Computer Methods and Programs in Biomedicine 206 (2021) 106129

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

journal homepage: www.elsevier.com/locate/cmpb

Prediction and visualization of acute kidney injury in intensive care unit using one-dimensional convolutional neural networks based on routinely collected data

Noriaki Sato[a,b], Eiichiro Uchino[a,b], Ryosuke Kojima[a], Shusuke Hiragi[b,c], Motoko Yanagita[b,d], Yasushi Okuno[a,*]

[a] Department of Biomedical Data Intelligence, Graduate School of Medicine, Kyoto University, 54 Shogoin Kawahara-cho, Sakyo-ku, Kyoto 606-8507, Japan
[b] Department of Nephrology, Graduate School of Medicine, Kyoto University, Japan
[c] Division of Medical IT and Administration Planning, Kyoto University Hospital, Japan
[d] Institute for the Advanced Study of Human Biology (ASHBi), Kyoto University, Kyoto, Japan

ARTICLE INFO

Article history:
Received 11 November 2020
Accepted 21 April 2021

Keywords:
Acute kidney injury
Convolutional neural networks
Intensive care unit
Visualization

ABSTRACT

Background: Acute kidney injury (AKI) occurs frequently in in-hospital patients, especially in the intensive care unit (ICU), due to various etiologies including septic shock. It is clinically important to identify high-risk patients at an early stage and perform the appropriate intervention.

Methods: We proposed a system to predict AKI using one-dimensional convolutional neural networks (1D-CNN) with the real-time calculation of the probability of developing AKI, along with the visualization of the rationale behind prediction using score-weighted class activation mapping and guided back-propagation. The system was applied to predicting developing AKI based on the KDIGO guideline in time windows of 24 to 48 h using data of 0 to 24 h after admission to ICU.

Results: The comparison result of multiple algorithms modeling time series data indicated that the proposed 1D-CNN model achieved higher performance compared to the other models, with the mean area under the receiver operating characteristic curve of 0.742 ± 0.010 for predicting stage 1, and 0.844 ± 0.029 for stage 2 AKI using the input of the vital signs, the demographic information, and serum creatinine values. The visualization results suggested the reasonable interpretation that time points with higher respiratory rate, lower blood pressure, as well as lower SpO2, had higher attention in terms of predicting AKI, and thus important for prediction.

Conclusions: We presumed the proposed system's potential usefulness as it could be applied and transferred to almost any ICU setting that stored the time series data corresponding to vital signs.

1. Introduction

Acute kidney injury (AKI) refers to a pathological condition that shows a sudden decrease in the glomerular filtration rate, and it replaced acute renal failure (ARF) as a new concept including the more subtle decline in kidney function [1]. AKI occurs frequently in in-hospital patients, especially in the intensive care unit (ICU), due to various etiologies including septic shock. Patients who develop AKI have increased mortality [2]; therefore, it is clinically important to identify the high-risk patients with AKI in a timely manner and perform the appropriate intervention.

Recently, machine learning techniques have been widely applied to the prediction of various clinical events including AKI. In the ICU settings, the important data, such as vital signs are usually stored densely compared with a normal in-hospital ward in a timely manner. These data are directly related to the pathophysiologies of AKI, including intrarenal hemodynamic changes [3]. Regarding the application to these time series data, the machine learning algorithms, particularly, neural networks with various architectures including recurrent neural networks (RNN), long short-term memory (LSTM), and one-dimensional convolutional neural networks (1D-CNN) have been frequently implemented. Caicedo-

* Corresponding author.
E-mail address: okuno.yasushi.4c@kyoto-u.ac.jp (Y. Okuno).

https://doi.org/10.1016/j.cmpb.2021.106129

Torress et al. applied 1D-CNN to predict death in ICU using the public database of Medical Information Mart for Intensive Care-III and provided visual interpretation by DeepLIFT [4]. Regarding AKI, Tomasev et al. utilized RNN to predict the probability of AKI within the next 48 h at any time point and achieved the high performance that was evaluated by metrics of the area under receiver operating characteristic curve (AUROC) of 0.927 [5]. Additionally, the 1D-CNN models were applied to predict AKI, in the study focused on the usage of clinical notes [6]. However, applying 1D-CNN to the prediction of AKI based on the densely collected vital sign data, especially with the visual interpretation of reasons underlying the prediction results has been scarcely investigated. Additionally, the calculation of the real-time probability of developing AKI is an important task, particularly in the ICU setting, where clinical events occur in hourly intervals.

The present work addresses the challenges of predicting AKI by densely collected dataset which is routinely available in almost every ICU by 1D-CNN. Among the machine-learning algorithms, one advantage of using CNN is that there are established and standardized algorithms to visualize the reasons behind the prediction [7–9]. Therefore, we propose a system to predict AKI which enables physicians to understand what time point is important for the prediction of AKI in the real-time and retrospective manner by visualizing which time point is important for prediction, for use in the real-world setting.

2. Materials and methods

2.1. Patient selection

We used a publicly disclosed dataset, eICU Collaborative Research Database [10], with over 200,000 patients' data collected in the period from 2014 to 2015, in multiple centers across the United States. The data were made available in the PhysioNet repository.

First, the demographic information was collected. The age above 89 was recorded "over 90" for deidentification, and we replaced "over 90" with the value of 90. Gender was categorized as male or female. We obtained the information regarding whether a patient had chronic kidney disease (CKD) based on pastHistory and the diagnosis table. In the pastHistory table, those who had "renal", and no "dialysis" in "pasthistoryvalue" were regarded as ones having CKD. Additionally, based on the diagnosis table, those with the words "chronic" and "kidney" in "diagnosisstring" column were regarded as having CKD. The acute physiology and chronic health evaluation diagnoses upon admission were categorized in groups according to the code provided officially (apache-groups.sql). The patients with the code "ARF" were not included in the analysis, and the remaining 20 disease categories were considered as the input features in a binary format. These yielded a total of 24 features with the demographic information of age, gender, baseline creatinine, and the presence of CKD, apart from the time-series information.

Subsequently, we excluded the patients aged younger than 18 years, those who passed away in 0 to T hours after admission. T referred to the hour after the admission of which the input features used in training were obtained. We excluded those who received continuous renal replacement therapy in 0 to T hours after admission. This was defined by having "hemodialysis|for acute renal failure" in treatmentstring in the treatment table in 0 to T hours after admission. Additionally, we excluded those who stayed in ICU less than T hours. We excluded those on chronic hemodialysis or peritoneal dialysis. These conditions were defined based on pastHistory table if the word "dialysis" was in in "pasthistoryvalue". Additionally, from the treatment table, those who have treatmentstring of "hemodialysis" (not "hemodialysis|for acute renal failure") were excluded.

2.2. Problem setting and the time series input features

The main objective of the present study was to predict the onset of AKI within the time period from 24 to 48 h after admission to ICU, using the routinely collected data in ICU between 0 to 24 h after admission. We included the time series data of systolic, diastolic blood pressure (denoted as SBP and DBP respectively), heart rate (HR), respiratory rate (RR), body temperature (TEMP), and SpO2 as the routinely collected data, that are available in almost every patient in every ICU.

We obtained the vital sign information per minute. For the values taken in the same minute, the mean value was calculated. The missing data for each time point were imputed by the feed-forward method, which yielded 1,440 features per vital sign per patient ("downup" in fill function of tidyr library). Subsequently, the values were grouped according to the 15-minute interval, and the mean value was derived for each interval. These yielded 576 features (96 time series for six channels) for the vital signs of each patient. Additionally, the time series of serum creatinine values (CRE) were included from the laboratory data. The vital signs and CRE were scaled for the input of 1D-CNN, RNN, and LSTM using StandardScaler function in scikit-learn [11].

The training and test datasets were split according to the ratio of 8:2 for cross-validation, and the training dataset was additionally split to a 9:1 ratio to obtain the validation dataset while training of 1D-CNN, RNN, and LSTM. Training, test, and validation split of the dataset were the same for all the model training and evaluation steps.

Additionally, we performed the same workflow applied to the problem of predicting developing AKI in 48 to 72 h using the data from 0 to 48 h after admission to ICU.

2.3. AKI definition

Baseline creatinine was defined as the lowest value of those recorded before seven days until admission to ICU. AKI was defined as follows: (a) CRE increase of Δ0.3 mg/dL or above of the lowest creatinine value of the past 48 h, (b) CRE equal to or above the baseline creatinine values multiplied by 1.5, (c) CRE equal to or above the baseline creatinine values multiplied by 2, (d) CRE equal to or above the baseline creatinine values multiplied by 3 and (e) CRE equal to or above 4.0 mg/dL fulfilling the lowest value of past 48 h equal to or below 3.7 mg/d, or CRE equal to or above the baseline creatinine values multiplied by 1.5. The criteria for AKI used in the present study roughly corresponded to the KDIGO creatinine criteria [1]. Those with the baseline creatinine value of 4.0 mg/dL or above were excluded from the analysis.

The outcomes were defined as stage 1 or above (definition a to e, denoted as stage 1 hereafter), and stage 2 or above (definition c to e, denoted as stage 2 hereafter). AKI was calculated for every CRE record obtained. The onset of AKI was defined as when the criteria were met first, and if the patient had any time point with CRE as defined, the patient was categorized to AKI group. Subsequently, we included those who developed AKI in the defined period and did not develop AKI before the defined period as AKI group. The patients who did not develop AKI within the defined period were subsampled to the same number based on the remaining population with the fixed seed (non-AKI group).

2.4. Machine learning and used libraries

We used the machine learning algorithm of 1D-CNN to perform prediction. For classification performance comparison and assessment, we considered RNN, LSTM, and XGBoost as the baseline ap-

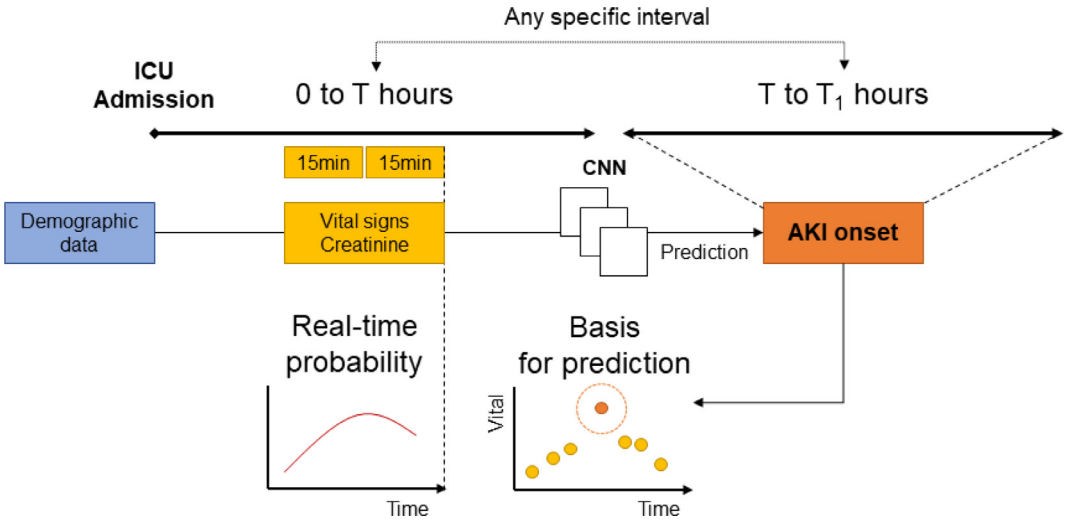

Fig. 1. The model overview.
The overview of the proposed workflow in the study.

proaches. We evaluated the performance of the considered models in terms of the area under the receiver operating characteristic curve (AUROC), with five-fold cross-validation.

The 1D-CNN model comprised one convolutional layer with the rectified linear activation function, followed by the batch normalization layer [12], a single global average pooling layer, and one dropout layer for each of creatinine and other vital signs. The final dense layer with the sigmoid function was used to make a prediction. The global average pooling layer was incorporated to enable the variable input length. The padding parameter was set to "causal", meaning that padding was performed in the manner that output of time point t does not depend on the input of time point $t+1$, which was suitable for modeling the temporal data. The RNN model had one RNN layer, and the LSTM model had one LSTM layer, along with the dropout and the dense layer. When the demographic information was incorporated into the 1D-CNN, RNN, and LSTM model, the corresponding input layer was concatenated to the output of a respective model, and the final output was derived using the sigmoid function. We used an Adam optimizer with the default setting for the optimization of parameters in all the models [12]. Binary cross entropy was calculated as a loss and optimized. The model was trained for 200 epochs, and weights of the best performance epoch regarding validation loss were saved.

For the input of RNN, LSTM, and XGBoost, the time series of vital signs with and without CRE and the combination with the demographic information was used. For the 1D-CNN models, an additional dataset including only CRE time-series was evaluated. Multiple filter numbers and kernel sizes were tested, and the best parameters regarding the 1D-CNN model with the time series with CRE and demographic information as input were chosen.

We used R library tidyverse for preprocessing of data [13]. We used machine learning and deep learning library scikit-learn and keras with tensorflow version 1.15 or 2.2.0 backend to construct neural networks [11,14,15]. We used Python library xgboost version 1.3.3 to construct XGBoost model, and the default parameter for xgboost training was used, and the objective parameter was set to "binary:logistic" [16]. The plotting and visualization were performed using ggplot2, ComplexHeatmap, firatheme, plotly, and patchwork [17–20]. When comparing patients' demographic information, one-way ANOVA was used for comparing continuous variables, and Chi-squared test was used for categorical variables.

2.5. Real-time prediction and retrospective visualization of the basis

To perform the real-time prediction of AKI, we predicted the probability of developing AKI at each 15 min interval. To achieve this, the input shape of the model was set to the corresponding time point, and the model was compiled, set weight of the best performance model and probability was predicted. Note that the weight of the best predictive model can be used for variable input, as the weight is dependent on filter number and kernel size, not on the input shape. Additionally, using the weight of the model in the fold with the best performance, we applied Score-weighted Class Activation Mapping (Score-CAM) to visualize the rationale behind the prediction in 1D-CNN, which was capable of interpreting which time point was important for prediction, for each patient retrospectively. The calculation of Score-CAM was described in the original paper [7]. Guided Score-CAM could be used to assess which channel was important for prediction and was calculated by multiplication of saliency map obtained by guided back propagation with Score-CAM values [9]. The absolute values of Guided Score-CAM were used as the score of the corresponding patient. We performed the analysis to show the validity of our approaches by the simulation of the intervention based on the real-time probability and guided Score-CAM. The detailed methods were described in Supplementary Text 1.

2.6. Availability of data and material

The source code of Python and R to reproduce the analysis can be found at https://doi.org/10.6084/m9.figshare.14555142.

3. Results

3.1. Predictive performance

The overall workflow is presented in Fig. 1. For the analysis predicting stage 2 AKI, 725 AKI patients and the same number of under-sampled patients were included in the analysis from the eICU database. The patient's demographic information is summarized in Table 1. There were significant differences between the non-AKI and AKI groups in age, ICU discharge offset and the baseline creatinine and the value was higher in AKI patients. Additionally, AKI patients had a higher frequency of having CKD. For the analysis including stage 1, 5342 patients were included. The

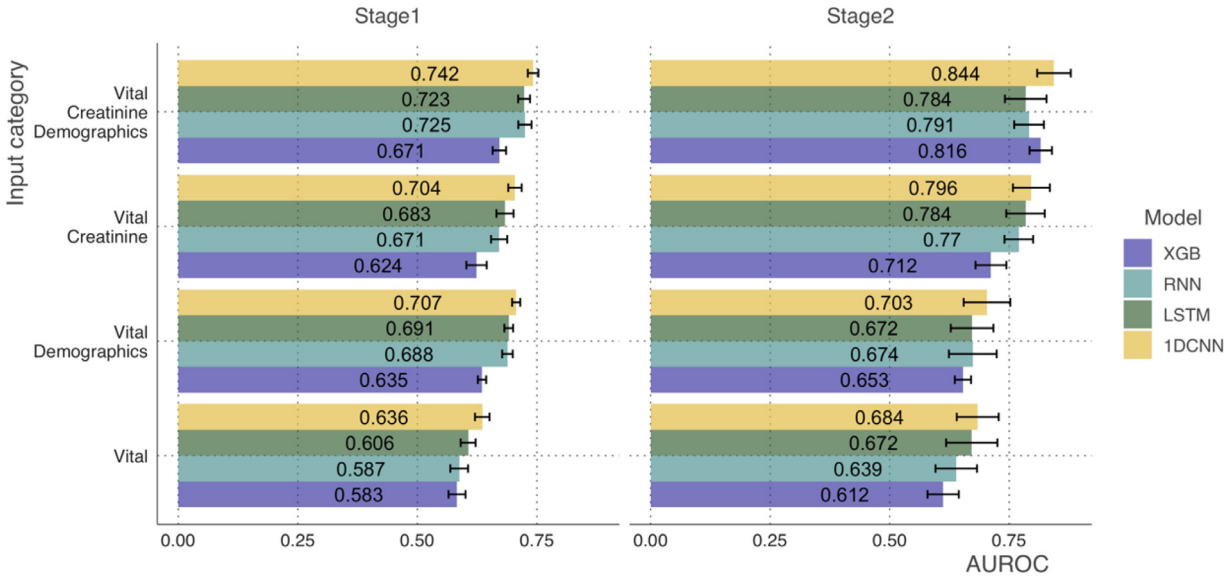

Fig. 2. The performance summary of prediction.
The summary of the performance for each model. The bar plot represents the mean values of five-fold cross-validation and the standard deviations. The color represents the model. XGB: XGBoost, RNN: recurrent neural network, LSTM: long short-term memory, 1DCNN: one-dimensional convolutional neural network.

Table 1
The demographics of patients included in the analysis.

Clinical Values	non-AKI (N = 725)	AKI (N = 725)	p-value
Age (mean (SD))	62.77 (16.55)	65.51 (14.44)	0.001
Gender = male (%)	383 (52.8)	417 (57.5)	0.081
Unit discharge offset (mean (SD))	5994.38 (6066.93)	8341.59 (7924.58)	<0.001
Baseline creatinine value (mean (SD))	1.18 (0.61)	1.36 (0.88)	<0.001
CKD present (%)	73 (10.1)	123 (17.0)	<0.001

Unit discharge offset: the timing of leaving ICU, presented as the minutes after admission.

Table 2
The number of the valid value for each feature.

VS	mean	SD	Max	min
HR	32.77	24.94	191	1
RR	31.08	24.19	189	1
SYS	25.14	20.45	179	1
DIA	25.14	20.45	179	1
SpO2	31.6	24.59	190	1
TEMP	9.7	12.68	126	1
CRE	1.66	1.17	12	1

VS: vital signs, SD: standard deviation, HR: heart rate, RR: respiratory rate, SYS: systolic blood pressure, DIA: diastolic blood pressure, TEMP: body temperature, CRE: serum creatinine value.

Table 3
The performance summary of each model.

Model	Stage1_ave	Stage1_std	Stage2_ave	Stage2_std	Input
1DCNN	0.636	0.012	0.684	0.036	VS
LSTM	0.606	0.013	0.672	0.044	VS
RNN	0.587	0.015	0.639	0.035	VS
XGB	0.583	0.014	0.612	0.027	VS
1DCNN	0.704	0.011	0.796	0.032	VS, CRE
LSTM	0.683	0.015	0.784	0.033	VS, CRE
RNN	0.671	0.014	0.77	0.024	VS, CRE
XGB	0.624	0.017	0.712	0.026	VS, CRE
1DCNN	0.742	0.009	0.844	0.029	VS, CRE, DEMOG
LSTM	0.723	0.01	0.784	0.035	VS, CRE, DEMOG
RNN	0.725	0.011	0.791	0.025	VS, CRE, DEMOG
XGB	0.671	0.011	0.816	0.019	VS, CRE, DEMOG
1DCNN	0.707	0.007	0.703	0.04	VS, DEMOG
LSTM	0.691	0.007	0.672	0.036	VS, DEMOG
RNN	0.688	0.009	0.674	0.041	VS, DEMOG
XGB	0.635	0.007	0.653	0.014	VS, DEMOG

Stage1_ave, Stage2_ave, Stage1_std, and Stage2_std: the average AUROC values and the standard deviation for prediction of stage 1 and 2 AKI, VS: vital signs, CRE: serum creatinine values, DEMOG: demographic information, RNN: recurrent neural network, LSTM: long short-term memory, 1DCNN: one-dimensional convolutional neural network, XGB: XGBoost.

patient's demographic information is summarized in Supplementary Table 1. The mean number of the valid values for each feature before imputation for all the patients included in the analysis of stages 1 and 2 is summarized in Table 2. CRE had a lower number of valid values compared to the other features. The performance of each classifier for the prediction of stages 1 and 2 is summarized in Table 3 and Fig. 2. Overall, the highest performance was obtained when the vital signs, demographic information, and CRE were combined (AUROC 0.742 ± 0.010 for stage 1 and 0.844 ± 0.029 for stage 2, mean ± standard deviation), compared to the other models. When only the CRE were used as input for prediction of stage 2 AKI in 24 to 48 h, the performance was lower than combined with CRE combined with vital signs as input (AUROC 0.759 ± 0.034 and 0.796 ± 0.032 respectively).

We additionally applied the same workflow for predicting developing AKI in 48 to 72 h using the features of 0 to 48 h after admission to ICU. For the setting, as same as the prediction of developing AKI in 24 to 48 h, the higher performance was obtained in 1D-CNN compared to the other approaches in both predictions of stage 1 and 2, when vital signs, demographic information, and CRE were combined (AUROC 0.698 ± 0.012 and 0.860 ± 0.026 for predictions of stage 1 and 2). The results for the prediction of de-

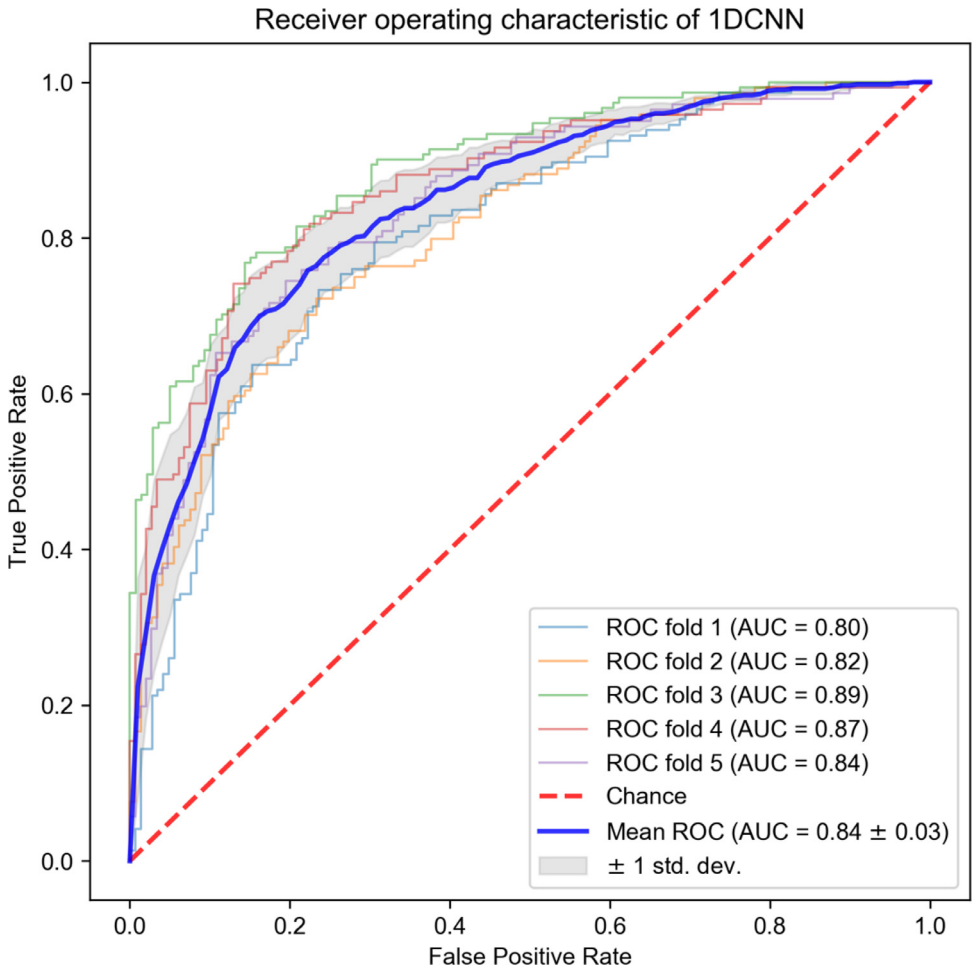

Fig. 3. The ROC curves of the 1D-CNN model.
The ROC curves of the 1D-CNN with the input of vital signs, demographic information, and creatinine values for all the folds in the cross-validation.

veloping AKI in 48 to 72 h were summarized in Supplementary Table 2.

We used the model of predicting the stage 2 AKI in 24 to 48 h for the downstream visualization analysis, with the weights in the best fold in cross-validation, which had the AUROC value of 0.890 (Fig. 3).

3.2. Real-time prediction and visualization

We calculated the real-time probability, Score-CAM and guided Score-CAM. Those with the highest three probability of developing AKI in the test dataset were visualized in Fig. 4 (the patient one to three hereafter). Note that all three patients visualized in the fold developed AKI in subsequent 24 to 48 h period and were true positives.

Specifically, for patient one, the average guided Score-CAM was high in SBP and DBP, and the time interval of lower SBP and DBP got higher attention, suggesting the patient developed AKI because of these vital signs. Additionally, higher RR and lower SpO2 got high attention. For patient two, The average guided Score-CAM was high in SpO2 and TEMP, and the time points with lower SpO2, as well as higher TEMP got high attention. Additionally, higher RR got high attention. For patient three, the higher RR and TEMP also got high attention, and the time interval with lower SpO2 and higher TEMP got high attention. For all the patients, the probability of developing AKI was continuously high from admission to ICU.

We showed one example of the patients who had a gradual increase in the probability of developing AKI in Supplementary Figure 1. This patient had high attention in SpO2, TEMP and RR. For this patient, we performed the additional analysis of the simulation of the intervention to show the validity of our proposed framework (Supplementary Figure 2). On all the experiments changing vital signs based on the real-time probability and guided Score-CAM, the final probability of developing AKI was low compared to the raw input.

Additionally, The visualization of the patients with the highest three probability in the fold with the second-highest performance (AUROC 0.866) is in Supplementary Figure 3 for the referencing purpose.

4. Discussion

In the present study, we applied the 1D-CNN model to the problem of prediction of AKI in ICU and proposed the system to help physicians understand which time point could be important for prediction in the real-time and retrospective manner. The obtained results indicated that although the performance of the proposed system was low compared to other studies, the real-time prediction and retrospective visualization of prediction were reasonable.

Visualization result of guided Score-CAM suggested that time points with lower SpO2 and blood pressure, and higher TEMP and RR had high attention and thus important for the prediction of AKI. One of the suspected reasons is that these vital signs are among the criteria for sepsis or septic shock [21], which is one of the important and common reasons for AKI in ICU and thus are consid-

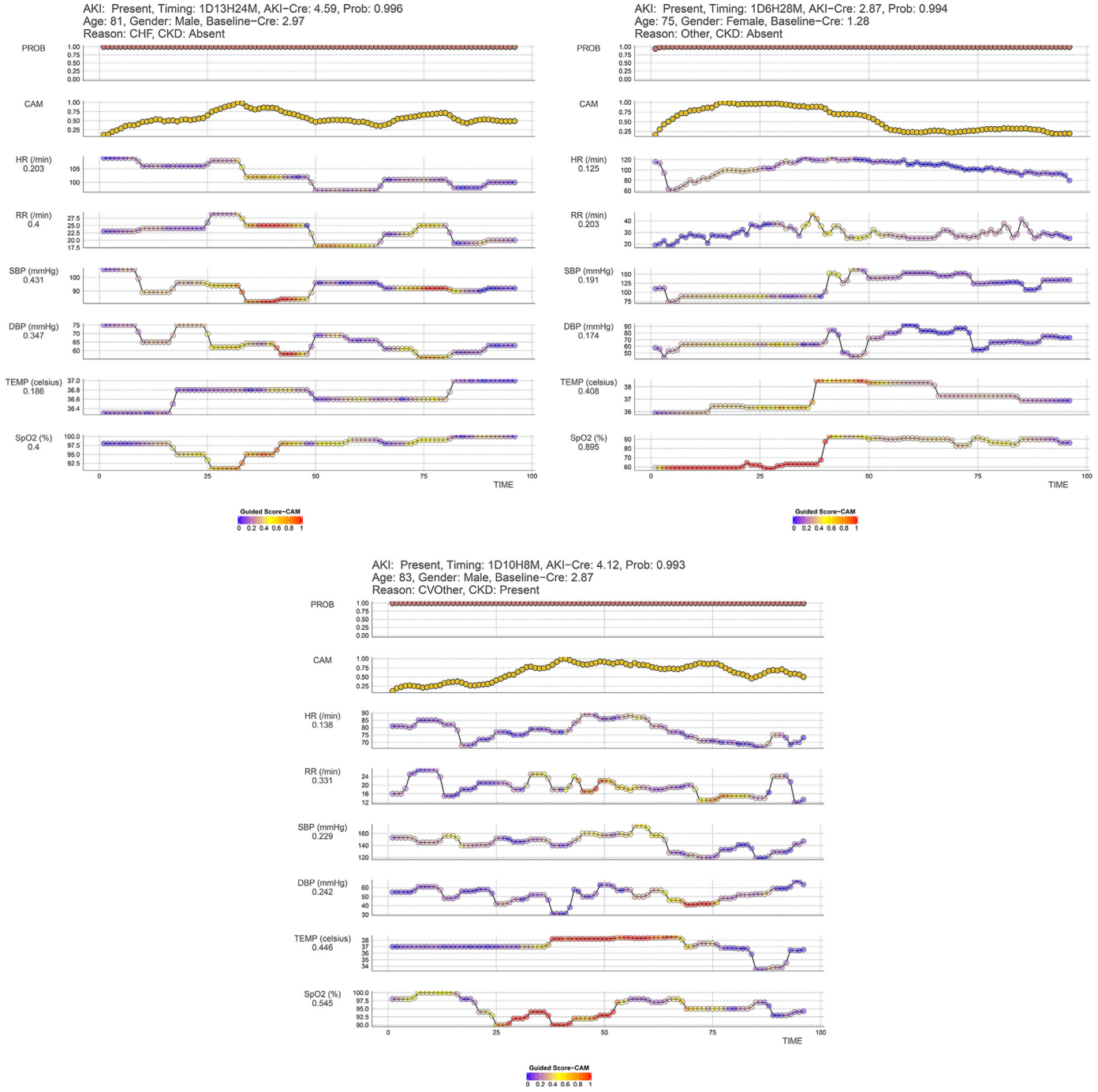

Fig. 4. Visualization of the rationale behind the prediction of developing AKI.
Information of patients with the highest three probability of developing AKI were listed. The *x*-axis showed a time point of 15 min interval. The line plot of PROB indicates the probability of developing AKI at the specific time point and CAM indicates Score-CAM values. HR, RR, SBP, DBP, TEMP, and SpO2 indicate the recorded value at the specific time point, and the color of the point indicates guided Score-CAM value. The values under each label in the *y*-axis represent the average value of guided Score-CAM for each vital sign.

ered to have high attention. Besides, cardiac arrest, in which vital signs show drastic changes, has been reported to cause AKI [22]. In this context, this suggested that the 1D-CNN algorithm is capable of correctly identify which time point has important information regarding the prediction of AKI. Additionally, the important vital signs for the prediction differ from patient to patient.

We compared the performance of 1D-CNN with those of XGBoost, RNN, and LSTM. When compared to RNN and LSTM, 1D-CNN achieved slightly better performance in the current setting. In one study, it was reported that CNN, or temporal CNN, which have the characteristic of architecture are causal, and that the architecture can take a variable length of the input, have superior properties compared with the RNN architecture, such as LSTM in the sequencing modeling task [23]. Additionally, CNN generally requires low memory and time to train compared to RNN or LSTM, which could be advantageous where the computational resource is scarce.

The major limitation of the present study was the low performance compared with the other study using the full set of features available in ICU [5]. The main reason underlying this was suspected

to be the selection of a time window, the usage of strict AKI criteria in terms of occurrence in a limited time window, and the usage of limited features, especially not using information regarding medications used in ICU, which is used in the other study [24], as medications can cause AKI in high frequency [25]. Although 1D-CNN has the advantage of the capacity to visualize the rationale behind the prediction and the capability of real-time prediction, other models could achieve high performance in the present comparison. Besides, in the situations like the number of valid input values before imputation was low, the algorithms could identify the highest or lowest values as important, which did not use time-series information sufficiently. As our model used CRE as the input, and the AKI category was defined using CRE, CRE seemed to play the major role in prediction. However, we showed that combining the easily accessible vital signs and demographic information with CRE can improve the predictive performance of developing future AKI to a certain extent. Additionally, although making changes in vital signs is clinically not feasible and impossible, we showed that the real-time calculation of probability and the calculation of the basis behind prediction by the proposed system could aid in reducing the probability of developing AKI in the present setting. The performance assessed by AUROC was lower in the prediction of stage 1 compared to stage 2 AKI. It was presumed that stage 1 included more subtle changes and a limited dataset that is routinely available did not have sufficient information to predict stage 1 AKI.

The approach proposed in the present study could be advantageous, as it could be applied and transferred to almost any ICU setting that stored the time series data of vital signs, without the need for the identification and linking of medication codes and translation of clinical notes.

5. Conclusions

In conclusion, we developed a prediction and visualization system of AKI using 1D-CNN and Score-CAM. The proposed system was aimed to help physicians involved in ICU to understand how the patient developed AKI retrospectively, and additionally realize the probability of developing AKI in a real-time manner.

Declaration of Competing Interest

Dr. Yanagita reports grants from AMED (Project for Elucidating and Controlling Mechanisms of Aging and Longevity), grants from AMED-CREST (Innovation for Ideal Medical Treatment Based on the Understanding of Maintenance, Change and Breakdown Mechanisms of Homeostasis among Interacting Organ Systems), grants from Grant-in-Aid for Scientific Research B (from 2017), grants from Grant-in-Aid for Scientific Research on Innovative Areas (from 2017), grants from Grant-in-Aid for Challenging Research (Exploratory), grants from Grant-in-Aid for Scientific Research on Innovative Areas (from 2018), grants from AMED (Platform Program for Promotion of Genome Medicine), grants from Grant-in-Aid for Scientific Research B (from 2020), outside the submitted work. The other authors declare no conflicts of interest.

CRediT authorship contribution statement

Noriaki Sato: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Visualization, Writing - original draft. **Eiichiro Uchino:** Methodology, Supervision, Writing - review & editing. **Ryosuke Kojima:** Methodology, Supervision, Writing - review & editing. **Shusuke Hiragi:** Supervision, Writing - review & editing. **Motoko Yanagita:** Supervision, Writing - review & editing. **Yasushi Okuno:** Conceptualization, Funding acquisition, Methodology, Project administration, Writing - review & editing.

Acknowledgments

This work was supported by JSPS KAKENHI Grant-in-Aid for Early-Career Scientists Grant Number 19K18321. The funding body had no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript. The preliminary results of the current study were presented at the American Society of Nephrology Kidney Week 2019 (November 7, 2019).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.cmpb.2021.106129.

References

[1] Kidney Disease: Improving Global Outcomes (KDIGO) Acute Kidney Injury Work Group, KDIGO clinical practice guideline for acute kidney injury, Kidney Int. Suppl. 2 (2012) 6.
[2] K. Doi, K. Negishi, T. Ishizu, D. Katagiri, T. Fujita, T. Matsubara, N. Yahagi, T. Sugaya, E. Noiri, Evaluation of new acute kidney injury biomarkers in a mixed intensive care unit, Crit. Care Med. 39 (2011) 2464–2469.
[3] L. Wan, S.M. Bagshaw, C. Langenberg, T. Saotome, C. May, R. Bellomo, Pathophysiology of septic acute kidney injury: what do we really know? Crit. Care Med. 36 (2008) S198–S203.
[4] W. Caicedo-Torres, J. Gutierrez, ISeeU: visually interpretable deep learning for mortality prediction inside the ICU, J. Biomed. Inform. 98 (2019) 103269.
[5] N. Tomašev, X. Glorot, J.W. Rae, M. Zielinski, H. Askham, A. Saraiva, A. Mottram, C. Meyer, S. Ravuri, I. Protsyuk, A. Connell, C.O. Hughes, A. Karthikesalingam, J. Cornebise, H. Montgomery, G. Rees, C. Laing, C.R. Baker, K. Peterson, R. Reeves, D. Hassabis, D. King, M. Suleyman, T. Back, C. Nielson, J.R. Ledsam, S. Mohamed, A clinically applicable approach to continuous prediction of future acute kidney injury, Nature 572 (2019) 116–119.
[6] M. Sun, J. Baron, A. Dighe, P. Szolovits, R.G. Wunderink, T. Isakova, Y. Luo, Early prediction of acute kidney injury in critical care setting using clinical notes and structured multivariate physiological measurements, Stud. Health Technol. Inform. 264 (2019) 368–372.
[7] H. Wang, Z. Wang, M. Du, F. Yang, Z. Zhang, S. Ding, P. Mardziel, X. Hu, Score–CAM: score-weighted visual explanations for convolutional neural networks, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020, pp. 24–25.
[8] R.R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, D. Batra, Grad–CAM: visual explanations from deep networks via gradient-based localization, in: 2017 IEEE International Conference on Computer Vision (ICCV), 2017, pp. 618–626.
[9] J.T. Springenberg, A. Dosovitskiy, T. Brox, M. Riedmiller, Striving for Simplicity: The All Convolutional Net, ArXiv [Cs.LG]. (2014). http://arxiv.org/abs/1412.6806.
[10] T.J. Pollard, A.E.W. Johnson, J.D. Raffa, L.A. Celi, R.G. Mark, O. Badawi, The eICU collaborative research database, a freely available multi-center database for critical care research, Sci. Data 5 (2018) 180178.
[11] P. Fabian, V. Gaël, G. Alexandre, M. Vincent, T. Bertrand, G. Olivier, B. Mathieu, P. Peter, W. Ron, D. Vincent, V. Jake, P. Alexandre, C. David, B. Matthieu, P. Matthieu, D. Édouard, Scikit-learn: machine learning in python, J. Mach. Learn. Res. (2011) https://dl.acm.org/doi/10.5555/1953048.2078195. (Accessed February 13, 2020).
[12] S. Ioffe, C. Szegedy, Batch normalization: accelerating deep network training by reducing internal covariate shift 37 (2015) 448–456. JMLR.org448
[13] H. Wickham, M. Averick, J. Bryan, W. Chang, L.D. McGowan, R. François, G. Grolemund, A. Hayes, L. Henry, J. Hester, M. Kuhn, T.L. Pedersen, E. Miller, S.M. Bache, K. Müller, J. Ooms, D. Robinson, D.P. Seidel, V. Spinu, K. Takahashi, D. Vaughan, C. Wilke, K. Woo, H. Yutani, Welcome to the tidyverse, J. Open Source Softw. 4 (2019) 1686, doi:10.21105/joss.01686.
[14] F. Chollet, Others, Keras, (2015). https://keras.io.
[15] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G.S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, X. Zheng, TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems, (2015). http://tensorflow.org/.
[16] T. Chen, C. Guestrin, XGBoost: a scalable tree boosting system, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Association for Computing Machinery, New York, NY, USA, 2016, pp. 785–794.
[17] H. Wickham, ggplot2: Elegant Graphics for Data Analysis, (2016). https://ggplot2.tidyverse.org.
[18] Z. Gu, R. Eils, M. Schlesner, Complex heatmaps reveal patterns and correlations in multidimensional genomic data, Bioinformatics 32 (2016) 2847–2849.

[19] E.-J. van Kesteren, vankesteren/firatheme: firatheme version 0.2.1, 2020. https://doi.org/10.5281/zenodo.3604681.
[20] P.T. Inc, Collaborative data science, Montréal, QC. (2015). https://plot.ly.
[21] M. Singer, C.S. Deutschman, C.W. Seymour, M. Shankar-Hari, D. Annane, M. Bauer, R. Bellomo, G.R. Bernard, J.-D. Chiche, C.M. Coopersmith, R.S. Hotchkiss, M.M. Levy, J.C. Marshall, G.S. Martin, S.M. Opal, G.D. Rubenfeld, T. van der Poll, J.-L. Vincent, D.C. Angus, The third international consensus definitions for sepsis and septic shock (sepsis-3), JAMA 315 (2016) 801–810.
[22] M. Rundgren, S. Ullén, M.P.G. Morgan, G. Glover, J. Cranshaw, N. Al-Subaie, A. Walden, M. Joannidis, M. Ostermann, J. Dankiewicz, N. Nielsen, M.P. Wise, Renal function after out-of-hospital cardiac arrest; the influence of temperature management and coronary angiography, a post hoc study of the target temperature management trial, Crit. Care 23 (2019) 163.
[23] S. Bai, J. Zico Kolter, V. Koltun, An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling, ArXiv [Cs.LG]. (2018). https://arxiv.org/abs/1803.01271.
[24] J.L. Koyner, K.A. Carey, D.P. Edelson, M.M. Churpek, The development of a machine learning inpatient acute kidney injury prediction model, Crit. Care Med. 46 (2018) 1070–1077.
[25] S. Ehrmann, J. Helms, A. Joret, L. Martin-Lefevre, J.-P. Quenot, J.-E. Herbrecht, D. Benzekri-Lefevre, R. Robert, A. Desachy, F. Bellec, G. Plantefeve, A. Bretagnol, A. Dargent, J.-C. Lacherade, F. Meziani, B. Giraudeau, E. Tavernier, P.-F. Dequin, Clinical research in intensive care and sepsis-trial group for global evaluation and research in sepsis (CRICS-TRIGGERSEP network), nephrotoxic drug burden among 1001 critically ill patients: impact on acute kidney injury, Ann. Intensive Care 9 (2019) 106.