An Explainable Multimodal Fusion Approach for Mass Casualty Incidents

Zoe Vasileiou[1,2][0000-0003-0634-6793], Georgios Meditskos[1][0000-0003-4242-5245], Stefanos Vrochidis[2][0000-0002-2505-9178], and Nick Bassiliades[1][0000-0001-6035-1038]

[1] School of Informatics, Aristotle University of Thessaloniki, Greece
`{zvasileiou,gmeditsk,nbassili}@csd.auth.gr`
[2] Centre for Research & Technology Hellas, Information Technologies Institute Thessaloniki, Greece `{zvasilei,stefanos}@iti.gr`

Abstract. During a Mass Casualty Incident, it is essential to make effective decisions to save lives and nursing the injured. This paper presents a work in progress on the design and development of an explainable decision support system, intended for the medical personnel and care givers, that capitalises on multiple modalities to achieve situational awareness and pre-hospital life support. Our novelty is two-fold: first, we use state-of-the-art techniques for combining static and time-series data in deep recurrent neural networks, and second we increase the trustworthiness of the system by enriching it with neurosymbolic explainable capabilities.

Keywords: pre-hospital life support · Explainable AI · deep neural network · neurosymbolic XAI · Mass Casualty Incident Management

1 Introduction

Mass Casualty Incidents (MCIs) are defined as "any event resulting in a number of victims large enough to disrupt the normal course of emergency and health care services" [8]. In MCIs, time is critical and the medical personnel should be aware of potential anomalies and any hazardous situations so as to assign degrees of urgency and decide the order of treatment of a large number of patients.

Deep Learning (DL) has been used to address different challenges in different domains, such as in Healthcare. Healthcare data includes various types of data, such as electronic health records (EHR) and raw signal values collected by ambient and wearable sensors as time-series. By integrating and fusing distributed and heterogeneous data, the complementarity of the multimodality is leveraged to acquire a consistent and accurate understanding of the situation. However, few studies have attempted to combine static and dynamic data, and most of them are focusing on the prediction of a specific disease.

Humans need to understand AI capability and effectively calibrate their trust. For building trustworthy clinical decision support systems, the model should be explainable and transparent through justifications. The main input of the clinical models are temporal data, but temporal explanations [10] is an unexplored and challenging task. A limited number of explainable clinical early warning systems that provide formal explanations have been developed and primarily focus on feature importance [12, 5]. Also, to the best of our knowledge, limited work has been done to combine static and dynamic information of the casualties for early warning systems, let alone to be enriched with explanations.

This paper presents work in progress from ongoing research to address the aforementioned challenges. Concretely, a Long short-term memory (LSTM) architecture, capturing the temporal dependencies in long time series, is followed. This RNN-based solution uses both the static and dynamic information of a casualty for predicting hazardous situations during a MCI, assisting the medical personnel in determining the need of medical treatment and transportation to hospitals. Additionally, the system is enhanced with explainable capabilities by providing indications how the model reached the final prediction by adopting neurosymbolic Explainability Artificial Intelligence (XAI) techniques.

The rest of the paper is structured as follows. A background and related work is provided regarding the LSTM models and existing works. Next, our methodology is presented by means of the architecture and the explainability aspects. Finally, we conclude the paper.

2 Background and Related Work

Recurrent Neural Networks (RNNs) are widely adopted in time-series classification or prediction in various kind of signals, as they allow the information to persist. RNNs are using loops for having a sort of memory. This kind of DL models enable the sequential and time-series to be represented such as the EHR and the long raw signal time sequences. The main disadvantage of RNNs is the problem of vanishing gradients that hinders the knowledge to be retained for long data sequences. LSTM models, a gated variant of RNN, can keep long-term memory by remembering long sequences of data since. Existing works were focused on the use of DL for the temporal data representation in EHR, facing various challenges, such as the data irregularity and data heterogeneity [11]. Mainly, RNN, LSTM and Gated Recurrent Units (GRUs) have been proposed for their suitability in representing temporal sequences.

Regarding the Early Warning Systems in the healthcare domain, various AI-powered solutions have been developed for predicting clinical deterioration [7]. An early detection system of heart failure onset [2] adopted a GRU model using EHR data as input. Other LSTM-based fusion approaches detect Alzheimer's progression [1] and predict early tachycardia [6]. Although temporal data entail several challenges, the opacity of the model is equally important as DL models are black-box models. Neurosymbolic XAI [3] can make black-box models transparent leveraging the inherent self-explainability of symbolic knowledge.

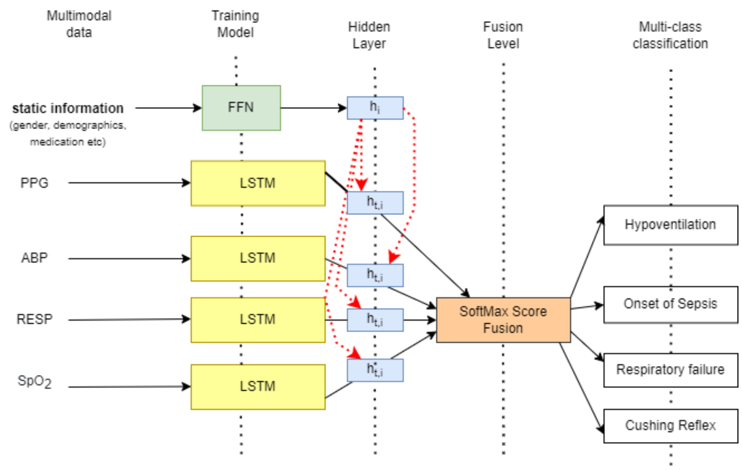

Fig. 1. Multimodal LSTM for decision support in pre-hospital life support in MCIs

3 Methodology

The proposed method provides situational awareness through explainable early warnings and decisions to the medical personnel. A DL-based approach is presented that weaves information from disparate modalities aiming at supporting medical personnel in the important decision for hospitalizing an MCI victim. The main modalities are Photoplethysmography (PPG), Blood Pressure (ABP), Respiratory Rate (RR), and Oxygen Saturation (SpO2).

A Multi-Source Information Fusion Engine integrates multiple sources and fuse data in an abstract way, aiming at detecting complex situations. We propose a LSTM-based multimodal DL algorithm for classifying the sensor data and the EHR. Our methodology combines dynamic and static data. About the dynamic data, the victims, e.g. first responders, are wearing equipment with sensors that are measuring time series. Static data include demographics, comorbidities and medication. For combining static and dynamic data, the LSTM is amalgamated with a feed-forward network. The static information is processed on a separate feed-forward network whose hidden state is concatenated with the hidden layers of the LSTMs. The late fusion scheme is applied as all the modalities are processed in different pipelines. Each modality is a physiological signal, thus processed by a LSTM model for capturing the temporal dependencies and extract patterns. The outputs of the LSTMs are concatenated and used by a softmax layer, a probability distribution over the classes, for making the final prediction. The prediction could be hazardous situations such as (i) respiratory failure, (ii) need for hypoventilation (iii) an onset of a medical emergency such as sepsis (iv) Cushing reflex, a serious situation usually seen in acute head injuries. The flow of our methodology is depicted in Figure 1.

The overall architecture is depicted in Figure 2. A pre-processing is performed for cleaning noisy and undesired signals, the data are segmented into fixed-size

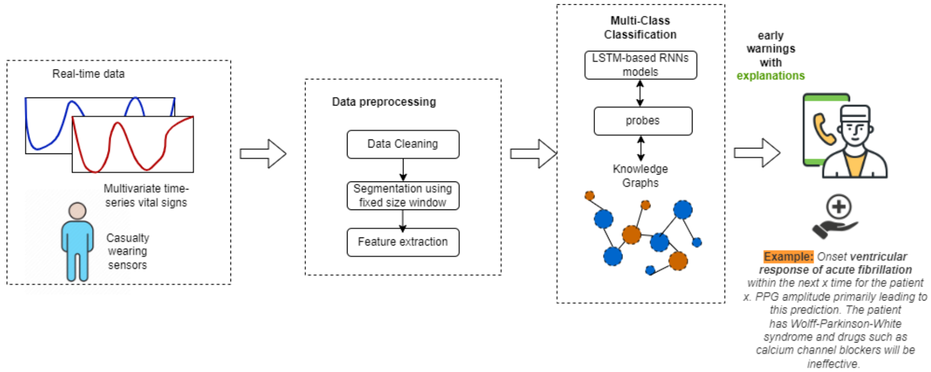

Fig. 2. The proposed architecture.

sliding windows, and then a feature generation is applied. Then, the features are sent as input to a multi-class classification algorithm that classifies the casualties into multiple classes by predicting risky situations. If a risky situation for a casualty is predicted as imminent, an early warning is sent to the medical personnel to incorporate it into their decisions about hospitalization and the level of hospital care that the casualty should receive.

One of the key features of our decision support system is that it is imbued with explainable capabilities which are prominent aspects in a multimodal environment. In order to foster an interpretable decision system, our neural system is endowed with symbolic functionalities forming a neuro-symbolic system. For providing symbolic justifications, a hidden layer analysis is performed for comprehending the concepts extracted by each activation node. A similar approach with the work of [4] is followed by using linear classifiers, known as probes, to be mapped to the intermediate layers running independently from the main model. Those probes are predicting whether a given concept was recognized by the model. The concepts entailed by the probes are mapped to ontology concepts through semantic annotation by populating knowledge graphs in a similar manner with the a mapping network approach [9] aligning artificial neural networks with ontologies. The knowledge graphs represent personalized patient data, information about diseases and clinical terminology leveraging ontologies, such as the SNOMED-CT[3] and ICD10[4].

4 Conclusion and Future Work

This paper presented a preliminary work on the development of an explainable Early Warning System that captures heterogeneous sources of a patient during a MCI. A LSTM-based architecture has been designed that amalgamates the static

[3] http://bioportal.bioontology.org/ontologies/SNOMEDCT
[4] https://bioportal.bioontology.org/ontologies/ICD10

and dynamic data of the casualty in a novel way and yields early warnings. By integrating the neural system with knowledge graphs, those early warnings are accompanied with explanations since symbolic approaches are inherently white-boxes. As next steps, we are working on finalising the implementation and testing the framework on real-world use cases[5].

Acknowledgements This work has received funding from the European Union's H2020 RIA projects NIGHTINGALE (101021957) and INGENIOUS (833435). Content reflects only the authors' view and the Research Executive Agency (REA) and the European Commission are not responsible for any use that may be made of the information it contains.

References

1. Abuhmed, T., El-Sappagh, S., Alonso, J.M.: Robust hybrid deep learning models for alzheimer's progression detection. Knowledge-Based Systems **213** (2021)
2. Choi, E., Schuetz, A., Stewart, W., Sun, J.: Using recurrent neural network models for early detection of heart failure onset. American Medical Informatics Association **24** (2016)
3. Futia, G., Vetrò, A.: On the integration of knowledge graphs into deep learning models for a more comprehensible ai—three challenges for future research. Information **11**(2) (2020)
4. Guillaume, A., Bengio, Y.: Understanding intermediate layers using linear classifier probes. ICLR (Workshop) (2017)
5. Lauritsen, S., Kristensen, M., Olsen, M., Larsen, M., Lauritsen, K., Jørgensen, M., Lange, J., Thiesson, B.: Explainable artificial intelligence model to predict acute critical illness from electronic health records. Nature Communications **11** (2020)
6. Liu, X., Liu, T., Kuo, P.C., Xu, H., Lan, K., Li, P., Ouyang, Z., Ng, Y., Yan, W., Zhang, Z., Li, D.: Top-net: Tachycardia onset early prediction using bi-directional lstm in a medical-grade wearable multi-sensor system. Medical Informatics **9** (2020)
7. Muralitharan, S., Nelson, W., Di, S., Mcgillion, M., Devereaux, P., Barr, N., Petch, J.: Machine learning-based early warning systems for clinical deterioration: A systematic scoping review. Journal of Medical Internet Research **23** (2020)
8. Organization, W.H.: Mass casualty management systems. strategies and guidelines for building health sector capacity. Geneva, Switzerland (2007)
9. de Sousa Ribeiro, M., Leite, J.: Aligning artificial neural networks and ontologies towards explainable ai. Artificial Intelligence **35**(6), 4932–4940 (2021)
10. Tonekaboni, S., Joshi, S., McCradden, M., Goldenberg, A.: What clinicians want: Contextualizing explainable machine learning for clinical end use. In: Machine Learning for Healthcare Conference, MLHC. vol. 106, pp. 359–380. PMLR (2019)
11. Xie, F., Yuan, H., Ning, Y., Ong, M.E.H., Feng, M., Hsu, W., Chakraborty, B., Liu, N.: Deep learning for temporal data representation in electronic health records: A systematic review of challenges and methodologies. Journal of Biomedical Informatics **126**, 103980 (2022)
12. Yang, M., Liu, C., Wang, X., Li, Y., Gao, H., Liu, X., Li, J.: An explainable artificial intelligence predictor for early detection of sepsis. Critical care medicine **48**(11) (2020)

[5] https://www.nightingale-triage.eu/