–1149, 2017
- Kabid HS, Samrat KD, Tahzib I, Mahbubur R: COVID faster R-CNN: A novel framework to diagnose novel coronavirus disease (COVID-19) in X-ray images. *Inf Med Unlocked.* **20**, 100405, 2020.
- Ezhilarasi R, Varalakshmi P: Tumor detection in the brain using raster R-CNN. *2018 2nd International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)* 388–392, 2018.
- Bhimireddy AR, Sinha P, Oluwalade B, Gichoya JW, Purkayastha S: Blood glucose level prediction as time-series modeling using sequence-to-sequence neural networks. *ECAI.* **2020**, 125–130, 2020.
- Sutskever I, Vinyals O, Le QV: Sequence to sequence learning with neural networks. *Proc of the 27th International Conference on Neural Information Processing Systems* **2**, 3104–3112, 2014
- Ryo O, Osamu, S, Yuki M, Masahiro M, Shusuke. H, Goshiro Y, Luciano S, Yuta N, Yoshikatsu H, Hiroshi T, Shigemi M, Akitaka T, Tomohiro K: Deep learning model to predict postoperative visual acuity from preoperative multimedia ophthalmic data. *Adv Biomed Eng.* **9**, 241–248, 2020.
- Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, Dehghani M, Minderer M, Heigold G, Gelly S, Uszkoreit J, Houlsby N: An image is worth 16x16 words: transformers for image recognition at scale. *International Conference on Learning Representations* 2021.
- Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser L, Polosukhin I: Attention is all you need. *NIPS.* **2017**, 5998–6008, 2017
- Lange C, Feltgen N, Junker B, Schulze-Bonsel K, Bach M: Resolving the clinical acuity categories “hand motion” and “counting fingers” using the Freiburg Visual Acuity Test (FrACT). *Graefes Arch Clin Exp Ophthalmol.* **247**(1), 137–142, 2009.
- Ferris FL 3rd, Kassoff A, Bresnick GH, Bailey I: New visual acuity charts for clinical research. *Am J Ophthalmol.* **94**(1), 91–96, 1982.
- Simonyan K, Zisserman A: Very deep convolutional networks for large-scale image recognition. *International Conference on Learning Representations* 2015, 2015
- Johnson SC. Some methods for classification and analysis of multivariate observations. *Psychometrika.* **32**(3), 241–254, 1967
- Rosser DA, Cousens SN, Murdoch IE, Fitzke FW, Laidlaw DA: How sensitive to clinical change are ETDRS logMAR visual acuity measurements? *Invest Ophthalmol Vis Sci.* **44**(8), 3278–3281, 2003.
- Kaiser PK: Prospective evaluation of visual acuity assessment: a comparison of Snellen versus ETDRS charts in clinical practice (An AOS Thesis). *Trans Am Ophthalmol Soc.* **107**, 311–324, 2009.

Ryo OTSUKI

Ryo OTSUKI is currently a Ph.D. student at Department of Social Informatics, Graduate school of Informatics, Kyoto University, Japan. He received his bachelor degree and master degree from College of Information Science and Engineering, Ritsumeikan University. His current research interest includes machine learning and medical informatics.



Osamu SUGIYAMA

Osamu SUGIYAMA received his PhD from School of Science for Open and Environmental Systems, Keio University in 2009. He is currently a Program-Specific Associate Professor in the Department of Real World Data Research and Development, Graduate School of Medicine, Kyoto University. His research fields are Medical Informatics, Communication Robot, Robot Audition. His research interests are in machine learning utilizing medical information and health promotion with personal health records. He is affiliated with the Robotic Society of Japan, the Japanese Society for Artificial Intelligence and the Architectural Institute of Japan.



Yuki MORI

Yuki MORI received his Doctor of Medicine, Kyoto University in 2014. He is currently an ophthalmologist and a graduate student at the Kyoto University Graduate School of Medicine. His research interests are in machine learning utilizing medical information and analysis of disease susceptibility genes. He is affiliated with Japanese Ophthalmological Society, Japan Ophthalmologist Association, Japanese Retina and Vitreous Society, and Japanese Society of Artificial Intelligence in Ophthalmology.



Masahiro MIYAKE

Masahiro MIYAKE received his PhD from Kyoto University Graduate School of Medicine in 2015, and Master of Public Health from Harvard School of Public Health in 2016. He is currently an assistant professor, Department of Ophthalmology, Kyoto University Graduate School of Medicine. His research fields are clinical ophthalmology, epidemiology, genetics, and clinical application of artificial intelligence. He is a board member of Japan Society of Artificial Intelligence in Ophthalmology (JSAIO) and affiliated with the Japanese Ophthalmological Society, Japanese Retina and Vitreous Society, and Japan Myopia Society.



Shusuke HIRAGI

Shusuke HIRAGI graduated medical school in 2012 and received medical license. He has been working as a nephrologist since then. He also received Ph.D. in medicine from Kyoto University in 2020 and is currently working as CIO at Kitano Hospital in Osaka, after finishing the term as an assistant professor of medical informatics laboratory at Kyoto University. His research interests are data analysis and secondary usage of health/medical information as well as data-driven health management. He is affiliated with the Japan Association for Medical Informatics, the Japanese Society of Nephrology, and the Japanese Society of Internal Medicine.



Goshiro YAMAMOTO

Goshiro YAMAMOTO received his BE, ME, and PhD in engineering from Osaka University. He is currently an associate professor in Kyoto University Hospital after being an assistant professor in Okayama University and Nara Institute of Science and Technology. His major research interest is human-computer interaction, digital transformation, and medical informatics.



Luciano Henrique de Oliveira SANTOS

Luciano Henrique de Oliveira SANTOS is currently an Assistant Professor at Kyoto University, Division of Medical Informatics. He received his bachelor's degree in Computer Science and his master's degree in Informatics both from University of Brasilia, Brazil, and his Ph.D. in Informatics from Kyoto University, Japan. His current research interests include pervasive games, augmented reality and gamification in different health application, such as rehabilitation therapy and supporting elderly people.



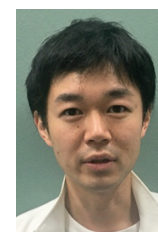
Yuta NAKANISHI

Yuta NAKANISHI received his Doctor of Medical Science, Kobe University in 2017. He is currently an Ophthalmologist, Graduate School of Medicine, Kyoto University. His research interests are in Medical imaging utilizing machine learning. He is affiliated with the Japanese Ophthalmological Society, and Japanese Society of Artificial Intelligence In Ophthalmology.



Yoshikatsu HOSODA

Yoshikatsu HOSODA specializes in retinal diseases. He received his MD from Kyoto University in 2011, and his PhD from Kyoto University Graduate School of Medicine in 2020. He is currently an ophthalmologist in the Osaka Red Cross Hospital. His particular interests include macular diseases, retinal surgery and cataract surgery. He is affiliated with the Japanese Ophthalmological Society and the Japanese Retina and Vitreous Society.



Hiroshi TAMURA

Hiroshi TAMURA is a professor of the Center for Innovative Research & Education in Data Science in Kyoto University. He graduated with Ph.D. from Kyoto University Graduate School of Medicine in 2006 and MS in epidemiology from Harvard School of Public Health in 2013. He has previously worked in Kobe City General Hospital from 1998–2002 at the department of ophthalmology as junior & senior resident. Research interests include ophthalmology, age-related macular degeneration, hospital management, epidemiology, and data science.



Shigemi MATSUMOTO

Shigemi MATSUMOTO is currently a program-specific Professor in the Department of Real World Data Research and Development, Graduate school of Medicine, Kyoto University. He works at Kyoto university hospital as one of head of medical oncologist. His subspecialty is gastrointestinal, hepatobiliary and rare cancer. He received his PhD from Graduate School of Medicine, Kyoto University in 2009. His research fields are medical information and health utilizing Real World Data (RWD) in Oncology. He is a member of JSMO, JSCO, ASCO, ESMO and JAMI.



Akitaka TSUJIKAWA

Akitaka TSUJIKAWA, MD, PhD graduated from the Kyoto University School of Medicine in 1993 and received his doctoral degree in ophthalmology from Kyoto University Graduate School of Medicine in 2001. In 1999–2000, he undertook a research fellowship at the Children's Hospital in Boston; he joined the faculty of the Department of Ophthalmology at Kyoto University Graduate School Medicine in 2005. After working as a Professor and Chairman of the Department of Ophthalmology at Kagawa University School of Medicine for three years, Dr. Tsujikawa is currently Professor and Chairman of Department of Ophthalmology at Kyoto University Graduate School of Medicine. Now, his interest focuses on the clinical investigations with the use of various latest imaging modalities, such as OCT, AO-SLO, FA, ICGA, and FAF. With the maximal use of these modalities, he contributed to our understanding of the pathophysiology of various macular diseases, especially wet age-related macular degeneration and polypoidal choroidal vasculopathy.

**Tomohiro KURODA**

Tomohiro KURODA is the CIO (the director of division of medical information technology and administration planning) of Kyoto University Hospital, and the professor of medical informatics in graduate school of medicine and graduate school of informatics of Kyoto University. Besides serving as IT manager of healthcare organization for a decade, he has been a researcher of applied informatics, especially mixed reality and ubiquitous computing applications for medicine and welfare, since he received Ph.D. in information science from Nara institute of science and technology (NAIST), Japan in 1998.

