Threshold	Precision	Recall	F1 score
82	0.21	25.76%	0.0042
88	0.23	20.45%	0.0045
96	0.30	15.15%	0.0058
99	0.28	9.85%	0.0054
100	0.34	5.30%	0.0064

Table 9 shows the results of predicting cardiac arrest within 24 hours as the threshold for the self-attention-based LSTM model. In the case of threshold 2, it has the best recall in the self-attention-based LSTM model, and there is no significant difference in precision. We discovered that the GRU model performs best when the models are evaluated with a default threshold of 76 for predicting cardiac arrest within 24 hours, and it performs best when the optimal threshold value is set for each model.

TABLE IX
PREDICTION OF CARDIAC ARREST WITHIN 24 HOURS BASED ON SELF-ATTENTION-BASED LSTM MODEL

Threshold	Precision	Recall	F1 score
1	0.27%	69.70%	0.0053
6	0.27%	65.15%	0.0053
21	0.26%	55.30%	0.0051
29	0.25%	50%	0.0050
51	0.31%	45.45%	0.0061
57	0.30%	40.15%	0.0059
68	0.28%	35.61%	0.0063
76	0.31%	28.79%	0.0060
81	0.30%	23.48%	0.0060
90	0.38%	17.42%	0.0075
96	0.40%	12.12%	0.0077
100	0.35%	3.79%	0.0064

Table 10 shows the results of predicting cardiac arrest within 72 hours as the threshold for the LSTM model. In the case of threshold 1, it has the best recall in the LSTM model, and there is no significant difference in precision.

TABLE X
PREDICTION OF CARDIAC ARREST WITHIN 72 HOURS BASED ON LSTM MODEL

Threshold	Precision	Recall	F1 score
1	0.73%	82.41%	0.0145
6	0.73%	77.47%	0.0146
16	0.73%	70.37%	0.0144
28	0.74%	66.67%	0.0147
39	0.76%	61.42%	0.0152
46	0.81%	55.25%	0.0159
50	0.85%	50.31%	0.0167
54	0.82%	45.06%	0.0160
59	0.81%	40.74%	0.0159
66	0.82%	35.19%	0.0160
70	0.79%	29.63%	0.0153
74	0.74%	24.69%	0.0143
78	0.63%	19.14%	0.0123
83	0.62%	15.12%	0.0118
95	0.79%	11.11%	0.0148
98	0.60%	6.79%	0.0110
100	0.60%	2.47%	0.0097

Table 11 shows the results of predicting cardiac arrest within 72 hours as the threshold for the GRU model. In the case of threshold 1, it has the best recall in the GRU model, and there is no significant difference in precision.

TABLE XI
PREDICTION OF CARDIAC ARREST WITHIN 72 HOURS BASED ON GRU MODEL

Threshold	Precision	Recall	F1 score
1	0.69%	94.44%	0.0136
4	0.68%	90.74%	0.0136
6	0.66%	85.80%	0.0131
10	0.64%	80.56%	0.0128
18	0.64%	75.62%	0.0126
28	0.65%	72.84%	0.0130
38	0.65%	66.67%	0.0128
44	0.62%	59.88%	0.0122
49	0.62%	55.25%	0.0122
53	0.61%	50.31%	0.0122
59	0.60%	44.75%	0.0118
66	0.57%	39.81%	0.0112
72	0.55%	35.19%	0.0108
80	0.58%	31.17%	0.0113
84	0.56%	25%	0.0110
89	0.60%	20.68%	0.0117
93	0.61%	15.74%	0.0117
98	0.62%	10.80%	0.0118
100	0.83%	5.25%	0.0144

Table 12 shows the results of predicting cardiac arrest within 72 hours as the threshold for the self-attention-based LSTM model. In the case of threshold 1, it has the best recall in the self-attention-based LSTM model, and there is no significant difference in precision. We discovered that the GRU model performs best when the models are evaluated with a default threshold of 28 for predicting cardiac arrest within 72 hours, and it performs best when the optimal threshold value is set for each model.

TABLE XII
PREDICTION OF CARDIAC ARREST WITHIN 72 HOURS BASED ON SELF-ATTENTION-BASED LSTM MODEL

Threshold	Precision	Recall	F1 score
1	0.66%	70.99%	0.0131
11	0.69%	66.05%	0.0137
28	0.66%	54.01%	0.0131
45	0.76%	50.31%	0.0150
50	0.73%	44.75%	0.0144
55	0.70%	39.81%	0.0138
68	0.77%	35.19%	0.0150
74	0.76%	30.56%	0.0148
80	0.78%	25.93%	0.0151
83	0.77%	21.30%	0.148
90	0.85%	15.74%	0.0161
95	0.77%	10.19%	0.0144
99	0.77%	6.48%	0.0138
100	0.84%	3.70%	0.0137

C. Cardiac Arrest Risk Scoring Based on GRU Model

We performed early prediction of cardiac arrest within 8 hours, within 24 hours, and within 72 hours. We found that the GRU model was better than other models. We evaluated the GRU model via test data. Figure 2 shows that the cardiac arrest risk score increases with the event time of cardiac arrest in cardiac arrest patients, and however, it shows low precision.

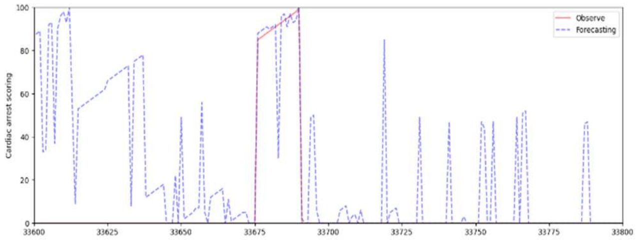

Fig. 2 Graph of cardiac arrest risk scoring

The MAE and RMSE of the GRU model are 52 and 62.69, respectively. Table 13 shows the results of early cardiac arrest prediction based on the GRU model via the cardiac arrest risk score.

TABLE XIII
EARLY PREDICTION OF CARDIAC ARREST BASED ON GRU MODEL

Time	Threshold	Precision	Recall	F1 score
Within 8 hours	92	0.07%	13.21%	0.0015
	3	0.11%	94.34%	0.0023
Within 24 hours	76	0.22%	32.58%	0.0044
	3	0.29%	96.97%	0.0058
Within 72 hours	28	0.65%	72.84%	0.0130
	1	0.69%	94.44%	0.0136

D. Discussion

Previous studies have focused on early cardiac arrest prediction. Despite the significance of the results in predicting cardiac arrest, it is difficult to determine the patient's risk level for cardiac arrest. We proposed a cardiac arrest risk score to provide significant information. To determine the patient's risk level, we assigned a score of 100 at the time of cardiac arrest. We performed multiple classifications rather than a linear regression, setting a higher weight to the time of the cardiac arrest. Table 14 shows the results of the LSTM model, the GRU model, and the self-attention-based LSTM model in predicting early cardiac arrest within 72 hours based on the weight of multiple classifications. We found that default weights have a little higher precision but poorer recall than optimized weights. If contiguous numbers have different priorities, researchers may consider multiple classifications.

TABLE XIV
PERFORMANCE RESULT OF CARDIAC ARREST WITHIN 72 HOURS BASED ON WEIGHT

Algorithm	Weight	Precision	Recall	F1 score
LSTM model	Default	4.90%	3.40%	0.0086
	Optimized	0.74%	66.67%	0.0147
GRU model	Default	5.07%	6.17%	0.0094
	Optimized	0.65%	72.84%	0.0130
Self-attention-based LSTM model	Default	2.85%	3.70%	0.0070
	Optimized	0.66%	54.01%	0.0131

In this study, clustering was used to extract similar patient populations. We hope that the information obtained from the retrospective study will assist us in identifying a group of patients who are similar to those who have experienced cardiac arrest. However, it has not been clinically validated.

TABLE XV
RESULT OF EWS AND OUR METHOD

Author	Algorithms	Precision	Recall	F1 score
Traditional EWS [5]	SPTTS	0.4%	60.7%	0.008
	MEWS ≥ 3	0.5%	63.0%	0.010
	MEWS ≥ 4	0.6%	49.3%	0.012
	MEWS ≥ 5	0.6%	37.3%	0.013
J. Kwon et al. [5]	Random forest	0.4%	75.3%	0.008
	Linear regression	0.2%	76.3%	0.004
	DEWS ≥ 2.9	0.5%	75.7%	0.010
	DEWS ≥ 3.0	0.5%	75.3%	0.010
	DEWS ≥ 7.1	0.8%	63.0%	0.0165
	DEWS ≥ 8.0	0.8%	60.7%	0.016
	DEWS ≥18.2	1.4%	49.3%	0.028
	DEWS ≥ 52.8	3.7%	37.3%	0.071
S. L. Javan et al. [8]	Stacking Early prediction within 6 hours	15%	74%	0.31
T. T Wu et al. [7]	XGboost	88.5%	73%	0.800
	Logistic regression	84.1%	58.7%	0.692
	Random forest	93.5%	46.0%	0.617
	Support vector machine	93.5%	46.0%	0.617
L. Yijing et al. [13]	XGBoost	-	86%	-
S. Hong et al. [9]	Logistic regression	-	75%	0.093
	Random forest	-	88%	0.198
	Recurrent neural network	-	84%	0.143
J. Kim et al. [12]	Char gated recurrent unit	0.8%	90.2%	-
Our method		0.11%	94.34%	0.0023

Our proposed research demonstrates a higher recall rate than other studies utilizing only vital sign data. Other studies demonstrate greater precision with clinical data. However, applying a model that predicts cardiac arrest based on clinical data to patients without measured clinical data is challenging. Our paper has three limitations. First, our dataset has a missing value. The vital signs of ICU patients are measured every 1 to 2 hours. However, the number of patients with no missing values is small. We use measured values to solve the missing values problem in this paper. However, a patient's health differs significantly from the last measured health condition. The missing values degrade machine learning training performance. Methods for interpolating missing values include substituting them with previous values, substituting them with subsequently measured values, or inferring additional items using the MICE algorithm. The MICE algorithm is a method for compensating for missing values based on another measuremenit, but vital signs data are measured at the same time. If one item is missing, other items are frequently missing. Second, our study was conducted as a single cohort at the Korea University Anam Hospital.

Accordingly, we have been unable to validate them through external sources. Third, our model is low precision.

IV. Conclusion

We proposed a cardiac arrest risk score and developed a prediction model for cardiac arrest risk score using ICU patients from the Korea University Anam Hospital. We extracted patient populations similar to cardiac arrest using clustering. Multiple classifications were used to predict cardiac arrest risk scores in a patient population with cardiac arrest-like characteristics. We evaluated the performance of each model, and our proposed GRU model had a high recall. We will improve the proposed model's precision and recall in the future. And we will conduct validation and multi-cohort studies using medical data from the Konyang University Hospital.

Acknowledgment

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ICAN(ICT Challenge and Advanced Network of HRD) program(IITP-RS-2022-00156439) supervised by the IITP(Institute of Information & Communications Technology Planning & Evaluation) and the Basic Science Research Program of the National Research Foundation (NRF) funded by the Korean government (MSIT) (NRF-2021R1A2C1009290).

References

[1] M. J. Holmberg *et al.,* "Annual incidence of adult and pediatric in-hospital cardiac arrest in the United States." *Circulation: Cardiovascular Quality and Outcomes,* vol. 12, no. 7, e005580, Jul. 2019.

[2] J. Kim, J. Jeong, and S. Kweon. "Incidences of Out-of-hospital Sudden Cardiac Arrest in the Republic of Korea, 2021." Public Health Weekly Report., CheongJu, the Republic of Korea, Dec. 2022.

[3] National Injury Information Portal , https://www.kdca.go.kr/injury/biz/injury/recsroom/rawDta/rawDtaDwldMain.do

[4] J. Kang *et al.,* "Association between time to defibrillation and neurologic outcome in patients with in-hospital cardiac arrest." *The American Journal of the Medical Sciences*, vol. 358, no. 2, pp. 143-148, Aug. 2019.

[5] J. Kwon *et al.,* "An algorithm based on deep learning for predicting in-hospital cardiac arrest." *Journal of the American Heart Association*, vol. 7, no. 13, e008678, Jul. 2018.

[6] J. H. Vähätalo *et al.,* "Association of silent myocardial infarction and sudden cardiac death." *JAMA cardiology*, vol. 4, no. 8, pp. 796-802, Jul. 2019.

[7] T. T. Wu *et al.,* "Machine learning for early prediction of in-hospital cardiac arrest in patients with acute coronary syndromes." *Clinical Cardiology*, vol. 44, no. 3, pp. 349-356, Feb. 2021.

[8] S. L. Javan *et al,.* "An intelligent warning model for early prediction of cardiac arrest in sepsis patients." *Computer methods and programs in biomedicine*, vol. 178, pp-47-58, Sep. 2019.

[9] S. Hong *et al.,* "Prediction of cardiac arrest in the emergency department based on machine learning and sequential characteristics: model development and retrospective clinical validation study." *JMIR medical informatics*, vol. 8, no. 8, e15932, Aug. 2020.

[10] D. Jang *et al.,* "Developing neural network models for early detection of cardiac arrest in emergency department." *The American journal of emergency medicine*, vol. 38, no.1, pp. 43-49, Jan. 2020.

[11] B. R. Matam, D. Heathe, and L. David, "Machine learning based framework to predict cardiac arrests in a paediatric intensive care unit: Prediction of cardiac arrests." *Journal of clinical monitoring and computing*, vol. 33, pp. 713-724, Aug. 2019.

[12] J. Kim *et al.,* "Development of a real-time risk prediction model for in-hospital cardiac arrest in critically ill patients using deep learning: retrospective study." JMIR Medical Informatics, vol. 8, no. 3, e16349, Mar. 2020.

[13] L. Yijing et *al.,* "Prediction of cardiac arrest in critically ill patients based on bedside vital signs monitoring." Computer Methods and Programs in Biomedicine, vol. 214, 106568., Feb. 2022.

[14] L. Ibrahim *et al.,* "Explainable prediction of acute myocardial infarction using machine learning and shapley values." *IEEE Access*, vol. 8, pp. 210410-210417, Nov. 2020.

[15] A. Sbrollini *et al.,* "Serial electrocardiography to detect newly emerging or aggravating cardiac pathology: a deep-learning approach." BioMedical Engineering OnLine, vol. 18, Feb. 2019.

[16] R. Ueno *et al.,* "Value of laboratory results in addition to vital signs in a machine learning algorithm to predict in-hospital cardiac arrest: A single-center retrospective cohort study." PloS one, vol.15, no. 7, e0235835, Jun. 2020.

[17] L. Nan *et al.,* "Development and validation of an interpretable prehospital return of spontaneous circulation (P-ROSC) score for patients with out-of-hospital cardiac arrest using machine learning: a retrospective study." *Eclinicalmedicine*, vol. 48, Jun. 2022.

[18] A. Shah *et al.,* "Smart cardiac framework for an early detection of cardiac arrest condition and risk." Frontiers in Public Health, vol. 9. 762303, Oct. 2021.

[19] A. Kachhawa, and J. Hitt, "An Intelligent System for Early Prediction of Cardiovascular Disease using Machine Learning." Journal of Student Research, vol. 11, no. 3, Aug. 2022.

[20] N. P. Desai *et al.,* "A Comparative Assessment Study on Machine Learning Classifiers for Cardiac Arrest Diagnosis and Prediction." In *International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems*, ICSES, Chennai, India, 2021, pp.. IEEE, 2021.

[21] R. Karthikeyan *et al.,* "Cardiac Arrest Prediction using Machine Learning Algorithms." *Journal of Physics: Conference Series*. vol. 1964, no. 6. Jul 2021.

[22] S. Baral *et al.,* "A novel solution of using deep learning for early prediction cardiac arrest in Sepsis patient: enhanced bidirectional long short-term memory (LSTM)." *Multimedia Tools and Applications*, vol. 80, pp. 32639-32664, Aug. 2021.

[23] J. Kim et al., "Predicting cardiac arrest and respiratory failure using feasible artificial intelligence with simple trajectories of patient data." Journal of clinical medicine, vol. 8, no. 9, 1336, Aug. 2019.

[24] Y. Chiu *et al.,* "Logistic early warning scores to predict death, cardiac arrest or unplanned intensive care unit re-admission after cardiac surgery." Anaesthesia, vol. 75., no. 2. pp. 162-170, 2020.

[25] S. L. Javan, and M. M. Sepehri, "A predictive framework in healthcare: Case study on cardiac arrest prediction." Artificial Intelligence in Medicine, vol. 117, 102099, Jul. 2021.

[26] D. Florence, B. Wulfran, and C. Alain, "Cardiac arrest: prediction models in the early phase of hospitalization." Current opinion in critical care, vol. 25, no. 3, pp. 204-210, Jun. 2019.

[27] F. Pedregosa *et al.,* "Scikit-learn: Machine learning in Python." the Journal of machine Learning research, vol 12 (2011): 2825-2830.

[28] N. V. Chawla et al. "SMOTE: synthetic minority over-sampling technique." Journal of artificial intelligence research, vol 16, pp. 321-357, Jun. 2002 16 (2002): 321-357.

[29] S. Hochreiter, and J. Schmidhuber, "Long short-term memory." *Neural computation*, vol. 9. no.8 pp. 1735-1780, Nov. 1997.

[30] K. Cho *et al.,* "Learning phrase representations using RNN encoder-decoder for statistical machine translation." *arXiv preprint arXiv:1406.1078*, 2014.

[31] A. Vaswani, et al. "Attention is all you need." Advances in neural information processing systems, 30, 2017.