Improved generalized performance of Hemodynamics Scenarios prediction with digital biomarkers by Conv1D approach

Momo Shirotori
Digital Strategy Department, Digital Transformation Unit
Chugai Pharmaceutical Co., Ltd.
Tokyo, Japan
shirotori.momo60@chugai-pharm.co.jp

Kato Kosuke
Japan Data Science Consortium Co. Ltd.
Tokyo, Japan
kosuke.kato@jdsc.ai

Takaya Hondo
Japan Data Science Consortium Co. Ltd.
Tokyo, Japan
takaya.hondo@jdsc.ai

Kijun Kim
Japan Data Science Consortium Co. Ltd.
Tokyo, Japan
kijun.kim@jdsc.ai

Kenton Tokuyama
Digital Strategy Department, Digital Transformation Unit
Chugai Pharmaceutical Co., Ltd.
Tokyo, Japan
tokuyama.kento26@chugai-pharm.co.jp

***Abstract*—Digital biomarkers (dB) provide valuable information for the continuous assessment of disease status in clinical practice. Hemodynamics is an important endpoint for evaluating a patient's status and, therefore, a continuous monitoring system using dB needs to be developed. In this study, we developed supervised learning modeling approaches for estimating hemodynamic scenarios of new patients using a time series dataset obtained from contact and contact-free sensors.**

Keywords—Machine learning, digital biomarker, wearable device, hemodynamics, multivariate time series classification

I. Introduction

Monitoring the health and disease status of a patient is important for understanding disease mechanisms. The recent development of wearable devices has enabled us to measure vital data continuously and remotely, and such data are referred to as digital biomarkers (dpmo) [1]. Some recent studies developed estimation models for a patient's condition or clinical outcome assessment using the measured values of dBM values. The developed estimators need to achieve a generalized performance for new patient prediction to use predictions in clinical practice. However, it has the challenge of being less accurate than predicting behavior in a new period within the same patient. In this study, we proposed the supervised-learning models and evaluated the generalized performance for hemodynamics scenario prediction for new patients using time series data obtained from contact and contact-free sensors.

II. Related Work

Slapničar et al. [2] reported the branched artificial neural network architecture to predict hemodynamics scenarios for the random CV. However, the evaluation of the other machine learning modeling approach was insufficient. Therefore, we engineered estimators using several machine learning techniques and evaluated their prediction accuracy in person-out CV.

III. Classifier Development for Hemodynamics Scenarios

We used a public dataset [3] obtained from previous research[1]. The dataset was collected from 30 healthy volunteers, and the subjects performed five different hemodynamic scenarios. This dataset contains bio-signal data measured simultaneously using contact sensors and contact-free sensors, which have 12 channels (i.e., features). We considered a multi-class classification problem that predicts five types of hemodynamic scenarios. For data preprocessing, we divided the signals into window lengths of 20 s with a 10 s overlap for each of the 12 channels of all subjects. Then, we normalized the values to obtain zero mean and unity variance. To evaluate the generalized performance, we used the person-out CV approach. Fig. 1 shows examples of dataset splitting for random and person-out CVs. The random CV that the previous study [2] used overestimates the performance of new patient prediction. This is because the randomly split data of the same subject can be included in both training and test datasets. The person-out CV approach can accurately evaluate the generalized performance because it splits the dataset into training and test datasets by each subject.

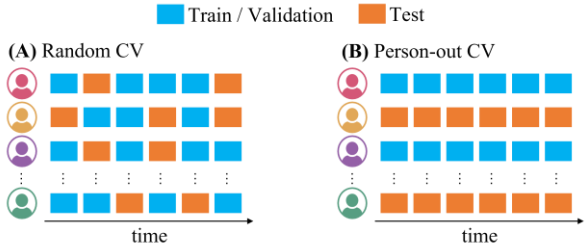

Fig. 1. Example of train/validation and test dataset splitting. (A) Random CV and (B) Person-out CV.

[1] https://figshare.com/articles/dataset/A_dataset_of_clinically_recorded_radar_vital_signs_with_synchronised_reference_sensor_signals/12186516/2 (accessed November 10, 2022).

IV. Experimental Results

We developed the 1D convolutional neural network models with adaptive average pooling (Conv1D), the bidirectional LSTM (Conv1D-BiLSTM), and the transformer (Conv1D-Transformer). We compared the MLP architecture of the previous study [2] with our developed three deep learning models and the LightGBM. For the LightGBM, we performed 8 types of feature engineering (Max, Min, Differences, Mean, Std, SAD, Peak_count, Peak_rrstd) for all channels. In deep learning model development, we performed the grid search of those hyperparameters shown in Table 1. We optimized hyperparameters by running the five-fold CV and calculated the average of performance measures: micro-accuracy, macro-recall, macro-precision, macro-F1, and macro-AUC. We set the learning rate to 0.0001, epoch size to 50, and batch size to 64 for all deep learning models.

Fig. 2 shows the architectures of the deep learning models when the macro-AUC was the best for person-out CV. Tables 2 and 3 give the average performance measures for random and person-out CVs, and the bold font denotes the best scores. The LightGBM (N_FEATURES =96) and the Conv1D were the best macro-AUC performance for random CV (0.979) and person-out CV (0.936), respectively. The accuracy of the person-out CV was poorer than that of the random CV in all models. In particular, the MLP model strongly decreased prediction accuracy in the person-out CV compared with the random CV. On the other hand, the performance of our models using the 1D convolutional neural network had only slightly different among CV types. These results suggested that the information of individual sensor differences can be considered more effectively by convolution with a filter between each input channel.

TABLE I. The Set of Hyperparameters We Invetigated

Hyperparameters	***Investigated values***
N_FEATURES (LightGBM)	[96, 72, 48, 24, 12]
N_HIDDEN_LAYER (LSTM)	[12, 24, 36, 48, 64, 78, 96, 108, 120]
N_LAYERS (TRANSFOMER)	[2, 4, 6, 8, 10, 12, 14]
CONV 1D_HIDDEN_CHANNEL	[[12, 12], [24, 12], [12, 24], [24, 24], [48, 12], [48, 24], [12, 48], [24, 48], [48, 48]]
ACTIVATION	[ReLU, Tanh, GELU]
NORM	[BatchNorm, LayerNorm]
NORM_IN_FIRST_CONV_BLOCK	[True, False]

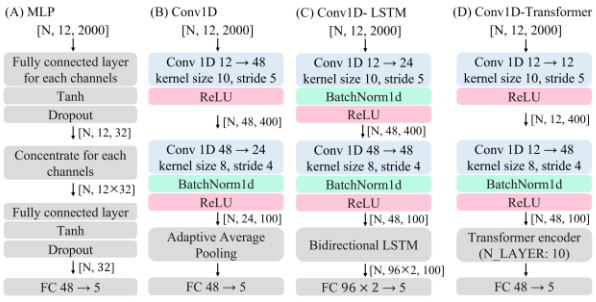

Fig. 2. The best model architectures for person-out CV. (A) MLP, (B) Conv1D, (C) Conv1D-BiLSTM, and (D) Conv1D-Transformer.

TABLE II. Mean Performance Measures For Random CV

	Micro-accuracy	***Macro-recall***	***Macro-precision***	***Macro-F1***	***Macro-AUC***
MLP	0.795	0.795	0.793	0.787	0.944
LightGBM	0.860	**0.876**	**0.887**	**0.874**	**0.979**
Conv1D	0.860	0.860	0.884	0.862	0.970
Conv1D-BiLSTM	0.856	0.856	0.880	0.858	0.974
Conv1D-Transformer	**0.869**	0.869	0.883	0.870	0.978

TABLE III. Mean Performance Measures For Person-out CV

	Micro-accuracy	***Macro-recall***	***Macro-precision***	***Macro-F1***	***Macro-AUC***
MLP	0.436	0.214	0.217	0.213	0.535
LightGBM	0.709	0.617	**0.700**	0.631	0.902
Conv1D	0.745	**0.788**	0.644	0.671	**0.936**
Conv1D-BiLSTM	0.771	0.763	0.677	**0.695**	0.930
Conv1D-Transformer	**0.747**	0.770	0.632	0.653	0.935

V. Conclusions

We aimed to develop a supervised-learning model for hemodynamics scenario prediction for new patients using time series data obtained from contact and contact-free sensors. We proposed three deep learning architectures using the 1D convolutional neural network models and compared them with the conventional MLP model to evaluate the generalized performance for random and person-out CVs. The experimental results for person-out CV showed that the Conv1D was the best performance and the MLP model strongly decreased prediction accuracy compared with random CV. We revealed that the information of individual differences of sensors can be considered more effectively by convolution with a filter between each input channel using Conv1D. This study provides valuable insights into the development of deep neural network models incorporating dpmo, which in turn contributes to the enhancement of patient health monitoring and a deeper understanding of disease mechanisms.

References

[1] S. Vasudevan, A. Saha, M. E. Tarver, and B. Patel, "Digital biomarkers: Convergence of digital health technologies and biomarkers," NPJ Digit. Med., Vol. 5, p.36, 2022.

[2] G. Slapničar, W. Wang, and M. Luštrek, "Classification of hemodynamics dcenarios from a public rradar dataset using a deep learning approach," Sensors, Vol. 21, No. 5, p. 1836, 2021.

[3] S. Schellenberger, K. Shi, T. Steigleder, A. Malessa, F. Michler, L. Hameyer, N. Neumann, F. Lurz, R. Weigel, C. Ostgathe, and A. Koelpin, "A dataset of clinically recorded radar vital signs with synchronised reference sensor signals," Sci. Data Vol. 7, p.291, 2020.

[4] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T. Y. Liu, "LightGBM: A highly efficient gradient boosting decision tree," Adv. Neural Inf. Process. Syst. Vol. 30, pp. 3146–3154, 2017.