BlobSnake: Gamification of Feature Selection for Human Activity Recognition

Reuben Kirkham
Digital Interaction Group.
Newcastle University.
r.kirkham@newcastle.ac.uk

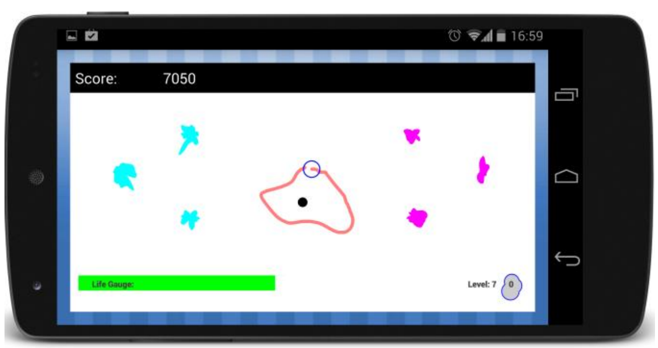

Figure 1: Illustration of BlobSnake. The goal of the game is to draw shapes that more effectively fit one side (representing one class of data) better than the other (whch represents different class or activity). The game incorporates a metric which is very similar to a snake fitting process from computer vision. The design is somewhat inspired by previous work that involved effectively visualizing sensor data [3]

Abstract
This paper discusses nascent work at Newcastle University's Digital Interaction Group, focused upon gameifying feature selection for Human Activity Recognition (HAR). The goals are two fold; the first is to mitigate the current need for a HAR expert to develop a feature selection for novel activity recognition problems, whilst the second is to address the need for science communication of this domain, especially in the legal setting. The initial game that has been developed – BlobSnake – is also briefly presented.

Author Keywords
Evidence, Gamification, Human Activity Recognition, Machine Learning, Science Communication.

Introduction
Human Activity Recognition (HAR) is an emerging field, which often involves the use of wearable sensors in order to detect specific activity patterns in the wild. Each system is hand developed to take a sensor stream – most likely from one or more inertial sensors (accelerometers or gyroscopes) and then points (or windows) in time are classified by a decision rule that partitions them into one of the particular activities under study, or a null (none of the above) category.

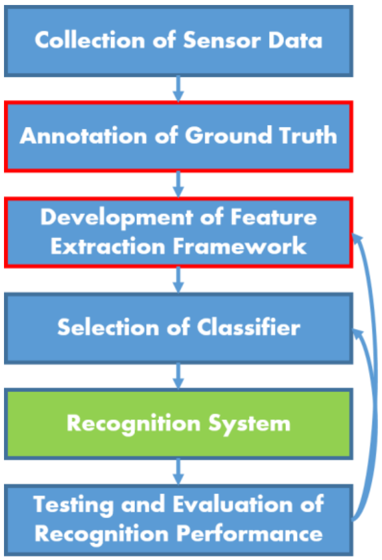

Figure 2: Illustration of the approach that is normally taken towards developing HAR systems. It can be seen that this involves a large iterative attempts at designing a system. In red outline, the two most challenging and burdensome aspects of this process are highlighted; these represent barriers to the wider deployment of HAR approaches.

Developing such a system is usually a non-trivial task. It involves the collection and precise temporal annotation of sensor data, followed by the use of an expert to develop a carefully selected set of features to use in a machine learning system (See **Figure 2**). Whilst it is true that standard features, such as the ECDF [2] can be used, it is usual that additional features should be added in order to improve the performance to a more reasonable level. The existing limitations are serious barriers to the wider adoption of HAR systems, and in particular have the effect of helping to restrict them to certain pluralities of typical (and usually, for instance, non-disabled) people that a system was originally trained upon. Reducing the amount of expertise required to develop a system based upon annotated training data would be of great assistance in broadening the reach of HAR systems. This is a primary goal of this work, by reducing the task of bespoke additional feature selection into being a matter of being a game that anyone can play, HAR can begin to become more inclusive as an endeavor. At the same time, there is a serious deficit in science communication, in that there is no existing work that attempts to explore HAR from a public engagement perspective. We attempt to fill this gap, noting that this application is most important in a legal context. This submission also explains our approach towards gamification of feature selection.

Science Communication, HAR, and the Law

If a HAR system could be used effectively in a legal context, then there would be profound benefits. Indeed, with the recent case involving FitBit, it appears that the legal community are on the verge of attempting to realize this potential opportunity. This author considers that such advances – if realized -are most likely to have an impact upon civil cases, particularly in respect of 'mental capacity', where someone's liberty is at stake, but to a civil standard (50.1%) rather than the criminal beyond reasonable doubt test. For example, in the case of *Neary vs Hillingdon Borough Council*, a dispute about the behavior of the plantiff's behavior, led to an extended deprivation of their own liberty; an effective recognition system on the line of Ploetz et al [5] might well have provided the court with sufficient evidence to address this, as well as empowering parties, such as a carer or supporter to more effectively oppose existing authorities and decision makers.

The current problem is that HAR systems have no real framework for which they could be used in a court context. The root of this problem lies with the fact that most judges are mathematically illiterate [4], with past judicial practice including accepting a p value of >0.5 as sufficient proof, before eventually switching to values more widely used in the scientific community; other cases have focused upon using actuarial calculations in order to reduce compensation payouts based upon gender and ethnicity. HAR represents a novel means of potentially providing evidence, and is therefore highly likely to encounter similar issues; this is especially of concern given the lack of agreed metrics for assessing a systems performance [6].

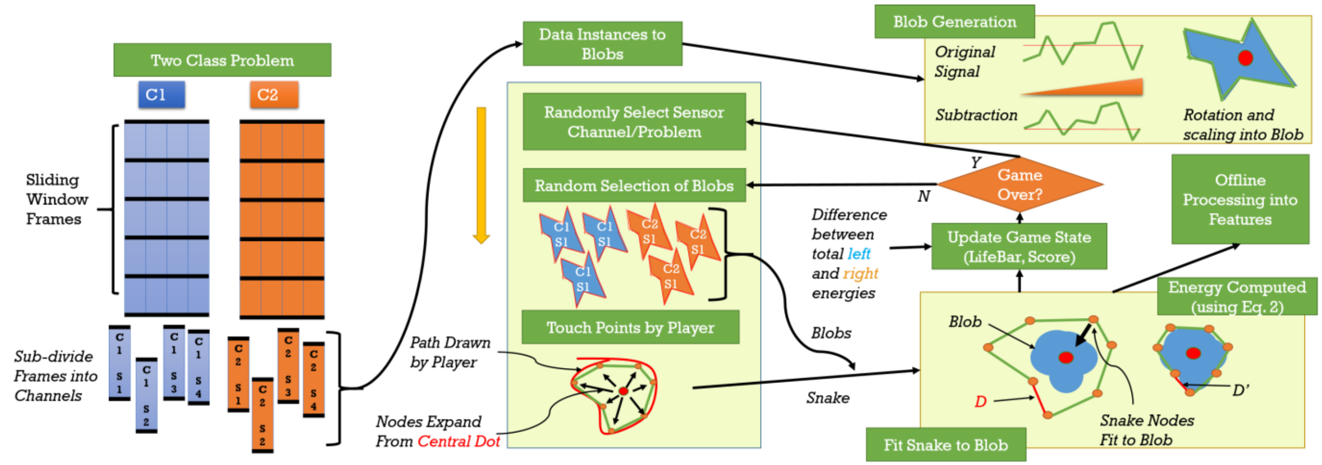

Figure 3: The algorithms involves in BlobSnake. On the left, the current two class machine learning problem that we will be using to explore the system in the context of an initial evaluation – although it is hoped that BlobSnake will scale to multi-class problems via a 1 vs all approach commonly used already in the machine learning

Skillful science communication – such as through BlobSnake – of the complexities and limitations of HAR to lay people, including lawyers and judges, is an essential step towards realizing the fairness of HAR usage in any evidential context. It is hoped that this work can be one initial step towards this far reaching goal, as well as assisting in furthering the debate around developing appropriate metrics for HAR systems that can be used in the legal setting.

BlobSnake

BlobSnake (**Figure 1**) is an Android game under development at Newcastle University. It comprises three components:

1. A script that deterministically maps existing sensor data into blobs that can be interacted within the game.
2. The game itself, where the players perform a feature generation task.
3. A simple algorithm that selects the most promising generated features for use in a system, in addition to standard features.

This workshop submission – for reasons of space – focusses upon briefly explaining the game itself (i.e. 1 and 2) We presume a sliding window approach, where sensor streams are chopped into fixed windows of N seconds (usually N =1), this is the most common approach in Human Activity Recognition [1]. The exercise of feature selection is a matter of finding functions that summarize data effectively as to aid a classifier in developing an effective decision rule.

Equation 1:

$$SF = \begin{cases} 1/|\Delta F| & \text{if}|\Delta F| \leq 2 \\ \log(1/|\Delta F|) & \text{otherwise} \end{cases}$$

$$\Delta F = |F_{min} - F_{max}|)$$

Equation 2:

$$E = \min_{S_r} \sum_{seg=1}^{n} abs\left(\frac{L_{seg}}{\sum L} - \frac{L'_{seg}}{\sum L'}\right)$$

Figure 4: Equations references in the text. Equation 1 is the scale factor used when translating data into blobs. Equation 2 is the energy function applied to a given snake with respect to each blob. Note we apply this in such a way as to be rotationally invariant.

Generating Blobs

We take a window of data, and from this a single channel of sensor data. This is then wrapped around as demonstrated in **Figure 3** in order to create a blob. The mapping here is deterministic, with minimal loss in sensor data. A dynamic scale factor (Eq1, **Figure 4**) is used to avoid signal noise being scaled into a meaningful blob size.

The Game

The player is presented with a number of blobs; on the left is a set of blobs from one class, on the right another. The goal is to draw an abstract shape which serves as a feature going forwards, namely that it will effectively assist a system in classifying – or distinguishing between – two categories. For each problem the player is presented with, the shape they draw is automatically evaluated, and the life bar adjusted, based upon the different between the left and right fitting energies (below).

Feedback and Fitting Metrics

The system uses an adaption of a snake fitting algorithm. This works by reducing the touch input into a series of points, and thus a polygon. This polygon is then deformed to fit each individual blob in turn, with the energy being the average change in segment wise length, normalized over the length of the polygon, with the formula being provided in Eq2 (**Figure 4).** The benefits of such an approach is that an identical approach can be used in the final (automatic) development of a system, as well as being efficient enough to quickly compute on mobile devices.

Conclusion

This submission has overviewed BlobSnake, a novel and exploratory approach towards engaging the public in Human Activity Recognition system development, as well as aiming towards contributing to science communication in a legal setting. This authors looks forward to discussing these concerns at the workshop itself, and beginning to advance these two agendas.

Acknowledgements

The author would like to thank those he has worked with on this submission: Patrick Olivier, Carlton Shepherd and Thomas Ploetz. This work was supported by an EPSRC DTA Award and a Google Scholarship.

References

[1] Bulling, A. et al. 2013. A Tutorial on Human Activity Recognition Using Body-worn Inertial Sensors. *ACM Computing Surveys*. (2013).

[2] Hammerla, N. et al. 2013. On Preserving Statistical Characteristics of Accelerometry Data using their Empirical Cumulative Distribution. *ISWC* (2013).

[3] Kirkham, R. et al. 2013. The Breaktime Barometer – An Exploratory System for Workplace Break-time Social Awareness. *Ubicomp 2013: Ubiquitious Computing* (2013).

[4] Meyerson, M.I. 2010. Significant Statistics : The Unwitting policy Making of Mathematically Ignorant Judges Significant Statistics : The Unwitting Policy Making of Mathematically Ignorant Judges. 37, 3 (2010).

[5] Plötz, T. et al. 2012. Automatic Assessment of Problem Behavior in Individuals with Developmental Disabilities. *Ubiquitous Computing 2012* (2012).

[6] Ward, J.A. and Gellersen, H.W. 2011. Performance Metrics for Activity Recognition. *ACM Transactions on Intelligent Systems and Technology*. 2, 1 (2011), 1–23.