Automatic Detection of Behaviours of Risk in People with Dementia using Unsupervised Deep Learning

Pratik K. Mishra †,‡,*
† University of Toronto, Toronto, Ontario, Canada
‡ KITE - Toronto Rehabilitation Institute, University Health Network, Toronto, Ontario, Canada

Keywords: Agitation, Long-term care, People with Dementia, Unsupervised learning

1. Background

Dementia is a syndrome that affects cognitive and language abilities to perform activities of daily living [1] and can lead to a lack of impulse control and impact the insight and judgment of a person [2]. With the progress of dementia, it becomes necessary to provide supervision and support to the people with dementia (PwD) in their activities of daily living, which is mostly fulfilled by long-term care (LTC) homes. However, the cognitive changes in PwD can contribute to changes in their behaviour that can place them and those around them at risk. Agitation and aggression are a few of the most notable of these behaviours of risk. As documented increasingly in the literature and media [3], these agitation behaviours can lead to resident-on-resident violence and workplace violence towards staff in LTC facilities. Many LTC centers have video surveillance installed to facilitate the digital monitoring of public spaces [4]. Video data contains vital information, which in conjunction with computer vision and artificial intelligence techniques provides an opportunity to detect episodes of agitation, which can serve as an important tool to raise alarms to take appropriate interventions and enable better care of PwD. However, current research is scarce in the use of video streams for automatic detection of agitation in PwD in care homes. Our team has already collected 600 days' worth of data (video and multimodal wearable sensors) in an LTC environment during a past study conducted at the Specialized Dementia Unit, Toronto Rehabilitation Institute, Canada [4]. I plan to utilize the collected surveillance data to build automated solutions for the detection of agitation/aggression in PwD.

2. Research Objectives

The objectives of my research plan are:

(i) To develop privacy-protective approaches that learn normal behaviours of PwD using the data from a single camera and automatically detect the infrequent and diverse episodes of agitation as an anomaly.
(ii) To develop approaches that can leverage data from multiple cameras in the dementia unit to automatically detect episodes of agitation and reduce false alarms.
(iii) To automatically identify the episodes of agitation that were unreported during data collection to better understand their frequency and diversity in individual patients, which in turn can be used to deliver personalized care.

3. Methodology

An important challenge in developing digital biomarkers for the agitation episodes in PwD is their infrequent nature and diversity. They occur infrequently in comparison to

*pratik.mishra@mail.utoronto.ca

normal activities [5]. As such, the amount of labelled data available for the agitation class is very less in comparison to the class of normal behaviours [5]. Traditional supervised deep learning methods require a comparable size of data for each class and the classes need to be predetermined. Hence, the supervised deep learning methods may not work well for this problem [6]. An unsupervised approach doesn't require labelled data, hence would be more suitable to detect agitation episodes in PwD. As part of a proof-of-concept study [7] conducted using the collected data, I investigated the use of surveillance videos of a LTC home to detect episodes of agitation in PwD. An unsupervised customized spatio-temporal convolutional autoencoder was developed that was trained on the normal behaviours in PwD and identified agitation as anomaly during testing. The results indicated that it is indeed possible to detect agitation as an anomaly in surveillance videos using the unsupervised deep learning approach. I plan to utilize the following methodologies to achieve the objectives outlined in Section 2.

(i) A major concern in the use of video data is the privacy concerns associated with it [8]. The existing body-pose estimation methods [9, 10] will be used to remove the facial identity information of the participants in the raw videos while keeping intact their structural information and develop unsupervised deep learning methods that will be trained on normal behaviours and detect agitation behaviour as an anomaly.

(ii) The video data from multiple camera viewpoints can provide a better understanding of an event and can help uncover the information which is usually blocked by obstacles in any single camera's line of sight. New multi-view and data fusion-based anomaly detection deep learning architectures will be developed to leverage data from multiple cameras and reduce false alarms.

(iii) A novel positive unlabelled learning deep learning architecture will be developed to identify unreported agitation events, where the relevant representations will be learned from unlabelled videos using a self-supervised approach, followed by a binary classifier to distinguish between agitation and normal behaviour video clips.

4. **Significance and Impact**

This research is the first step in the development of software capable of monitoring video streams in real-time to detect agitation in PwD as an anomalous event. Nursing home facilities that have video surveillance systems would be able to make use of this software to increase the situational awareness and response time of staff. This is a highly plausible future technology, as video-based systems for detecting falls are already in commercial use in some nursing homes [11]. These technologies work by alerting staff when an event is detected, and by including a brief video for the staff to review and confirm the event. These systems have proven useful for event review, by providing clinically important information (e.g., confirming whether someone fell or was pushed) to facilitate better assessments and clinical decision-making (e.g., risk assessment or decision to transfer to hospital) [11]. Similarly, capturing the moments leading up to and during an agitation event can help to clarify triggers and provide context for the event, to help caregivers understand the behaviour and personalize the appropriate intervention. While agitation detection is a more complicated deep learning problem than fall detection, I am uniquely positioned to tackle this problem with our dataset and my expertise.

Acknowledgements

This research plan is supported by Dr. Alex Mihailidis and Dr. Shehroz S. Khan.

References

[1] J. Cerejeira, L. Lagarto, and E. Mukaetova-Ladinska. "Behavioral and psychological symptoms of dementia". In: *Frontiers in neurology* 3 (2012), p. 73.

[2] *Dementia*. Accessed: 19 January 2021. 2020. URL: https://www.who.int/news-room/fact-sheets/detail/dementia.

[3] *Seniors' violence in B.C. care facilities claims 16 lives*. Accessed: 27 Jan 2016. 2016. URL: https://www.cbc.ca/news/canada/british-columbia/seniors-death-violence-report-1.3422295.

[4] S. S. Khan, T. Zhu, B. Ye, A. Mihailidis, A. Iaboni, K. Newman, A. H. Wang, and L. S. Martin. "Daad: A framework for detecting agitation and aggression in people living with dementia using a novel multi-modal sensor network". In: *2017 IEEE International Conference on Data Mining Workshops (ICDMW)*. IEEE. 2017, pp. 703–710.

[5] S. Spasojevic, J. Nogas, A. Iaboni, B. Ye, A. Mihailidis, A. Wang, S. J. Li, L. S. Martin, K. Newman, and S. S. Khan. "A pilot study to detect agitation in people living with dementia using multi-modal sensors". In: *Journal of Healthcare Informatics Research* 5.3 (2021), pp. 342–358.

[6] S. S. Khan and M. G. Madden. "One-class classification: taxonomy of study and review of techniques". In: *The Knowledge Engineering Review* 29.3 (2014), pp. 345–374.

[7] S. S. Khan, P. K. Mishra, N. Javed, B. Ye, K. Newman, A. Mihailidis, and A. Iaboni. "Unsupervised Deep Learning to Detect Agitation from Videos in People with Dementia". In: *IEEE Access* (2022).

[8] B. Ye, S. S. Khan, B. Chikhaoui, A. Iaboni, L. S. Martin, K. Newman, A. Wang, and A. Mihailidis. "Challenges in collecting big data in a clinical environment with vulnerable population: Lessons learned from a study using a multi-modal sensors platform". In: *Science and engineering ethics* 25.5 (2019), pp. 1447–1466.

[9] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh. "Realtime multi-person 2d pose estimation using part affinity fields". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, pp. 7291–7299.

[10] M. Bidstrup, J. V. Dueholm, K. Nasrollahi, and T. B. Moeslund. "Privacy-Aware Anomaly Detection Using Semantic Segmentation". In: *International Symposium on Visual Computing*. Springer. 2021, pp. 110–123.

[11] E. Bayen, S. Nickels, G. Xiong, J. Jacquemot, R. Subramaniam, P. Agrawal, R. Hemraj, A. Bayen, B. L. Miller, G. Netscher, et al. "Reduction of time on the ground related to real-time video detection of falls in memory care facilities: observational study". In: *Journal of medical internet research* 23.6 (2021), e17551.