Computers and Electronics in Agriculture 93 (2013) 17–26

Contents lists available at SciVerse ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Sow-activity classification from acceleration patterns: A machine learning approach

Hugo Jair Escalante [a,*], Sara V. Rodriguez [b], Jorge Cordero [b], Anders Ringgaard Kristensen [c], Cécile Cornou [c]

[a] Computational Sciences Department, Instituto Nacional de Astrofísica, Óptica y Electrónica, Tonantzintla, 72840 Puebla, Mexico
[b] Graduate Program in Systems Engineering, Universidad Autónoma de Nuevo León, San Nicolás de los Garza, 66450 Nuevo Leon, Mexico
[c] HERD – Centre for Herd-oriented Education, Research and Development, Department of Large Animal Sciences, University of Copenhagen, Groennegaardsvej 2, DK-1870 Frederiksberg C, Denmark

ARTICLE INFO

Article history:
Received 21 June 2012
Received in revised form 17 December 2012
Accepted 13 January 2013

Keywords:
Accelerometer measurements
Logitboost with trees
Pattern classification
CLOP
Sow-activity classification

ABSTRACT

This paper describes a supervised learning approach to sow-activity classification from accelerometer measurements. In the proposed methodology, pairs of accelerometer measurements and activity types are considered as labeled instances of a usual supervised classification task. Under this scenario sow-activity classification can be approached with standard machine learning methods for pattern classification. Individual predictions for elements of times series of arbitrary length are combined to classify it as a whole. An extensive comparison of representative learning algorithms, including neural networks, support vector machines, and ensemble methods, is presented. Experimental results are reported using a data set for sow-activity classification collected in a real production herd. The data set, which has been widely used in related works, includes measurements from active (*Feeding*, *Rooting*, *Walking*) and passive (*Lying Laterally, Lying Sternally*) activities. When classifying 1-s length observations, the best method achieved an average recognition rate of 74.64%, for the five activities. When classifying 2-min length time series, the performance of the best model increased to 80%. This is an important improvement from the 64% average recognition rate for the same five activities obtained in previous work. The pattern classification approach was also evaluated in alternative scenarios, including distinguishing between active and passive categories, and a multiclass setting. In general, better results were obtained when using a tree-based logitboost classifier. This method proved to be very robust to noise in observations. Besides its higher performance, the suggested method is more flexible than previous approaches, since time series of any length can be analyzed.

1. Introduction

Automated monitoring of animal behavior enables oestrus, health disorders, and animal-welfare in general to be supervised on a large scale. It is therefore an important research area within livestock production. Recent research and development have targeted animal activity recognition, since the recognition of basic animal activities can help to detect and monitor important events such as oestrus, pregnancy or parturition. Data collected from sensors physically-attached to animals have been successfully used to classify the activities of individual animals when housed in groups (Cornou and Lundbye-Christensen, 2010; Firk et al., 2002; Umstatter et al., 2008). The main motivation behind physically-attached sensors is to gather real-time (first hand) information of the animals' behavior. In addition, sensors such as infrared and accelerometers are affordable and accurate, which make them suitable tools for commercial research. The activities of dairy cows, sows and other species have been monitored and classified using data collected from these types of sensor.

The present work focused on the classification of individual sows' activity using accelerometers measurements. Previous studies (Cornou and Lundbye-Christensen, 2008; Cornou and Lundbye-Christensen, 2010) used dynamic linear models to classify different sow activities. In particular, Cornou and Lundbye-Christensen (2010) used a Multi-Process Kalman Filter (MPKF) which achieved excellent classification results for passive (lying laterally, LL, and lying sternally, LS) and active (feeding, FE, rooting, RO, and walking, WA) activities. The authors reported a 64.4% average recognition rate. The current study aimed at improving the recognition performance obtained by the MPKF for active (FE RO WA) and passive (LL and LS) sow activities by classifying accelerometer data using a supervised machine learning process that neglects time dependencies between sample-measurements. Specifically, the classification of time series is approached as a standard

* Corresponding author.
E-mail address: hugojair@inaoep.mx (H.J. Escalante).

0168-1699/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.compag.2013.01.003

(atemporal) pattern classification task that can be solved with a variety of techniques (Duda et al., 2000; Hastie et al., 2009). In this way, time series of arbitrary length (duration) can be analyzed by combining the predictions provided by the model for the elements (acceleration measurements) of the series. This formulation offers a wide flexibility for the on-line monitoring of animals. Furthermore, it was hypothesized that four accelerometer measurements (axes x,y,z, and the norm of the acceleration vector) recorded at an instant (1 Hz in this study) are informative enough to discover and recognize sow activities.

Using the data set from Cornou and Lundbye-Christensen (2010), classification results are generated by applying six of the most representative classifiers from the fields of machine learning and pattern recognition (Duda et al., 2000; Hastie et al., 2009; Saffari and Guyon, 2006): neural networks (`neural`), support vector machines (`SVM`), Naïve Bayes (`naive`), a linear classifier (`zarbi`), random forest (`RF`) and logitboost-with-trees (`logitboost`). The performance of these classifiers is evaluated under different scenarios.

The main contributions of the study presented in this paper are as follows:

- A highly-effective supervised-learning approach to sow-activity classification where time dependencies between measurements are ignored.
- The proposed approach is able to classify measurements recorded at an instance (a second) of time, facilitating the real-time monitoring of animal-behavior in practice.
- In addition, a method combining predictions made at the observation (second) level for classifying time series of varying length is proposed.

The remainder of this paper is organized as follows: Section 2 describes the method used to obtain the accelerometer data of five types of sow activities. Section 3 follows with background information on pattern classification and leads into Section 4 where the supervised learning approach is presented. Section 5 reports the classification results obtained using the six classification methods for the five activity types. The classifiers are tested on an individual basis, per activity, and as a multiclass problem, across all activities. The classifiers are also tested under different data-input scenarios, using data samples at 1 Hz (an observation) and for time series of accelerometer data of 2 min (a series of observations). Section 6 concludes with the findings of this study and outlines future work directions.

2. Acceleration measurements

2.1. Time series recording

Time series of acceleration measurements were collected for 11 group-housed sows, in a Danish production herd. These experimental sows were housed in a dynamic group of approximately 100 sows, where the pen was 22.45 m long by 12.45 m wide. Resting areas were straw-bedded and activity areas had solid or slatted floors.

The accelerometers and a battery package was placed on a box fitted on each experimental sow using a neck collar, so that the box was placed on the lowest part, i.e. at the bottom of the neck, for each of the 11 sows. Acceleration data were measured in three dimensions using a digital accelerometer (LIS3L02DS from STMicroelectronics) four times per second, 24 h a day, during 20 days. Furthermore, the sows were video recorded 24 h a day. Video recordings were used to identify the types of activity that the experimental sows (individually marked on their back) were performing.

2.2. Data set construction

This study used the two data sets from Cornou and Lundbye-Christensen (2008), Cornou and Lundbye-Christensen (2010). Five types of activity were included: feeding (FE), rooting (RO), walking (WA), lying sternally (LS) and lying laterally (LL).

The data sets contain extracts (observations) of time series corresponding to each of the five activities. Each extract is a 4D vector of measurements, with values for the three-dimensional axes x, y and z and the length of the acceleration vector $a = \sqrt{x^2 + y^2 + z^2}$. A learning (or training) data set was used to train discriminative models and a test data set was used to evaluate the classification methods.

- The learning data set includes 46 series of 10 min: 6, 7, 11, 11 and 11 series, respectively for FE, RO, WA, LS and LL.
- The test data set includes 490 series of 2 min: 84, 79, 107, 110 and 110, respectively for FE, RO, WA, LS and LL.

Video recordings were used to select the series' extracts. The procedure was carried out by a single person, who simultaneously analyzed the video and noted the start and end of activities on the printed time series. Since a change of activity can be visualized on the series (for one or more axes), this ensured a good concordance between the activity and the series' extract. Any overlapping of activity (especially between RO and FE) was reduced to a minimum. Moreover, missing data and the fact that sows perform more rarely RO and FE activities resulted in a smaller number of series for these activities.

The two data sets differ in terms of time series' length. For the learning data set, a length of 10 min was chosen in order to have sufficient training time for the development of the classification method, and considered as a maximum length (especially with respect to FE), since the extract should contain an activity performed continuously. For the test data set, a length of 2 min was considered as sufficient to recognize a given activity, and short enough to reduce overlapping of different activities. The data set used in this study is described in more details in Cornou and Lundbye-Christensen (2010).

3. Pattern classification

A wide range of pattern classification methods have been developed (Duda et al., 2000; Hastie et al., 2009), as pattern classifiers are important core components within machine learning and pattern recognition systems. The supervised pattern-classification process involves finding a map between observations (inputs) and labels (outputs), given a set of data for which the correspondence between inputs (observations) and outputs (labeled data) are known. In this study, during the process of classifying sow activities, the four accelerometer measurements are considered to be the observations and the labels correspond to the different activities to be recognized. Classic pattern classification problems include: handwritten digit recognition, spam filtering and face recognition. In this work, observations are 4D accelerometer measurements and the labels correspond to the different activities to recognize (Section 4).

Let us consider a data set $\mathcal{D}$ formed by N pairs in the form $(\mathbf{x}_i, y_i)$, where $\mathbf{x}_i \in \mathbb{R}^d$ is an observation, d is the dimensionality of the observations, and $y_i \in C$ indicates the corresponding class label for $\mathbf{x}_i$, where $C = \{1, \ldots, K\}$ for a problem of K classes or labels associated to the problem at hand.

The classification problem for $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}_{1,\ldots,N}$ consists of finding a function f of the form $f : \mathbf{x}_i \in \mathbb{R}^d \to y_i \in C$, from the paired samples in $\mathcal{D}$. The learned function f must be able to classify