LETTER

GUI System to Support Cardiology Examination Based on Explainable Regression CNN for Estimating Pulmonary Artery Wedge Pressure

Yuto OMAE[†a)], ***Member***, **Yuki SAITO**[††], **Yohei KAKIMOTO**[†], **Daisuke FUKAMACHI**[††], **Koichi NAGASHIMA**[††], **Yasuo OKUMURA**[††], ***and*** **Jun TOYOTANI**[†], ***Nonmembers***

SUMMARY In this article, a GUI system is proposed to support clinical cardiology examinations. The proposed system estimates "pulmonary artery wedge pressure" based on patients' chest radiographs using an explainable regression-based convolutional neural network. The GUI system was validated by performing an effectiveness survey with 23 cardiology physicians with medical licenses. The results indicated that many physicians considered the GUI system to be effective.

key words: *pulmonary artery wedge pressure, cardiology examination, convolutional neural network, regression activation map*

1. Introduction

As heart failure is often fatal, early detection of its symptoms is invaluable. Usually, the cardiac state is estimated by measuring the pulmonary artery wedge pressure (PAWP) via right heart catheterization. However, this method is invasive and presents risks of complications to patients [1]. Alternatively, physicians can estimate PAWP based on chest radiograph [2]. However, owing to the subjectivity of this method, the estimation accuracy depends significantly on the physician's skill. Therefore, the development of objective and non-invasive PAWP-measurement methods is required [3]. In this context, a classification convolutional neural network (CNN) has been proposed to detect PAWP above a threshold of 18 mmHg [4]. Further, regression CNN (R-CNN) has been developed to estimate PAWP based on chest radiographs [5]. These CNNs output class/regression activation maps (CAM/RAM)—thus, the activated region can be investigated while estimating PAWP (henceforth, such CNNs are referred to as "explainable CNN").

Although the systems proposed in the aforementioned studies are effective, the suitability of CNNs for clinical use by physicians has not been investigated—only the development of the models and their reliability evaluation have been reported. To address this shortcoming, we develop a graphical user interface (GUI) to estimate PAWP and output RAM based on chest radiographs of patients. The proposed system, depicted in Fig. 1, is based on the R-CNN developed by Saito et al. [5]. The development process and estimation of its reliability was reported in a previous study [5]. In this article, we explain the functions of the GUI system and report the results of an effectiveness survey performed by 23 cardiology physicians with medical licenses.

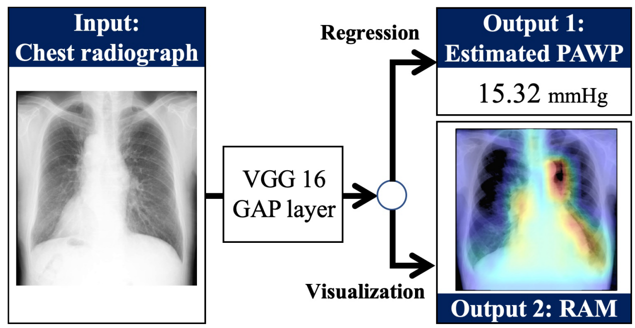

Fig. 1 Developed R-CNN [5].

2. The Developed GUI System

A screenshot of the GUI system is presented in Fig. 2, and its functions are listed in Table 1.

First, we describe Function 1, i.e., "Estimating PAWP", listed in Table 1. When the physician pushes the "Select radiograph" button illustrated in Fig. 2, the system opens a window for choosing a patient's chest radiograph. Subsequently, the chest radiograph selected by the physician is input into the R-CNN, and the system displays the estimated PAWP in the bottom-right corner, as depicted in Fig. 2. In the illustrated case, the estimated PAWP is 21.55 ± 0.80 mmHg. By referring to Forrester's subset [7], when the estimated PAWP exceeds 18 mmHg, the system displays "High risk". Saito et al. [5] developed the R-CNN by optimizing the hyperparameters to minimize the validation error via 5-hold cross validation. Corresponding to an estimated PAWP of $a \pm b$ mmHg, a and b indicate the average and standard deviation of estimations obtained using five models.

Now, we explain Functions 2 and 3 listed in Table 1. The explainable R-CNN reported in [5] outputs RAM dur-

Manuscript received July 21, 2022.
Manuscript revised October 6, 2022.
Manuscript publicized December 8, 2022.

[†]The authors are with the Department of Industrial Engineering and Management, College of Industrial Technology, Nihon University, Narashino-shi, 275–8575 Japan.
[††]The authors are with the Division of Cardiology, Department of Medicine, Nihon University School of Medicine, Tokyo, 173–8610 Japan.
a) E-mail: oomae.yuuto@nihon-u.ac.jp

DOI: 10.1587/transinf.2022EDL8059

Fig. 2 Screenshot of the developed GUI system. The heat map is a RAM [6], in which the red region signifies the region the R-CNN considers important for estimating PAWP. Conversely, the blue region is considered unimportant.

Table 1 Functions.

id	functions
Function 1	Estimating PAWP
Function 2	Viewing RAM
Function 3	Scaling RAM
Function 4	Viewing structure of R-CNN
Function 5	Viewing reliability of estimation

ing the estimation of PAWP. The heat-map of RAM is depicted in Fig. 2, and physicians can investigate the activated region of the R-CNN during the estimation of PAWP. Omae et al. reported that PAWP estimation error is high when the activated region contains areas in addition to the cardiac region [8]. In other words, RAM verification by physicians before the utilization of the estimated PAWP is important. To this end, we propose a function for scaling RAM. The scale bar is depicted in Fig. 2.

Next, we describe Functions 4 and 5 listed in Table 1. As the primary function of the GUI system is PAWP estimation, the estimation method and its reliability are of interest to physicians. This was the motivation to develop these functions. When the physician pushes the "R-CNN structure" button illustrated in Fig. 2, the system displays the R-CNN layer structure, as depicted in Fig. 1 in [5]. Moreover, when the physician pushes the "Reliability information" button illustrated in Fig. 2, the system displays a scatter plot about the ground truth PAWP and estimated PAWP, their correlation coefficient, and the absolute estimation error. The scatter plot and their correlation coefficient are illustrated Fig. 3 in [5].

3. Survey for Evaluation of Effectiveness

3.1 Method

To evaluate the proposed GUI system, 23 cardiology physicians with medical licenses in Japan, participated in a questionnaire survey. All the physicians were experienced cardiologists with at least three years of experience. We created a video explaining how to use the GUI system, the functions listed in Table 1, the estimation reliability of R-CNN [5], and the results obtained after inputting two patients' X-ray images (high risk / low risk of heart failure). This video was played for the physicians. Subsequently, the physicians were asked to respond to the question "Do you think the k-th function is effective?". As five functions are considered, five such questions were considered. In addition, physicians answered the question "Do you feel you want to use the GUI system?". Thus, in aggregate, six questions were considered. All responses were recorded on a six-point Likert-scale with the following levels—1: very negative, 2: negative, 3: slightly negative, 4: slightly positive, 5: positive, 6: very positive. In addition, physicians recorded their overall impressions about the GUI system.

3.2 Results and Discussion

As each physician responded to each question, all data points from the effectiveness survey were used for analysis. The results are illustrated in Fig. 3. The horizontal and vertical axes in the figure represent the different responses and the numbers of physicians who selected it, respectively. The responses to each question are recorded in separate figures. "P. rate" and "N. rate" denote the rates of positive and

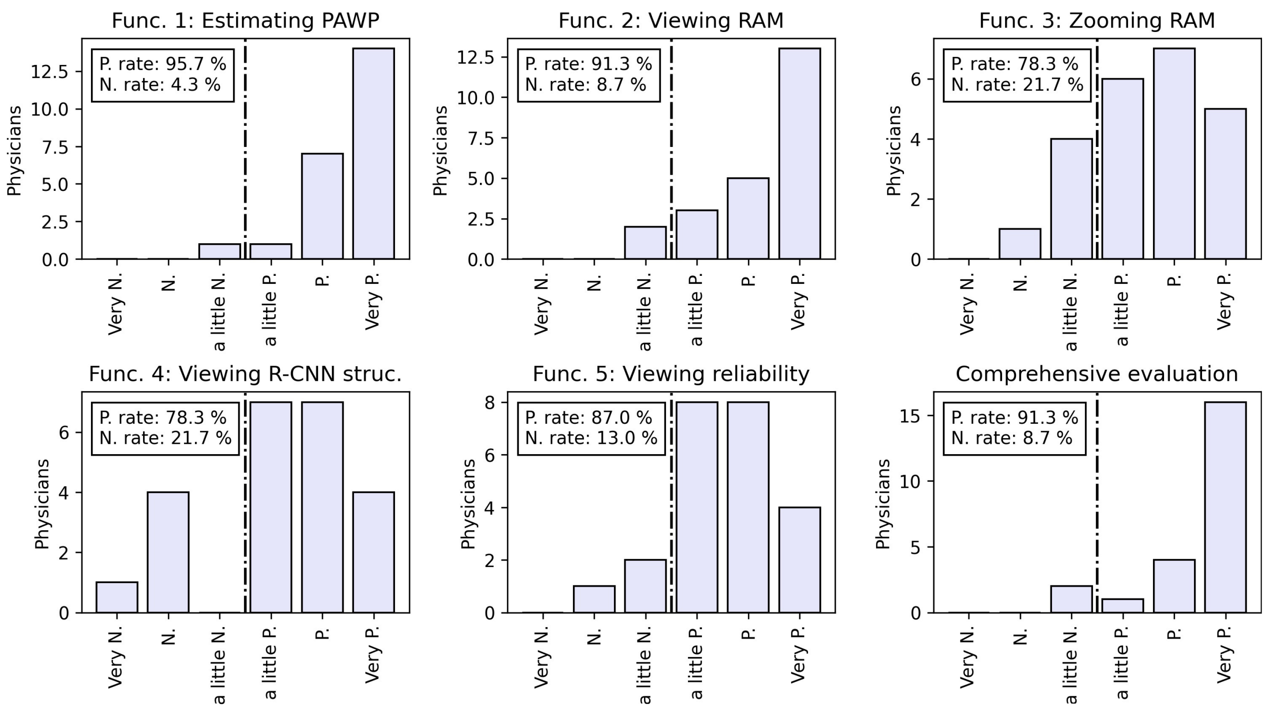

Fig. 3 The results of the effectiveness survey performed by 23 physicians. The horizontal axis represents the physicians' choices. "N." and "P." indicate negative and positive, respectively. The vertical axis represents the number of physicians. Functions 1 ~ 5 are presented in Table 1. "Comprehensive evaluation" indicates the response to the question "Do you feel you want to use the GUI system?".

negative answers, respectively.

Regarding Function 1 (estimating PAWP), the most common response was "very positive" with a high overall positive rate (exceeding 95%). Right heart catheterization is an invasive PAWP measurement method, which may lead to complications in patients. In contrast, the proposed GUI system only requires chest radiograph of patients for PAWP estimation—thus, this method is non-invasive and objective. The overall perception of high effectiveness of Function 1 among physicians is attributed to this.

Regarding Function 2 (viewing RAM), high effectiveness is noted from the responses, as in the case of Function 1. As RAM represents the activated region in the R-CNN during the estimation of PAWP, it is essential for the verification of the reliability of the proposed method. Moreover, when the estimated PAWP is high, the activated region represents an area of anomalous cardiac function. Therefore, RAM visualization is important to physicians. For the same reason, Function 3 (scaling RAM) is also important. We attribute the perceived effectiveness of Functions 2 and 3 among physicians to this reason.

Regarding Function 4 (viewing R-CNN structure), although the positive rate was approximately 80%, a certain number of negative responses is noted. This is attributed to the technical nature of CNNs, which are advanced systems in information science, that is difficult to understand for medical professionals.

Regarding Function 5 (viewing reliability of estimation), the positive rate was 87%. This function is related to the scatter plot, correlation coefficient, and absolute error between the ground truth PAWP measured via right heart catheterization and the PAWP estimated via R-CNN. The high positive rate is attributed to the importance of estimation reliability during clinical application.

Finally, we describe the responses to "comprehensive evaluation". "Very positive" was the most common response, with an overall positive rate of 90%. Therefore, we concluded that many medical physicians consider the proposed GUI system to be highly effective. The following comments were also recorded:

- During the review of a high number of chest radiographs obtained from regular health check-ups, the proposed GUI system can help avoid the manual overlooking of anomaly cardiac states.
- The proposed GUI system aids physicians without a major in cardiology to estimate PAWP based on chest radiographs.

As mentioned in the introduction, manual estimation of PAWP from chest radiographs is highly subjective. Therefore, the review of a high number of chest radiographs is time-consuming and anomalous cardiac states may easily be overlooked. Further, this method depends on the skills of individual physicians. The proposed GUI system ameliorates these shortcomings, and the positive responses of physicians in this article are attributed to this reason.

4. Conclusions

In this article, we reviewed the effectiveness of a GUI system that incorporates the explainable R-CNN developed by Saito et al. [5] using an effectiveness survey comprising 23 cardiology physicians with medical licenses in Japan. The results indicated that many physicians considered the proposed GUI system to be effective. We aim to introduce the proposed system into clinical medical diagnosis in accordance with the Japanese guidelines for artificial intelligence-based medical systems [9]. Additionally, verifying whether using the GUI system improves the diagnostic performance of physicians is very important. Therefore, we plan to examine this aspect in future work.

References

[1] M.M. Hoeper, S.H. Lee, R. Voswinckel, M. Palazzini, X. Jais, A. Marinelli, R.J. Barst, H.A. Ghofrani, Z.-C. Jing, C. Opitz, H.-J. Seyfarth, M. Halank, V. McLaughlin, R.J. Oudiz, R. Ewert, H. Wilkens, S. Kluge, H.-C. Bremer, E. Baroke, and L.J. Rubin, "Complications of right heart catheterization procedures in patients with pulmonary hypertension in experienced centers," Journal of the American College of Cardiology, vol.48, no.12, pp.2546–2552, Dec. 2006.

[2] T.A. McDonagh, M. Metra, M. Adamo, R.S. Gardner, A. Baumbach, and M. Böhm, "2021 ESC guidelines for the diagnosis and treatment of acute and chronic heart failureDeveloped by the Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC) With the special contribution," European Heart Journal, vol.42, no.36, pp.3599–3726, Sept. 2021.

[3] Y. Chen, E. Shlofmitz, N. Khalid, N.L. Bernardo, I. Ben-Dor, W.S. Weintraub, and R. Waksman, "Right heart catheterization-related complications: a review of the literature and best practices," Cardiology in Review, vol.28, no.1, pp.36–41, Jan. 2020.

[4] Y. Hirata, K. Kusunose, T. Tsuji, K. Fujimori, J. Kotoku, and M. Sata, "Deep learning for detection of elevated pulmonary artery wedge pressure using standard chest x-ray," Canadian Journal of Cardiology, vol.37, no.8, pp.1198–1206, Aug. 2021.

[5] Y. Saito, Y. Omae, D. Fukamachi, K. Nagashima, S. Mizobuchi, Y. Kakimoto, J. Toyotani, and Y. Okumura, "Quantitative estimation of pulmonary artery wedge pressure from chest radiographs by a regression convolutional neural network," Heart and Vessels, vol.37, no.8, pp.1387–1394, Feb. 2022.

[6] Z. Wang and J. Yang, "Diabetic retinopathy detection via deep convolutional networks for discriminative localization and visual explanation," Proceedings of the thirty-second AAAI conference on artificial intelligence, pp.514–521, 2018.

[7] A. Galdo, John A and Riggs, A Rickard and Morris, "Acute Decompensated Heart Failure," US Pharmacist, vol.38, no.2, pp.2–8, 2013.

[8] Y. Omae, Y. Saito, D. Fukamachi, K. Nagashima, Y. Okumura, and J. Toyotani, "Relationship between Regression Activation Map of R-CNN for Estimating Pulmonary Artery Wedge Pressure from Chest Radiographs and Generalization Score," IEICE Technical Report, vol.122, no.94, pp.52–54, 2022.

[9] Ministry of Economy Trade and Industry, "Medical Image Diagnosis Support System," 2019. https://www.meti.go.jp/policy/mono_info_service/healthcare/iryou/downloadfiles/pdf/47_guideline.pdf, accessed at July 12, 2022.