HUMAN ACTIVITY RECOGNITION

WHICH DEEP LEARNING APPROACHES WORK BEST?

INTRODUCTION & AIMS

Human activity recognition (HAR) is the process of identifying people's actions and activities, typically from data captured by sensors such as cameras, accelerometers and gyroscopes.

Over the years, many different algorithms have been applied to HAR with various amounts of success, but recently approaches using deep learning neural network (NN) techniques have shown to be the most promising [1]. However, the question as to which method gives the highest performance is still unanswered. In light of this, it is the aim of my summer project to identify which NN architectures are most effective at HAR by conducting a systematic review involving several HAR datasets.

By identifying the most effective methods and determining any specific situations where there are exceptions, I aim to contribute to the literature in a way that helps engineers to select the most appropriate techniques for their future applications.

MEASURES

In most HAR datasets, a few activities tend to appear more than others and this causes a problem when using accuracy as a measure of performance.

Instead, we can use the average F1 score for each class [2]. F1 score balances recall and precision metrics as defined in Figure 1, meaning that misclassifying minority classes is just as important as misclassifying majority ones. F1 is used when misclassifications have varying costs.

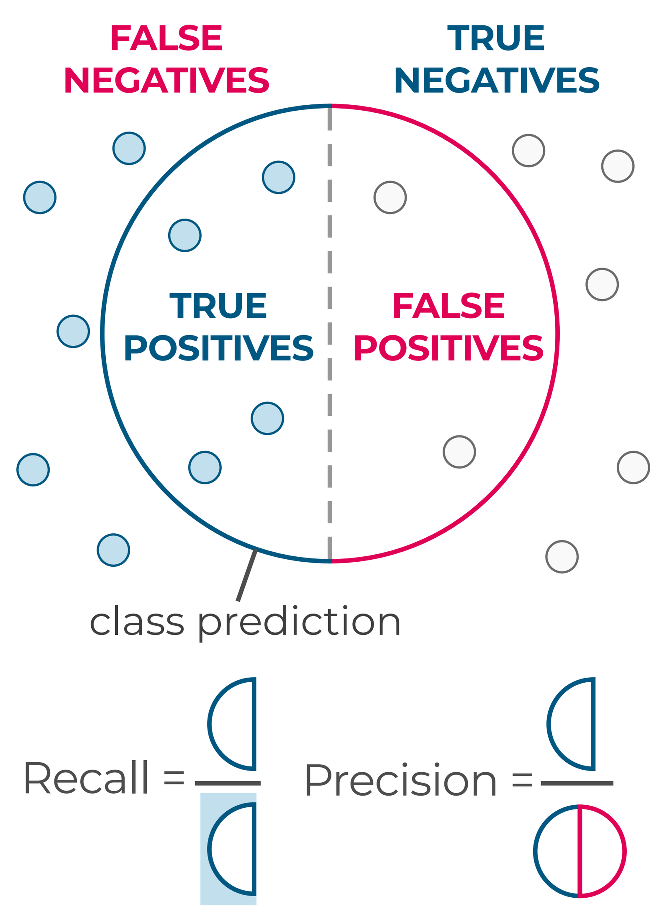

Fig 1. Visualising recall & precision. Adapted from [3]

METHODS

The main techniques to be evaluated are 1-dimensional (1D) convolutional NN (CNNs) and Long Short-Term Memory (LSTM) NNs. CNNs work by learning which features to extract from the signal automatically [4], and their translational and scale invariance properties, which result from the convolution and pooling operations (Fig. 2), make them suitable for signal processing.

LSTM NNs are also useful for processing signals [5]. This is made possible by the internal memory of the system, indicated by **c** in Fig. 3, and the feedback loops from one time-step **t** to the next **t+1**. Hybrid architectures using both techniques will also be evaluated.

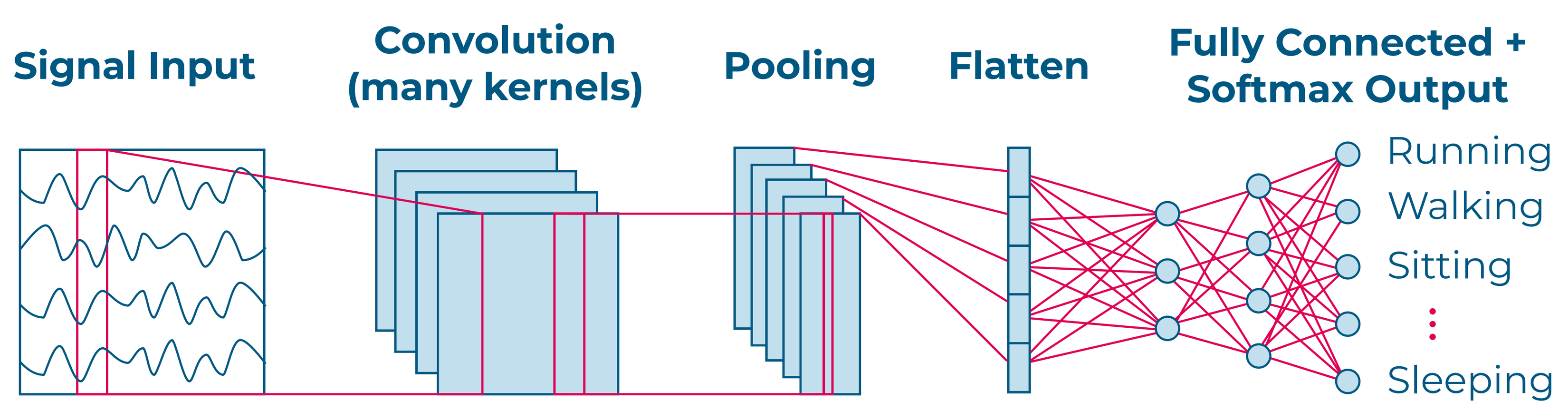

Fig 2. Simplified Convolutional Neural Network with 1D time convolution.

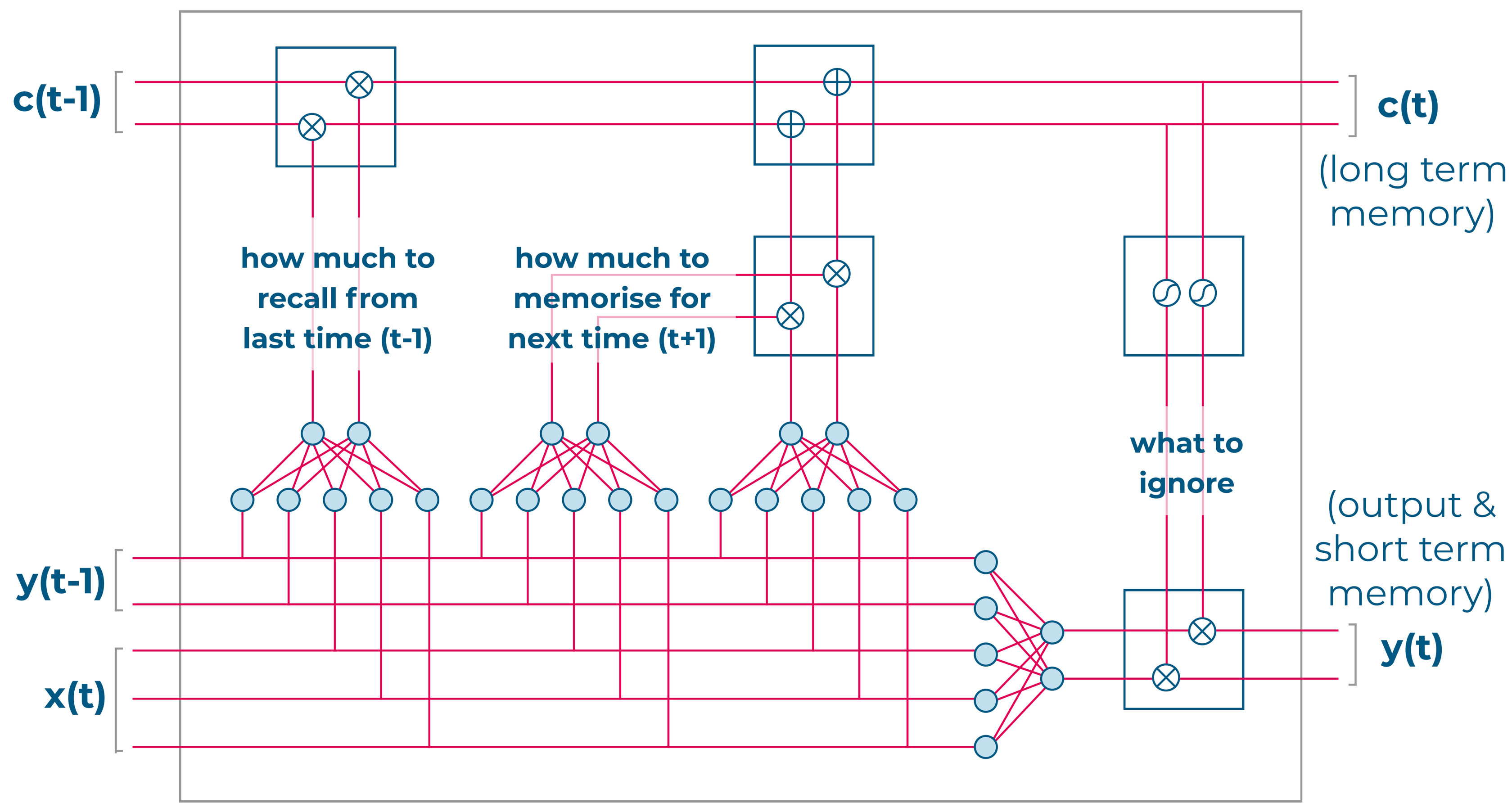

Fig 3. Layer from a Long Short-term Memory (LSTM) network. **x(t)** is the input to the layer at time **t.** Adapted from [7]

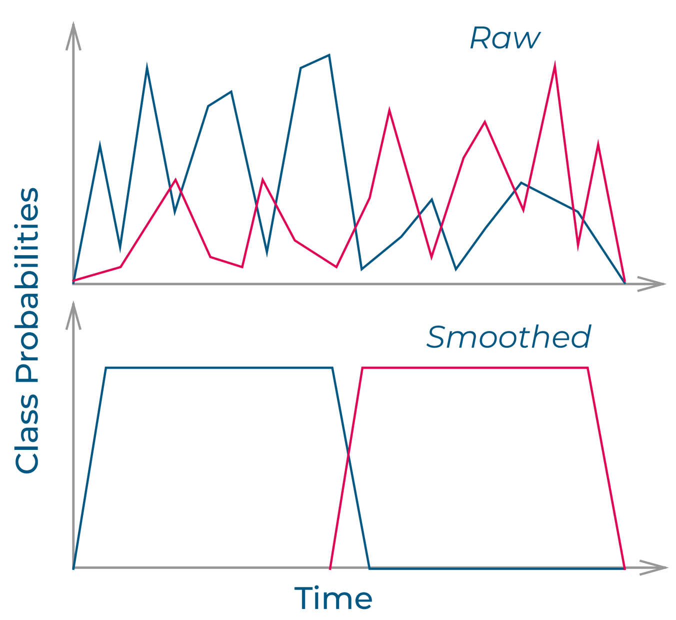

Fig 4. K-Nearest Neighbour voting for sequential classifications [6].

Since HAR data is typically sequential in nature, in order to use techniques which require input vectors of fixed length (like CNNs), the data must be classified in steps. This causes a problem with sequential predictions fluctuating between classes, and so smoothing methods are required. In this project, new kinds of smoothing techniques, like the one illustrated in Fig. 4, will be explored.

MOTIVATIONS

Improving the performance of HAR devices has many potential benefits. For example, improving fall detection devices used by the elderly would reduce the number of dangerous falls missed by current devices. Additionally, if more complex behaviours can be classified, HAR could be used to better analyse the activities of an individual to improve sports training. Developing the capabilities to identify unusual behaviours also has security applications for predicting possible threats, for example in airport security systems.

REFERENCES

[1] Ordóñez, F.J. and Roggen, D., 2016. Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition. Sensors, 16(1), p.115.

[2] Sokolova, M. and Lapalme, G., 2009. A systematic analysis of performance measures for classification tasks. Information Processing & Management, 45(4), pp.427-437.

[3] Walber. November 2014 - https://en.wikipedia.org/wiki/Precision_and_recall#/media/File:Precisionrecall.svg

[4] LeCun, Y. and Bengio, Y., 1995. Convolutional networks for images, speech, and time series. The handbook of brain theory and neural networks, 3361(10), p.1995.

[5] Hochreiter, S. and Schmidhuber, J., 1997. Long short-term memory. Neural computation, 9(8), pp.1735-1780.

[6] Mitchell, H.B. and Mashkit, N., 1992. Noise smoothing by a fast k-nearest neighbour algorithm. Signal Processing: Image Communication, 4(3), pp.227-232.

[7] Prugel-Bennet 2018, COMP6208 LSTM Lecture Notes. https://secure.ecs.soton.ac.uk/notes/comp6208/lectures/lstm.pdf

Presented by Gregory Walsh - MSc Data Science
Supervisor Kate Farrahi
May 2018

UNIVERSITY OF Southampton