Computer Vision-based Approach to Maintain Independent Living for Seniors

[1]Zelun Luo, [1]Alisha Rege, [1]Guido Pusiol, [2]Arnold Milstein, [1]Li Fei-Fei, [2]N. Lance Downing
[1]Stanford Vision Lab, Stanford University, [2]Department of Medicine, Stanford University

Introduction

As the first baby-boomers begin to reach retirement age, the U.S. has begun to experience a shift in the age demographics of its population that has significant implications for Medicare spending and the federal budget. Thus far, computer vision has not had significant application in healthcare but holds significant potential for automation and augmentation in the monitoring and recognition of health and healthcare behaviors[1]. In this work, we introduce a computer vision-based system to allow seniors to live independently at home. Depth and thermal sensors are used as they hold advantages over wearable sensors and RGB cameras: they do not require daily wearing, provide richer information, avoid personal identification, and operate in different lighting conditions. Our prototypes are lightweight and easy to install, requiring only power source and internet access. We introduce a new open-source dataset: Thermset[2], which contains 214 hours of thermal video collected at senior bedrooms. A convolutional neural network (CNN) is used to detect five clinical states: background, people present, standing, sitting, and sleeping. These states are important on their own and necessary to identify more complex activities that are more likely to be actionable by caregivers.

Methods

We trained a CNN-based feed forward network on Thermset (Figure 2) to detect whether an activity occurs in a video frame. Our network is inspired in AlexNet[3], and it consists of five convolutional layers. Each layer is followed by a ReLU nonlinearity layer, a batch-normalization layer, and a dropout layer. The output is the multi-label classification of the target activities, obtained with a softmax function. At training stage, 21,960 frames (4,392 from each category) were used. For evaluation, we do not include any frame of a scene that appears in the training set.

Results

Figure 1 shows the confusion matrix for the algorithm. The matrix shows that an average 76% correct classification rate on the five aforementioned clinical states. We were only recently able to begin collecting data from depth sensors so do not have results to present here. As we add data diversity by collecting more data and using depth data, we expect a much more robust system in the future.

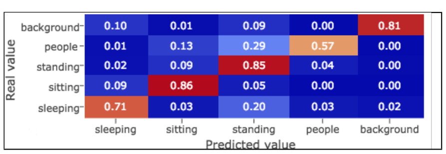

Figure 1: Confusion Matrix on Thermset

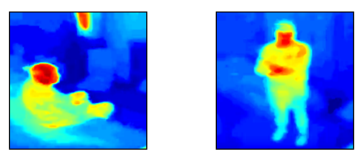

Figure 2: Thermset Dataset Examples

Conclusion

This work shows the viability of using thermal signals for monitoring elderly citizen, and can potentially allow for them to receive the care that they require from the comfort of home. In our future work, we plan to train a model that recognizes a wider range of clinical states, including day and night reversals, eating, falls, slowed movements, unstable transfers, front door loitering, chair and bed immobility, urinary frequency, restlessness, and fever. By detecting these clinical states, we aim to identify elder patients who are at risk for requiring long-term care, and to provide feedback to caregivers that would support their safe and independent living.

References

1. Gulshan, V. D. et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 316, 24022410 (2016)
2. Pusiol, P (n.d.). Thermset [Scholarly project]. In Github. Retrieved from https://github.com/blusa/research/raw/master/Thermset.
3. A. Krizhevsky, I. Sutskever, and G. Hinton. ImageNet classification with deep convolutional neural networks. In NIPS, 2012.