

Robust Sensor-Orientation-Independent Feature Selection for Animal Activity Recognition on Collar Tags

JACOB W. KAMMINGA, University of Twente, The Netherlands

DUC V. LE, University of Twente, The Netherlands

JAN PIETER MEIJERS, University of Twente, The Netherlands

HELENA BISBY, University of Twente, The Netherlands

NIRVANA MERATNIA, University of Twente, The Netherlands

PAUL J.M. HAVINGA, University of Twente, The Netherlands

Fundamental challenges faced by real-time animal activity recognition include variation in motion data due to changing sensor orientations, numerous features, and energy and processing constraints of animal tags. This paper aims at finding small optimal feature sets that are lightweight and robust to the sensor's orientation. Our approach comprises four main steps. First, 3D feature vectors are selected since they are theoretically independent of orientation. Second, the least interesting features are suppressed to speed up computation and increase robustness against overfitting. Third, the features are further selected through an embedded method, which selects features through simultaneous feature selection and classification. Finally, feature sets are optimized through 10-fold cross-validation. We collected real-world data through multiple sensors around the neck of five goats. The results show that activities can be accurately recognized using only accelerometer data and a few lightweight features. Additionally, we show that the performance is robust to sensor orientation and position. A simple Naive Bayes classifier using only a single feature achieved an accuracy of 94% with our empirical dataset. Moreover, our optimal feature set yielded an average of 94% accuracy when applied with six other classifiers. This work supports embedded, real-time, energy-efficient, and robust activity recognition for animals.

CCS Concepts: • **Computing methodologies** → **Supervised learning; Feature selection; Classification and regression trees;**

Additional Key Words and Phrases: Animal Activity Recognition, Sensor Orientation, Embedded Systems, Machine Learning, Decision Tree, Naive Bayes

This research was supported by the Smart Parks Project, which involves the University of Twente, Wageningen University & Research, ASTRON Dwingeloo, and Leiden University. The Smart Parks Project is funded by the Netherlands Organisation for Scientific Research (NWO).

Authors' addresses: Jacob W. Kamminga, University of Twente, Drienerlolaan 5, 7522NB, Enschede, The Netherlands, j.w.kamminga@utwente.nl; Duc V. Le, University of Twente, Drienerlolaan 5, 7522NB, Enschede, The Netherlands, v.d.le@utwente.nl; Jan Pieter Meijers, University of Twente, Drienerlolaan 5, 7522NB, Enschede, The Netherlands, j.p.meijers@utwente.nl; Helena Bisby, University of Twente, Drienerlolaan 5, 7522NB, Enschede, The Netherlands, h.c.bisby@student.utwente.nl; Nirvana Meratnia, University of Twente, Drienerlolaan 5, 7522NB, Enschede, The Netherlands, n.meratnia@utwente.nl; Paul J.M. Havinga, University of Twente, Drienerlolaan 5, 7522NB, Enschede, The Netherlands, p.j.m.havinga@utwente.nl.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2018 Association for Computing Machinery.

2474-9567/2018/3-ART15 \$15.00

<https://doi.org/10.1145/3191747>

Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, Vol. 2, No. 1, Article 15. Publication date: March 2018.

ACM Reference Format:

Jacob W. Kamminga, Duc V. Le, Jan Pieter Meijers, Helena Bisby, Nirvana Meratnia, and Paul J.M. Havinga. 2018. Robust Sensor-Orientation-Independent Feature Selection for Animal Activity Recognition on Collar Tags. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 1, Article 15 (March 2018), 27 pages. <https://doi.org/10.1145/3191747>

1 INTRODUCTION

Animal behavior is a commonly used and sensitive indicator of animal welfare and can provide rich information about their environment [25, 34, 39]. Previous research has shown that the physical behavior of animals can be an early indicator of diseases, pain, and heat-stress. It also provides information about social interaction within a herd [7]. The behavior of groups of animals can even be an indicator of the occurrence of environmental events such as forest fires [39], poaching activities [3], and environmental problems [34]. The Movement Ecology field argues that the movement of individual organisms is associated with major problems including habitat fragmentation, climate change, biological invasions, and the spread of pests and diseases [30]. A major challenge in movement ecology research is to explicitly link the statistical properties of movement patterns to specific internal traits and behaviors [30]. Therefore, enriching location data with actual behavioral statistics will greatly support movement ecology research. For example, the amount of time that an animal spends performing various activities such as eating, fighting, or running can help researchers understand the animal's health, its stress levels, or changes in its environment. Collaring technology has immense potential for research in the field of basic and applied animal ecology [8]. Animals have been collared in many studies for varying research purposes, including animal identification, tracking, and health monitoring. Current collars are usually equipped with Global Positioning System (GPS) sensors. More recently, collars have additionally been equipped with motion sensors such as accelerometers and gyroscopes. Equipping collars with advanced processing techniques supports a wide range of existing and future applications. A long lifetime and smaller size of a monitoring collar require a design with minimal complexity without compromising accuracy. The objective of this work is to meet these requirements by discovering a small subset of features that can be used for real-time, energy-efficient, and robust activity recognition of animals. We select features through a combination of filtering and wrapping techniques – the so-called embedded method. The filter is based on the Chi-squares Test, while, the wrapper is based on Sequential Forward Selection [15]. Since the Decision Tree (DT) and Naive Bayes (NB) classifiers have low power consumption for the predicting phase, we perform Forward selection with the DT and NB classifiers on motion data recorded from various sensor orientations. This results in small efficient subsets of features that are lightweight and robust to sensor orientation, while still providing a very high accuracy value, i.e. above 95 %. We also test the subsets of optimal features on other classifiers such as Neural Network (NN), Support Vector Machine (SVM), Linear Discriminant Analysis (LDA), and k-Nearest Neighbors (k-NN). The results show that the average accuracy is at least 80 % for typical activities such as being stationary, walking, eating, running, and trotting. The high performance signifies that, even though the features set was optimized for either the Decision Tree (DT) or Naive Bayes (NB) classifiers, the optimal features presented in this paper can be reused with other classifiers while maintaining good accuracy.

1.1 Challenges

Monitoring and activity recognition of livestock and wildlife in widespread and remote areas using collar tags is challenging for a number of reasons. First, activity recognition should be executed in real-time, on a collar tag, while activities are being performed, since real-time activity recognition greatly enhances the value of many applications in livestock and wildlife management. Moreover, real-time activity recognition provides an embedded system with context regarding an animal's current activity, which can allow the system to efficiently adapt its resource usage, e.g. the device can sleep when the animal is sedentary. Second, a collar tag should last a lifetime, especially in the case of wildlife monitoring. It is dangerous, expensive, stressful for the animals,

and sometimes impossible to recapture the animals in order to charge or replace the collar tag. This means that the energy consumption of any processing taking place on the collar tag must be kept to a minimum. Third, collar tags have limited computational resources, mainly due to energy, weight, and size constraints. Finally, animal tags are subject to rough environments and intense movement behavior and are likely to shift and rotate throughout the day. Many studies assume that the orientation of a sensor on a body is fixed when classifying activity [19, 36, 54]. However, the variability in sensor-orientation causes significant errors if activity classifiers are sensitive to sensor orientation. Therefore, activity classifiers must be insensitive to sensor orientation.

Numerous features can be calculated from summaries (often expressed as *windows* or *frames*) of raw sensor data. Due to resource limitations of collar tags, it is important to calculate only those features that are most sensitive to changing behaviors and robust to sensor orientation. Thus, complexity should be kept to a minimum, and the trade-off between activity recognition performance and resource efficiency should be carefully evaluated. Complexity can be found in two parts of the application; first, in the calculation of features from the sensor data, which act as datapoints for the classification task, and second, in the machine learning algorithm that is used for this classification. It is our challenge to find an optimal combination of the two, which both minimizes computational expense and maximizes performance.

Human activity recognition through wearables, such as smartphones and smartwatches, is currently a popular research topic [42]. Robustness to sensor orientation has been widely researched in the field of human activity recognition. Sensor orientation independence can be tackled in two ways [42]: (i) using orientation-independent features: to only use those features that are insensitive to orientation [38, 43]; (ii) transformation of the input signal: for which in most cases – for all input data – the coordinate system of the mobile device is transformed into a global coordinate system before activity classification [12, 32]. However, continuous calibration of raw measurement data is computationally intensive and, as aforementioned, should be avoided for animal tags. Studies on human activity recognition which claim to use orientation-independent features often make use of the magnitude of the 3 axes of the accelerometer [38, 43]. However, multiple features can be derived from the 3D vector of the accelerometer, and to the best of our knowledge an extensive analysis of the 3D vector has not yet been performed.

In summary, the monitoring of animal behavior using collar tags in remote areas faces the following main challenges:

- Limited energy supply;
- Large number of features to be used for online activity recognition with limited processing capability of the collar tags;
- Sensitivity to sensor orientation.

1.2 An Overview of our Approach

Our main objective is to discover a small optimal set of features that are orientation-independent and can discriminate between various activities using a low-complexity classifier. Therefore, from the raw data collected by collar tags, we propose the extraction of only the most informative orientation-independent features.

Extracted features are used locally – by the collar tag – to classify the animal activities using a lightweight classification approach such as the Decision Tree. Summary results of activity recognition are then sent to a sink node, using a low-cost and long range communication link such as a Low Power Wide Area Network (LPWAN), that has recently become popular in the Internet of Things (IoT) paradigm. By doing so we aim to resolve the problem of energy consumption and communication overhead.

To evaluate this approach, analysis was performed on a total of 126 features, captured by six 3D accelerometers and 3D gyroscopes in different orientations on five farm animals. As a result, global individual movements and

herd interaction can be analyzed by a central server using an advanced inference approach that utilizes the characteristics which were derived locally on the tags.

1.3 Contributions

The main contributions of this paper are:

- This work extends research on human activity recognition and provides a robust technological basis for many applications in movement ecology, livestock management, and wildlife conservation
- We reveal features that are orientation-independent, robust, and suitable for implementing animal activity recognition on real-time embedded platforms
- We show the significance of inertial sensors over orientation sensors in activity recognition performance
- We evaluate the performance of the selected features on 5 different goats
- We develop a video-based toolkit for labeling activities on motion sensor data
- We verify the classification performance of 6 different classifiers with our feature sets
- We make our dataset publicly available at [18]

The rest of the paper is organized as follows: Firstly, related work is discussed in Section 2, after which data acquisition and pre-processing methods are described in Section 3. Subsequently, the proposed feature selection methodology is detailed in Section 4, before the results of various evaluations are presented in Section 5. Finally, conclusions are drawn in Section 6.

2 RELATED WORK

In recent years, there has been a considerable rise in interest in activity monitoring of livestock and wildlife using sensors and embedded devices. Existing approaches that identify animal behavior rely on data-loggers, the subsequent collection of data, and centralized processing [6–9, 14, 21, 24, 25, 31, 41, 46, 48, 50, 53]. In real-world applications, these approaches require transferring data to a central location. However, the transmission demands high bandwidth which dramatically reduces the precious battery life of a collar tag due to the high energy consumption of radios. Recent studies acknowledge the potential of collaring applications and have evaluated offline activity recognition of cows [7, 9, 14, 25, 48], sheep [24, 46], and vultures [31]. Smith et al. [44] studied features in cattle behavior models, using a greedy search to identify feature subsets that were most effective in classifying activities of steers. The authors elaborated on this work [45] and classified five general activities of dairy cows. By means of an ensemble method, the authors trained multiple binary classifiers for each activity independently. Embedded feature selection was used to select a subset of features that produced the best performance for each activity. The orientation of the sensor attached to the cows was fixed in both studies [44, 45] and the authors did not consider orientation-dependency. The goal of the work presented here is to find features that are orientation-independent, to evaluate these features on various classifiers, and to investigate the impact of the feature-subset size. We simultaneously recorded the activities from six unique orientations in order to find features that are truly orientation-independent and evaluate them on data from animals that were not used in the feature selection process (unseen animals).

Real-time activity recognition systems with wearables have been widely studied for humans, and have a certain overlap with animal activity recognition. Yang et al. [52] consider only time-domain features and disregard frequency-domain methods such as FFT and wavelet analysis in order to limit time and processor consumption in their implementation of a real-time activity recognition system. They state that *Mean*, *Root Mean Square (RMS)*, and *Standard Deviation* are the most common and practical time-domain features to be used for activity recognition. They state that *mean* offers both the highest accuracy and the easiest implementation. By calculating only the *mean* value of the accelerometer data, 87.55% accuracy in the classification of five activities was achieved using a Naive-Bayes classifier. Liang et al. [23] built on this by applying a 'two-step feature extraction'

method for classifying 11 activities using data from a smartphone's onboard tri-axial accelerometer. Although the aforementioned papers describe the extraction of basic features such as *mean*, *standard deviation*, and *magnitude* in order to reduce the feature dimension, feature selection remains a useful, yet unexplored, issue to distinguish the most prominent of a large array of features. Zhang et al. [54] investigate feature selection for human activity recognition, stating that high-quality feature selection is 'essential' for the improvement of classification accuracy. They discussed three methods: Sequential Forward Selection (SFS), Relief-F, and Single Feature Classification (SFC). SFS, which was found to deliver the best performance, adds features one by one sequentially, evaluating the accuracy for each sequential combination. This method is the simplest form of greedy feature selection, which is employed by Marais et al. [24] to classify sheep behavior, finding the maximum and minimum values for each axis in a frame as the most important feature out of *mean*, *standard deviation*, *variance*, *skewness*, *kurtosis*, *energy*, *frequency-domain entropy*, *correlation between axes*, and *average signal magnitude*. This combination of features is widely used in activity classification [2, 4, 25, 37]. Previous approaches either select features by using randomly selected training and test data from the full dataset [24], providing similar proportions of each specimen's data in both the training and the test set by visual or statistical analysis [33], or by comparing various feature combinations found in the literature to reach the highest performing combination [16, 36]. However, these methods do not aim to optimize feature sets for robust orientation-independent classification. Lester et al. [22] address the problem of creating a generic algorithm suited to an individual 'out-of-the-box' by training and testing their algorithm on a dataset that contains activity data from a large and diverse group of individual animals. However, they do not address robustness of the selected features.

A number of works have reported the effects of various placements of sensors on the human body [1, 26, 28]. Error in activity recognition due to sensor orientation is often tackled by ensuring that the device orientation is defined in advance relative to the subject's orientation [6, 48, 51]. Ngo et al. [32] use a tilt-correction method and an orientation-compensative matching algorithm is applied in order to solve the remaining relative sensor orientation angle between training and test data. Florentino-Liano et al. [12] transform the measurements from the device's frame into a fixed frame using a rotation matrix. However, these approaches require undesirable additional computational expense. Another solution to this problem is to use orientation-independent features such as the magnitude of an accelerometer as implemented in [43] and [38]. Fixing a device to an animal's body in a particular orientation is impractical and orientation compensating techniques are not always efficient. Therefore, we aim to find features that are orientation-independent and efficient, yielding accurate activity recognition.

3 DATA ACQUISITION AND PRE-PROCESSING

This section first presents a data acquisition system with sensor nodes that comprise motion and orientation sensors such as 3D accelerometers and 3D gyroscopes. A description of how the sensor data are pre-processed for feature selection is then given. All experiments with the animals complied with Dutch ethics law concerning working with animals.

3.1 Data Collection

In order to find orientation-independent features, movement data should be recorded simultaneously by sensors in various orientations. Therefore, data were collected from six sensor nodes that were fixed with different orientations to a collar around the neck of goats. In practice, activity recognition only requires one sensor node, however, the main purpose of this exercise was to investigate the effect of sensor node position and orientation on activity recognition. For our data collection, we recorded the activities of five goats on two farms in the Netherlands. Goats were chosen because they belong to the family of *Bovidae*, together with many other wild animals such as antelope and wildebeest, and are therefore representative of several wild animals. Additionally, goats are widely available in the Netherlands and easy to work with. In total, data were collected from three

domestic pygmy goats and two bigger, more wild goats. The three pygmy goats are shown in Figure 1. Figure 2 shows the two larger goats along with a sensor collar in Figure 2b.

We collected data during 7 days in 2 periods of approximately 4 hours between 08:00 and 17:00 with 2 collars simultaneously and a sampling rate of 100 Hz; data acquisition was bound to these time periods because of the opening times of the petting zoo, safety of the animals (they needed to be watched at all times), and battery lifetime of the sensor nodes. The sensor nodes among the collars were synchronized with a precision of <100 ns. Figure 3 shows a 3D sketch of the collar and the sensor positions with their respective x , y , and z -axes. All efforts were put in to maximize the difference in positions and orientations between the six sensors for all three x , y , and z axis. Since a sensor node placed on the left would give similar data to one placed on the right, due to the symmetry, we allocated sensor nodes on one side of the goat's neck as much as possible. An identical configuration was used for each animal. The collars were loosely attached to the animals and prone to rotation around the animal's neck during the day. When the observer noticed the collar had rotated, it was rotated back to the orientation shown in Figure 3. Therefore, most of the collected data for each sensor can be roughly coupled to the locations denoted in Figure 3 although the sensor locations were not fixed.

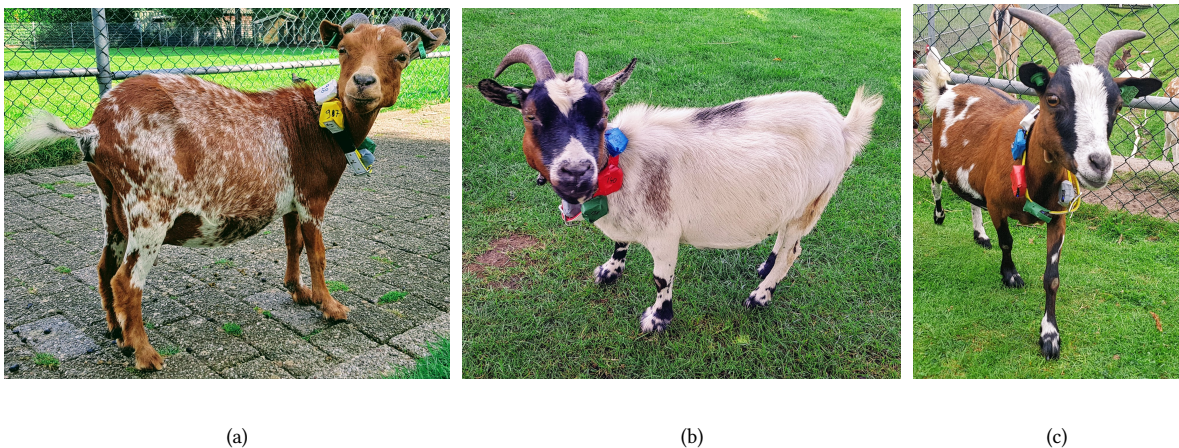


Fig. 1. Three smaller Pygmy goats with attached sensors in different orientations. (a) Goat G1. (b) Goat G2. (c) Goat G3.

We used the ProMove-mini [47] sensor nodes from Inertia Technology, which contain a 3-axis accelerometer and 3-axis gyroscope. Both sensors were sampled at 100 Hz such that it is possible to down-sample the data later on for further investigation. The activities that were observed during the day are listed in Table 1. The animals were recorded on video from various angles during the day. The videos were later used as ground truth for the labeling process.

3.2 Data Labeling

We developed a labeling application using a Matlab GUI [27]. A screen capture of the application is shown in Figure 4. Clock timestamps from the ProMove-mini nodes were used to obtain a coarse synchronization. The labeling application was used to further synchronize videos with sensor data by adjusting the offset. The magnitude of the accelerometer vector $M(t)$, expressed in Equation (1), is displayed to visualize the sensor data. The data can be labeled by clicking at the point representing a change in behavior on the graph. The activity that belongs to the data following the selected point in time can then be selected from a drop-down menu and is added to the graph. A file with activity label and timestamp tuples is instantly updated when an annotation



Fig. 2. Two larger goats. The main purpose is to investigate the effect of sensor position and orientation on activity recognition. (a) Goat G4. (b) Goat G5 with attached sensors at different orientations.

is added. The visualization of the sensor data and the high synchronization achieved with the video allow the annotator to accurately label the activity associated with the sensor data.

All data for all animals were annotated according to the behaviors listed in Table 1. The stop marker for one activity is also the start marker for the following activity, if the following activity is of any other type than *unknown*. Transitions between activities were not excluded from the data, thus some datapoints include a transition phase to another activity. When a goat was performing multiple activities simultaneously, the activity that was mainly exercised was chosen as the label. For example, when a goat was eating while slowly walking, this activity was labeled as *walking*. In order to minimize label bias, all labeled data were visually inspected and corrected by a single person. All efforts were put in to ensure high quality of the labeling process.

The composition and size of the datasets are shown in Table 2. Not all recorded activities were equally often performed in time by all goats. Therefore, the *other activities* class is comprised of the activities with the smaller amounts of data: *fighting*, *shaking*, *climbing-up*, *climbing-down*, *rubbing*, and *food-fight*. It can be seen that goats G1 to G3 do not often perform *trotting* and *running* and spend most of their time *eating* and being *stationary*. The datasets for Goats G4 and G5 are generally more balanced. Segments of data were excluded from the dataset when the activity of the goat for that segment could not clearly be recognized in the video (e.g. when the goat has moved behind an obstacle or other animals). All excluded data that could not be labeled is denoted as *Null* in Table 2.

3.3 Splitting Data for Training and Testing

All sensor data were segmented according to the labeling file generated by the labeling application. One segment holds data for each consecutive activity performed by the animal. Each segment has a different length since the animal performs each activity with a varied time duration. Information leakage may occur when window overlap is used and overlapping windows are concurrently used as training and test datapoints. Therefore, segments,

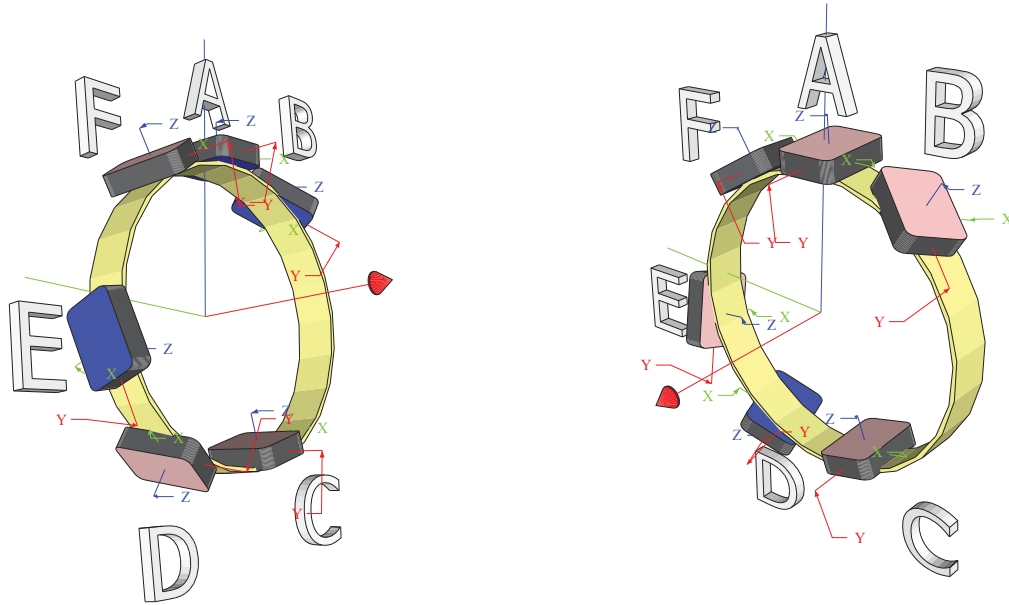


Fig. 3. A 3D sketch of the collar with the attached sensor units seen from two angles. The red arrow denotes the front of the animal. The blue and green axis denote vertical and horizontal axis, respectively. The blue and red colored faces denote the bottom and top side of the sensor nodes, respectively. The sensor nodes have been labeled A to F clockwise.

instead of windows, were divided into training, cross-validation, and test sets for all sensor orientations. The cross-validation sets were used during the feature selection process. The test sets were only used at the end to assess the performance of the final feature set.

K-fold cross-validation [40], also referred to as rotation estimation, ensures that each datapoint has been used at least once as training, cross-validation, and test data. In each iteration k we divided the labeled segments into three respective datasets, rotating over the whole dataset. 10-fold cross-validation was used to obtain 10 respective training, cross-validation, and test sets. This method divides the data into 80% training, 10% cross-validation, and 10% test data. Both the feature selection process and performance assessment were repeated 10 times.

Since some activities may have a long duration, the data among the activities may be unbalanced, which results in a significant bias for the activity learning and testing phase. For example, a *stationary* activity can last either 20 minutes or 5 seconds; if the longer duration is used as training and the shorter as testing, then the activity *stationary* might have a larger ratio of training/testing than other activities. To eliminate this imbalance, each segment was recursively split into smaller segments until each segment had a maximum length of 10 seconds before dividing the segments into the three training, cross-validation, and test data sets. The effect of the splitting is shown in Figure 5. It can be seen in Figure 5a that the ratios between the 3 sets vary for each activity when the data are not recursively split to segments with a maximum length of 10 seconds. Figure 5b shows the proportions when the data were split, it can be seen that the ratios are more balanced between the activities.

Table 1. Observed activities during the day

Activity	Description
Stationary	The animal is lying on the ground or standing still, occasionally moving its head or stepping very slowly.
Eating	The animal is pulling fresh grass out of the ground, eating hay from a pile or twigs/grains on the ground.
Walking	The animal is walking. The pace of walking varies from very slowly to nearly trotting.
Trotting	This is the phase between walking and running. The animal is not galloping rapidly but walking very quickly and is therefore in a trot state.
Running	The animal gallops.
Fighting	The animal is fighting with another animal. This consists of banging its head against another animal's head or body. A goat often stands on the back of its legs, drops itself to the ground and drives its horns into another animal.
Shaking	The animal is shaking its entire body in a very rapid motion, often followed by rapidly shaking its head for a brief moment. On a few occasions, the animal only shakes its head.
Climbing up	The animal is walking/jumping up onto an object. For example climbing into a shelter, or jumping onto a wooden bench.
Climbing down	The animal is walking/jumping down from an object. For example climbing out of a shelter, or jumping off a wooden bench.
Rubbing	The goat is pressing its body against a fence and walking while it keeps pushing itself towards the fence.
Food fight	This occurs when food is dropped in a group of animals. All animals are pushing each other trying to reach the food.

Table 2. Composition of the datasets per goat. The columns denote the durations and proportions of data from 6 sensors per activity.

Activity	G1		G2		G3		G4		G5	
	duration (min)	fraction	duration (min)	fraction	duration (min)	fraction	duration (min)	fraction	duration (min)	fraction
Stationary	944	28.8%	314	19.1%	679	35.4%	1053	37.0%	406	42.2%
Walking	441	13.5%	312	19.0%	227	11.8%	433	15.2%	191	19.9%
Trotting	10	0.3%	16	1.0%	12	0.6%	45	1.6%	24	2.4%
Running	9	0.3%	5	0.3%	6	0.3%	46	1.6%	23	2.3%
Eating	1278	39.1%	639	38.9%	771	40.3%	379	13.3%	202	21.0%
Other activities	72	2.2%	36	2.2%	36	1.9%	4	0.1%	9	0.8%
Null	519	15.9%	322	19.6%	187	9.7%	889	31.2%	110	11.4%
Total (min)	3273		1644		1918		2849		965	
Total (hrs)	54.6		27.4		32.0		47.5		16.1	

After the division of segments into three groups, each segment of data were separated into windows of data. We used a window size of 2 seconds and a 50 % window overlap. For each window, calculation of the features described in Section 4.1 was performed.



Fig. 4. Screenshot of the labeling application

4 FEATURE SELECTION

This section details the proposed approach to find orientation-independent features. During the feature selection process, the data from goat G4 were used because this dataset contains both the largest quantity of data and the best balance between the activities. A graphical representation of the approach is shown in Figure 6, wherein each colored box represents an inner loop in the process. First, we select only 3D-vector features since they are theoretically robust to sensor orientation [43]. Second, for each fold k of the data with selected 3D-vector features, a filter method is applied in order to suppress the least interesting features. The core of our feature selection is an embedded method, which selects features through simultaneous feature selection and classification [15]. The embedded method comprises a wrapper that uses classification performances as criteria to select the most informative features. The classifier is used within the wrapper as a black box to weight feature subsets based on their recognition performance. The embedded method was executed with the Forward Selection scheme for all 10 folds with various maximum feature set sizes denoted by n in the range of $\{1, \dots, 10\}$ so that a total of 100 instances of $F_{k,n}$ were obtained. Finally, the feature selection results of 10 folds were used to select the most optimal feature set for each size n . The aforementioned steps are described in more detail in the following subsections.

4.1 Feature Calculation

A selection was made consisting of time and frequency-domain features that are typically used for activity recognition [4, 10, 13, 16, 19, 24, 25, 36, 37, 42, 54]. Although frequency-domain features are more complex to calculate than time-domain features [11], no features were excluded because of their higher complexity because this study aims to perform an exploratory analysis and to find those features that are most robust to sensor orientation. Thus, Table 3 denotes the features that were calculated for each window of data. The total number of

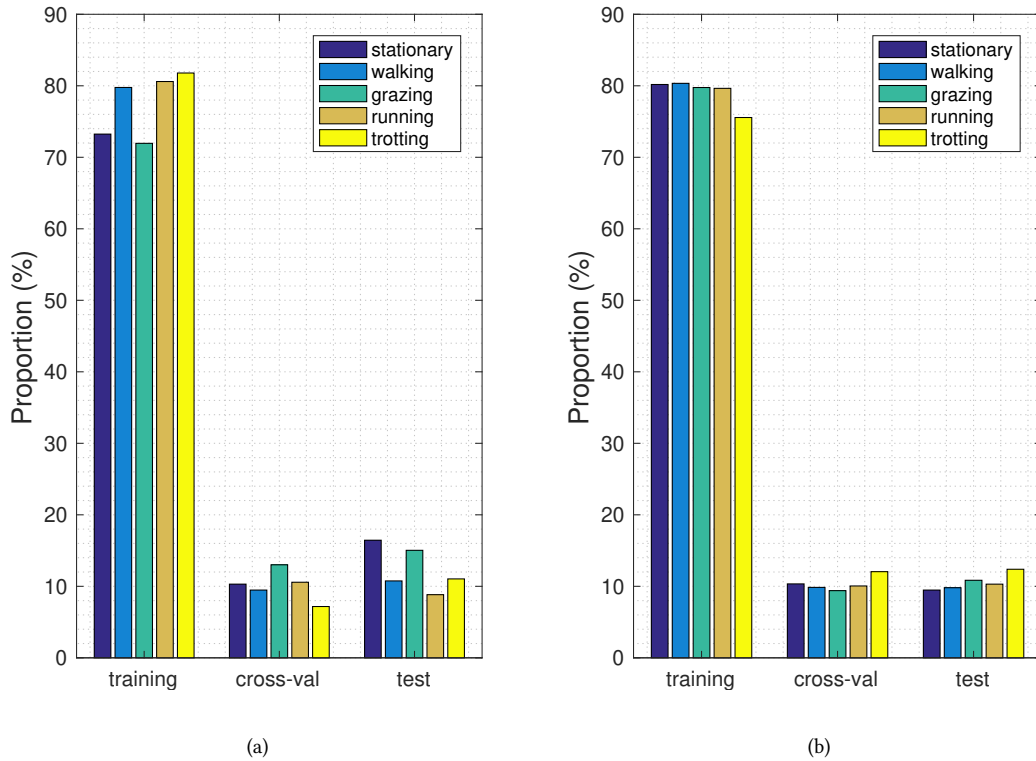


Fig. 5. Distribution of data over training, cross-validation, and test sets. (a) No recursive splitting of segments. (b) Recursive splitting of segments to maximum length of 10 seconds

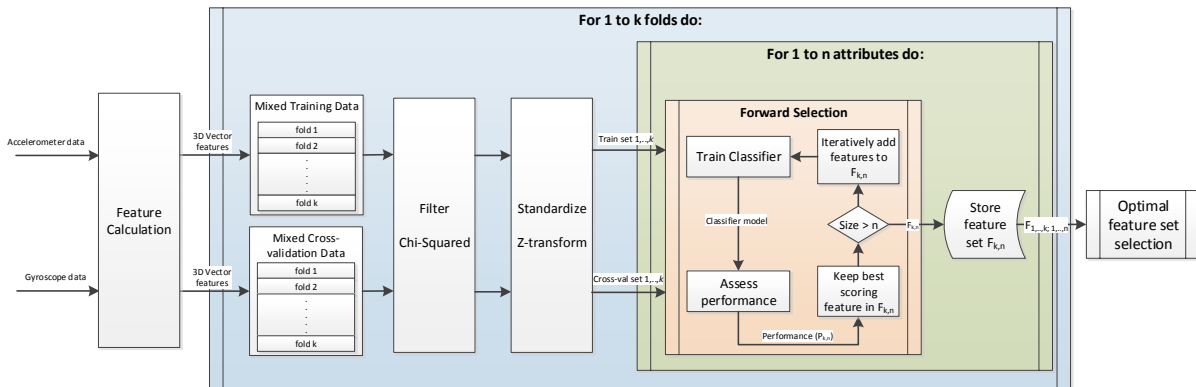


Fig. 6. Embedded feature selection process. Each colored box denotes an inner loop

126 features is calculated by multiplying the number of derived features with the number of sensors and axes used in the selection process.

Table 3. Features that were calculated for each window of data from all sensors and all their axes

Feature	Description	Number of features
Maximum	Maximum value	6
Minimum	Minimum value	6
Mean	Average value	6
Standard deviation	Measure of dispersion	6
Median	Median value	6
25 th percentile	The value below which 25 % of the observations are found	6
75 th percentile	The value below which 75 % of the observations are found	6
Mean low pass filtered signal	Mean value of DC components	6
Mean rectified high pass filtered signal	Mean value of rectified AC components	6
Skewness of the signal	The degree of asymmetry of the signal distribution	6
Kurtosis	The degree of 'peakedness' of the signal distribution	6
Zero crossing rate	Number of zero crossings per second	6
Principal frequency	Frequency component that has the greatest magnitude	6
Spectral energy	The sum of the squared discrete FFT component magnitudes	6
Frequency entropy	Measure of the distribution of frequency components	6
Frequency magnitudes	Magnitude of first six components of FFT analysis	36
Total		126

Our goal is to find orientation-independent features, therefore, a 3D vector was calculated from the sensors' individual axes. The orientation-independent magnitude of the 3D vector is defined as:

$$M(t) = \sqrt{s_x(t)^2 + s_y(t)^2 + s_z(t)^2}, \quad (1)$$

where, s_x , s_y , and s_z are the three respective axes of the sensor. $M(t)$ was calculated from both the gyroscope and accelerometer data. All data were standardized by means of a Z-transformation, obtaining a standard score of each feature value. Standardization does not affect the performance of Decision Tree (DT) and Naive Bayes (NB) classifiers.

4.2 Filter: Chi-Squared Test

Since filters are faster than wrapper methods they can be used as a pre-processing step to reduce space dimensionality and overcome over-fitting. Moreover, filters can provide a more generic feature set that is not specifically tuned for a given classifier [15]. In order to determine the relevance of each feature with respect to the activity label, the Chi-Squared test [35] was applied to the feature set. The Chi-Squared test is a statistical technique that is used to determine the independence between features. The test determines if a distribution of observed frequencies differs from the theoretically expected frequencies [35]. The Chi-Squared statistic is expressed as

$$\chi^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i}, \quad (2)$$

where O_i is the observed frequency and E_i is the expected frequency. The irrelevant features were excluded prior to searching for the optimal set with an embedded approach; a weight was calculated for each feature and the 30 % lowest scoring features were excluded.

4.3 Embedded Forward Selection

After applying the Chi-Squared filter, the remaining features were used as the input for the Embedded Forward Selection method. Forward selection starts with an empty set of features F and tests the performance of the classifier with each isolated feature of the given dataset. Only the feature that gives the highest increase in performance is added to the subsequent feature set F . This set is then used as the basis for the next round of adding features until one of the stopping criteria, detailed in the following paragraph, was met. The Decision Tree (DT) and Naive Bayes (NB) classifiers are both known to have good performance and low complexity [20], therefore, we consider them as the two most promising classifiers for real-time activity recognition and used them within the Forward Selection. Forward Selection was performed separately twice, first a DT was used within the embedded method and second a NB classifier, resulting in two feature sets.

In order to investigate the effect of the number of features used, denoted as n , on the recognition performance, we used a maximum number of features as the first stopping criterion. By incrementing the maximum number of features from 1 to 10, 10 feature sets with incremental sizes were obtained for both the DT and NB classifier. The second stopping criterion was defined as a minimum increase in performance. The iteration was terminated when the absolute performance did not increase at all, so that the first criterion was enforced as much as possible. Three speculative rounds were implemented in order to avoid getting stuck in local optima.

During the forward selection process, the classifier was trained using merged training data from all 6 sensors with various orientations. Simultaneously, the classifier was tested with merged cross-validation data from all 6 sensors. The datapoints were always shuffled after merging to prevent bias of the classifier towards a single orientation. Due to the fact that the classifier was trained with mixed data from all orientations, orientation-dependent features would not perform well. Therefore, the result of the feature selection process is a feature set that is robust against sensor-orientation.

The process described above was repeated ten times by using a 10-fold cross-validation of our sensor datasets, leaving out a test set for each fold. This approach resulted in 10 feature sets F for both the DT and NB classifiers and for each size n of F in a range of $\{1 \dots 10\}$. Ultimately, the test sets were only used in order to assess the performance of the final feature set.

Most research related to activity recognition is focused on the accuracy performance of a technique. However, there are other criteria to be considered when selecting a technique [20], including: i. CPU and memory complexity, ii. sensitivity to irrelevant features, iii. sensitivity to continuous versus discrete features, iv. sensitivity to noise, v. bias and variance of classifiers, vi. storage space required during training and classification stages, vii. possibilities for use as an incremental learner (online Machine Learning (ML)), viii. ease of use, related to the number of classifier or runtime parameters to be tuned by the user, and ix. comprehensibility of a classifier.

Because our study focuses on real-time activity recognition for animals, low complexity, memory footprint, and accuracy are equally important. Ease-of-use and comprehensibility of the classifier greatly ameliorate effective implementation in a low-cost and energy-efficient embedded system. To show the applicability of the selected features on different classifiers, the performance of both feature sets was assessed and compared for the following classifiers: i. Decision Tree (DT), ii. Neural Network (NN), iii. Support Vector Machine (SVM), iv. Naive Bayes (NB), v. Linear Discriminant Analysis (LDA), and vi. k-Nearest Neighbors (k-NN). A description of the functionality and properties of these classifiers is presented in [20]. The performance results are shown and discussed in Section 5.

4.4 Optimal Feature Set Selection

Finally, the optimal feature set was selected from the results of the 10 folds. In order to study the effect of the number of features (size of n) on the classification performance we selected 10 optimal feature sets incremental in sizes $n = \{1 \dots 10\}$ for both the DT and NB classifier. The optimal feature set selection algorithm is shown in Algorithm 1. The embedded selection process resulted in a 3-dimensional binary feature weight map. The 3 dimensions are all features H , feature-set-sizes n , and folds k . For every set-size n in the binary map F , Algorithm 1 selected the features that were most often selected over 10 folds and included them in the final optimal feature set $U_{opt_{1\dots n}}$.

ALGORITHM 1: Optimal feature set selection

Data: $F_{1\dots H, 1\dots n, 1\dots k}$, F is a binary feature weight map for all features H , set-sizes n , and folds k
Result: $U_{opt_{1\dots n}}$, the optimal feature sets for all sizes n

```

for  $n = 1$  to 10 do                                     /* where  $n$  is the feature set size */
  for  $k = 1$  to 10 do                                       /* where  $k$  is the fold */
    for  $i = 1$  to  $H$  do                                       /* where  $H$  is the total number of features */
      if  $F_{i, n, k}$  then
         $S_{i, n} = S_{i, n} + 1$ ; /* count selections for feature  $i$  in feature-set-size  $n$  over folds  $k$  */
      end
    end
  end
end
for  $n = 1$  to 10 do
   $sort(S_{1\dots H, n}, \text{descending})$ ;                               /* sort the feature scores */
   $U_{opt_n} = S_{1\dots n, n}$ ;                                       /* pick the  $n$  best scoring features */
end

```

5 EVALUATION

In this section, the results of the feature selection are presented, and subsequently the orientation-independent activity recognition performance of the DT and NB optimal feature sets.

All classifiers were implemented in RapidMiner [29], and no performance fine-tuning was performed for any of the classifiers. We used information gain as the splitting criterion for the DT, with a maximal tree depth of 10, and application of both pruning and pre-pruning. The NB classifier used Laplace smoothing to prevent high influence of zero probabilities. For the remaining classifiers we used generic parameter settings.

5.1 Evaluation Criteria

During feature selection, the performance of the classifiers was evaluated by the accuracy measure, denoted as

$$\text{accuracy} = \frac{tp + tn}{tp + tn + fp + fn}, \quad (3)$$

where tp denotes true positives, tn denotes true negatives, fp denotes false positives, and fn denotes false negatives. In the context of activity recognition tp denotes the number of true datapoints from activity α that were also classified as activity α ; tn denotes the number of datapoints from all other activities ($\neg\alpha$) that were not classified as activity α ; fp denotes the number of datapoints from all other activities ($\neg\alpha$) that were classified as activity α ; fn denotes the number of true datapoints from activity α that were classified as any of the other activities. These values were obtained from the confusion matrix of each evaluation. For a given activity α , an

ideal classifier correctly identifies all true datapoints (i.e. 100 % tp) and also guarantees that when a datapoint is classified as activity α it is in fact not another activity than α (i.e. 100 % tn). Therefore, the accuracy, expressed in Equation (3), which takes both tp and tn into account, was used during the feature selection. Often the F1 measure is used for evaluation purposes and is expressed as

$$F1 = \frac{2 \cdot tp}{2 \cdot tp + fp + fn}. \quad (4)$$

We used the accuracy measure for feature selection because the F1 score does not take tn into account and we consider tn to be important for activity recognition. Additionally, the overall performance of the classifiers is also evaluated with the recall, precision, and F1 measures in Section 5.3.4. Recall and precision are expressed as

$$\text{recall} = \frac{tp}{tp + fn}, \quad (5)$$

$$\text{precision} = \frac{tp}{tp + fp}. \quad (6)$$

For a given activity α , precision only takes into account the quantity of classified datapoints (the $tp + fp$). Recall takes all the datapoints of activity α into account (the $tp + fn$). Recall is also referred to as sensitivity or true positive rate. When activity α has a high recall, this means that the classifier classified most of the datapoints of activity α . Precision is also referred to as Positive Predictive Value (PPV). High precision means that substantially more classified datapoints were correct than incorrect. The F1 score is also referred to as F-score or F-measure and can be interpreted as a weighted average of the precision and recall. An F1 score of 1 is optimal and 0 is worst.

5.2 Feature Selection

Figure 7a shows the selected features when the DT classifier is used within the wrapper. Each color represents a feature and the size of the box represents the number of times that the feature is selected over 10 folds. When a feature is selected less than 10 times over all 10 folds and feature set sizes n , it is considered to be insignificant and is not shown in Figure 7.

Table 4. Optimal feature sets for DT and NB classifiers. *acc* and *gyro* are abbreviations for *accel_3dvector_norm* and *gyro_3dvector_norm*, respectively

Size(n)	1	2	3	4	5
DT	accel_mag_6	accel_mag_6	accel_mag_6	accel_mag_6	accel_mag_6
		accel_std	accel_freqEntropy	accel_freqEntropy	accel_freqEntropy
			accel_std	accel_std	accel_std
				accel_twenty_fith_p	accel_mag_1
					accel_twenty_fith_p
NB	accel_std	accel_std	accel_std	accel_std	accel_std
		accel_freqEntropy	accel_freqEntropy	accel_freqEntropy	accel_freqEntropy
			gyro_twenty_fith_p	gyro_twenty_fith_p	gyro_twenty_fith_p
				accel_median	accel_twenty_fith_p
					accel_median

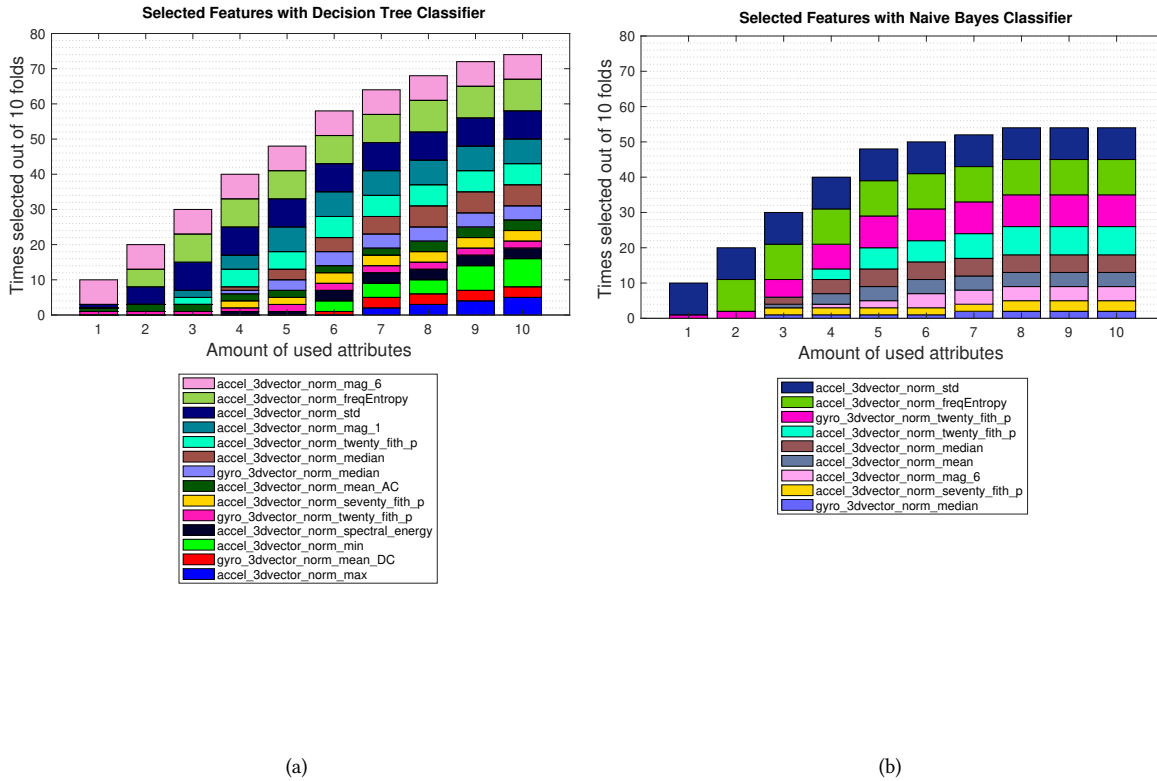


Fig. 7. Optimal feature sets for each number of features that are included. The size of each box denotes the number of times that feature is included in the feature set over 10 folds. (a) Selected features using Decision Tree classifier. (b) Selected features using Naive Bayes classifier

Forward Selection with the DT classifier selects various features over 10 folds when the size of the feature set is $n > 1$; this means that the DT classifier is less sensitive to the variety of features. Figure 7a shows that the 6th magnitude in the norm of the accelerometer’s 3D vector is the best-performing feature to use with the DT classifier if a single feature is used. The optimal feature set comprises the features that are most often selected over 10 folds and is shown in Table 4, in which each column denotes the feature set with size n . When a box in Figure 7 is similar in size to the largest box, this means that interchanging some of the features in our optimal set (Table 4) with these other high-scoring features in Figure 7 will not significantly degrade the performance. For example, the standard deviation of the accelerometer’s 3D vector performs equally well as the frequency entropy feature. The optimal feature sets were used to assess the performance on both the mixed data of all sensors and each individual sensor. These performance results are presented in Section 5.3.

The same feature selection is performed when the NB classifier is used within the embedded method. The selection results are shown in Figure 7b. The results show that NB is more sensitive to the selected features than DT because the same features are selected more consistently over the 10 folds. It is well known that NB is sensitive to irrelevant features [17]. Therefore, the embedded approach selects features more consistently for NB than for DT. The optimal feature set with the NB classifier is shown in Table 4. The gyroscope’s 25th percentile outperforms the accelerometer features with a minor difference. Thus, the performance will degrade minimally, if

any, when the gyroscope is not used for activity recognition. Figure 7b shows that the *gyro_twenty_fith_p* feature can be exchanged with either the *accel_twenty_fith_p* or *accel_median*.

5.3 Optimal Feature Set Performance

This section presents the classification performances when the data were characterized with the feature sets shown in Table 4.

5.3.1 Evaluation of Individual Orientations and Positions. The aim of the experiment presented here is to investigate the difference in classification performance when a single accelerometer is used on a collar at different positions and in diverse orientations. Moreover, the results presented here not only show the impact of different feature subset sizes, they also give an insight on the effect of individual features in these optimal subsets.

In this experiment we used the data from goat G4 because this dataset contains both the largest quantity of data as well as the most even balance between activities. The performance evaluation on the other goats is discussed in Section 5.3.4. In order to show that the selected features are truly orientation-independent, the performance of the classifier is assessed with data from each sensor node individually. All sensor nodes simultaneously recorded the same number of activities but did so from diverse orientations and locations, as shown in Figure 3. Both DT and NB feature sets were used to characterize the training and test sets of each individual sensor's data. Both the DT and NB classifiers were trained and tested with data from each individual sensor. The performances of both classifiers, shown in Figure 8, are very high and the variance among the sensors is low. The minor difference between the sensors' performances can be explained by the different locations of the sensors. Figure 8a shows that position E and C benefited from the *accel_median* feature that was included in the subset with 6 features. This feature is not included in the smaller subset sizes because the feature selection was optimized for all orientations simultaneously and not for individual location performance. Figure 8b shows that the NB classifier selected features that perform more consistently throughout all positions. Because NB classifiers are known to be very sensitive to the presence of redundant and/or irrelevant attributes [5], they typically select more consistent features that perform well on average over all positions. For both classifiers, the sensors C and F are the best overall performing sensors. These sensors were mostly positioned at the bottom and top of the goat's neck, respectively. This indicates that the best location on the neck for animal activity recognition is either at the top or bottom of the neck. The difference in performance among various orientations and locations is small ($\pm 2\%$ with 3 features), and activity recognition will be robust against rotation of the animal tag around the animal's neck.

5.3.2 Evaluation of DT and NB Classifiers with data from Diverse Sensor Orientations. The goal of this experiment is to assess and compare the performance of DT and NB classifiers which were tested with data from diverse orientations. A single DT and NB classifier were trained with mixed data from all six sensor orientations (A-F) with diverse orientations. The optimal feature sets were used to characterize the mixed data's training and test sets. The results are presented in Figure 9, which shows that both classifiers are able to classify activities with high accuracy using just a single feature. The results in Figure 9 show that the optimal subset size is 3 features and adding more than 3 features does not increase the performance of the classifiers. NB performs slightly better than DT for all feature set sizes.

In order to evaluate the validity of our feature sets, they were compared with randomly-drawn feature sets, excluding the optimal set. The results are shown in Figure 10. The performances of the random feature set are always lower than the optimal feature set which shows that our feature sets are indeed optimal. The difference in performance between the optimal and random feature set is larger for the NB classifier. This is as expected because NB is more sensitive to irrelevant features than DT.

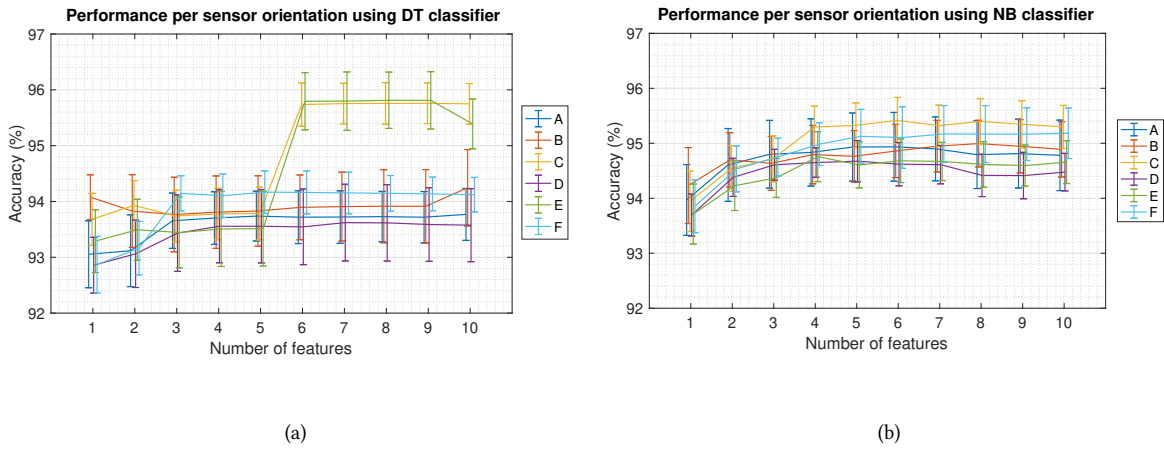


Fig. 8. Performance for each sensor orientation. Error bars denote standard deviation over 10 folds. (a) Performance using DT classifier and DT feature set. (b) Performance using NB classifier and NB feature set.

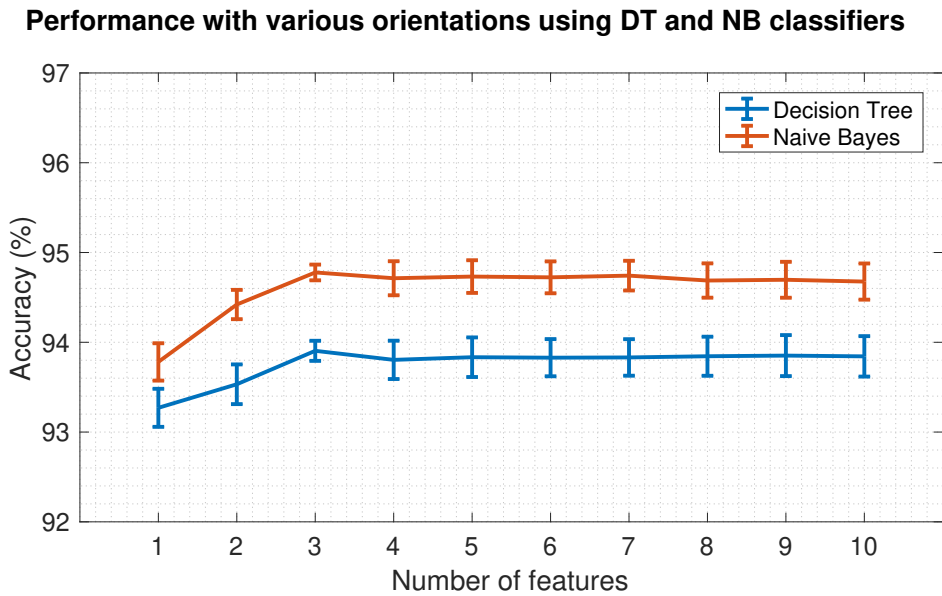


Fig. 9. Performance for mixed data set with DT algorithm using DT feature set and NB algorithm using NB feature set. Error bars denote standard deviation over 10 folds

5.3.3 *Assessing the Performance of the Features with Various Types of Classifiers.* The aim of this experiment is to evaluate the performance of various types of classifiers other than DT and NB. We used our optimal feature

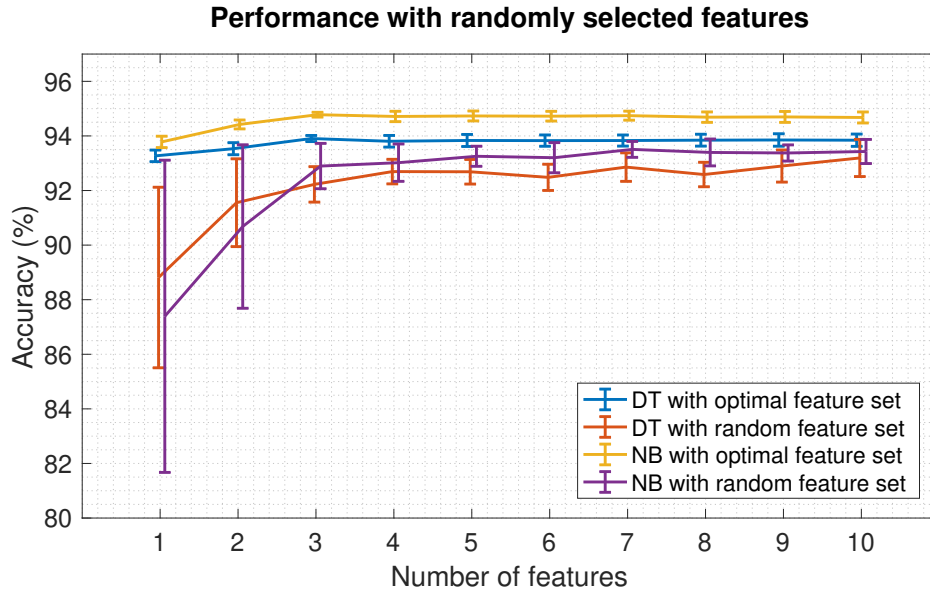


Fig. 10. Performances of optimal feature set and randomly selected feature sets. The mixed data set was used. Error bars denote standard deviation over 10 folds

sets to characterize the mixed-orientation activity data and assessed the performance of the features and the effect of the subset size with 6 types of classifiers.

Figures 11a and 11b show the performance of 5 classifiers for different sizes of the feature sets. The results in Figure 11a are obtained with the DT feature sets and Figure 11b shows results that are obtained with the NB feature sets. The fact that performance increases with the use of more features signifies that the feature set is not optimal for these classifiers. An optimal feature set can only be found when feature selection is tuned for that specific classifier. The NN classifier is expected to be among the best performing algorithms and Figure 11 shows that the optimal sets perform at least as good on a NN than on the classifiers they were selected with in the embedded method. Figure 12 shows that the genericity of both feature sets is very similar except when $n = 1$. The NB feature set includes more generic features than the DT feature set (both LDA and SVM perform better with 2 features because the *accel_freqEntropy* is included).

The general performance does not significantly increase when more than 3 features are included. For all the classifiers, except LDA and SVM, the performance is already 93 % with just 1 feature.

5.3.4 Evaluation of Feature Sets on Unseen Goats. The aim of the experiment described in this section is to evaluate the performance of the optimal feature sets on the goats that were not used in the feature selection process. The optimal feature sets were tested on 4 other goats that wore the collar. The optimal feature sets were also tested with unseen data of the goat that were used for our feature selection. The activities of each goat were recorded concurrently from 6 positions in various orientations, as shown in Figure 3. For each animal, the classifiers were trained and tested with data from that animal, using the optimal feature sets. Training and testing a classifier with data from various individuals would be possible through methods such as leave-one-subject-out, which are generally used to assess the generalization of classifiers over multiple individuals; however, such an

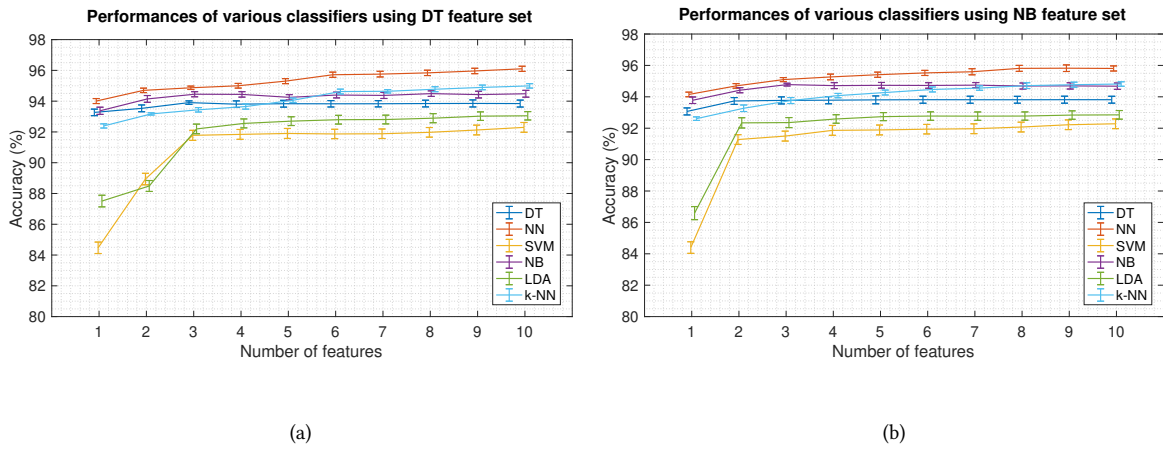


Fig. 11. Performances of multiple algorithms using the mixed data set. Error bars denote standard deviation over 10 folds (a) using DT feature set. (b) using NB feature set

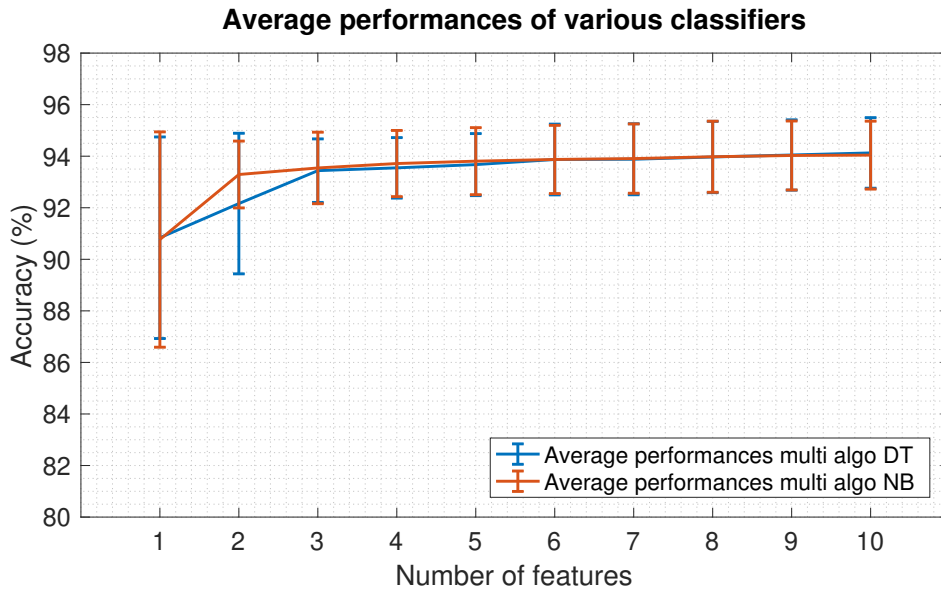


Fig. 12. Average performances of multiple algorithms with DT and NB feature feature sets. Error bars denote standard deviation over 6 algorithms

assessment brings up a range of issues regarding generic activity recognition, which is beyond the scope of this paper. In what follows, we discuss the evaluation of the optimal features on unseen goats with both the DT and NB classifiers.

The performance of the DT and NB classifiers can be seen in Figures 13a and 13b, respectively. Because the data from G4 were used for the feature selection process, the features are most optimal for this animal. G5 is physically most similar to G4 and Figure 13a shows that the performance of G5 is even a little higher than G4. Table 2 shows that the *eating* activity is better balanced for G5 and has a 6 % higher accuracy (shown in Table 8). The difference in accuracy for the *eating* activity of G5 explains the overall, although minor, higher performance for G5. G1, G2, and G3 are all three goats from a different family. They are smaller, have shorter legs, and do not trot and run as much as the more wild goats. We believe these physical differences to be the main reason of the drop in performance for goats G1, G2, and G3. However, the accuracy of activity recognition of these goats is still above 89 % while the classifiers' parameters and data were not tuned or optimized for multiple subjects. The results in Figures 13a and 13b show that the NB subset was able to find an optimal set for all animals with 3 features as the performance does not significantly increase with more features. With the DT subset the *gyro_3dvector_norm_median* feature was included at $n = 8$ and improved the performance of goat G1 and G2 with 3 %. Using a gyroscope consumes more energy, therefore this is a trade-off between accuracy and power consumption. However, the results of the other goats' performances show that using a gyroscope is not necessary if the classifier is properly tuned for a given or multiple individuals. Power consumption and other considerations for real-time implementation are further discussed in Section 5.3.5.

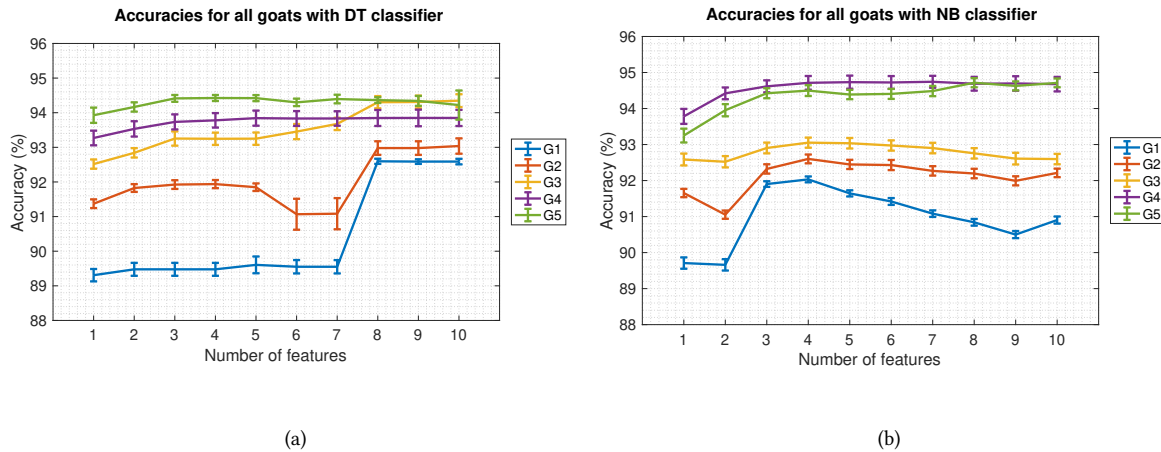


Fig. 13. Performances of DT and NB classifiers using the mixed data set of 5 different goats. Data from goat G4 was used exclusively during the feature selection process. (a) using DT feature set. (b) using NB feature set

We evaluated the performance of 6 types of classifiers with test-data from all 5 goats. The classifiers were evaluated with both the DT and NB optimal feature subsets with 3 features. The results in Figure 14 show that the relative differences in performance is similar for all goats. The best overall performance was obtained with data from G4 because the feature sets were selected with cross-validation data from this animal. For all the goats other than G4 the classifier that was used during feature selection performed the best, closely followed by the NN classifier. The results show that the NN classifier generalizes well over multiple animals. The highest overall accuracy for all classifiers and animals was obtained with the NB feature set.

The confusion matrices of the DT and NB classifiers for goat G4 are shown in Tables 5a and 5b, respectively. The classification performances in terms of accuracy, recall, precision, and F1 scoring are presented in Table 6. The confusion matrices and performances for goat G5 are shown in Tables 7 and 8, respectively. The optimal feature sets with $n = 3$ were used to characterize the data. The values in the confusion matrices are summed up

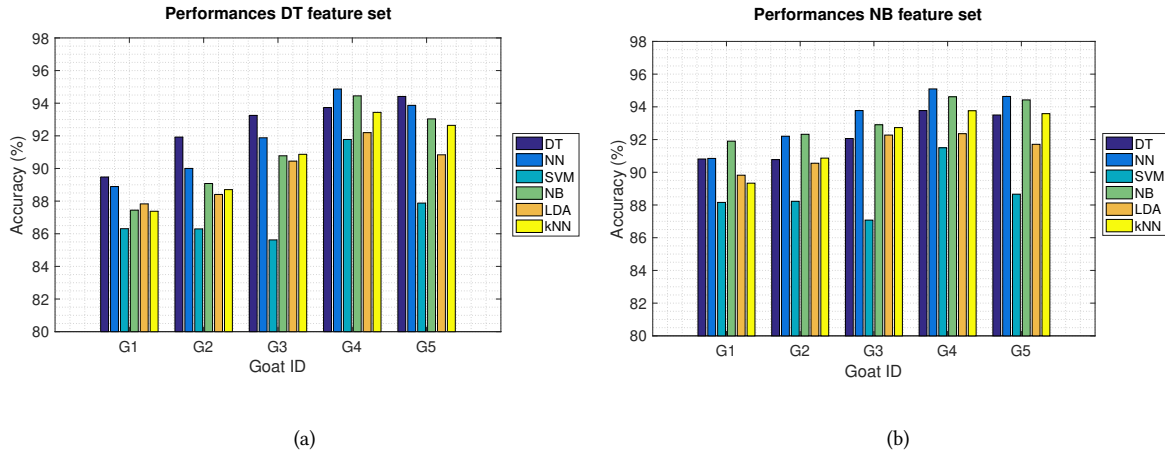


Fig. 14. Performances of multiple algorithms using the mixed data set of 5 different goats using 3 features. (a) using DT feature set. (b) using NB feature set

over 10 folds. The numbers in Tables 5 and 7 express the number of test datapoints that have been classified as that activity. Each datapoint represents a 2-second window of raw sensor input data. The sum of each column denotes the total number of true datapoints for that activity over 10 folds.

The results in Table 6 show that both classifiers perform similar with average accuracies of 93.90 % and 94.78 %, respectively. For both goats there is no overall significant difference in performance between the two classifiers other than NB performing slightly better (0.8 % higher F1 score). The DT outperforms NB for the activities *walking*, *eating*, and *trotting*, as is reflected in the higher F1 score. For the activity *eating* DT has a higher recall and lower precision than NB; this means that more datapoints of this activity were recognized by the DT, but the ratio of correctly classified datapoints was higher with NB.

Because the classifiers' performances are very similar for both goats, in the following we combine the evaluation of individual activities. Overall, all activities are well detected, except for the *other activities* activity. The activities *stationary* and *walking* have the best F1 scores and are well detected and distinguished from other activities. The activity *trotting* has the single lowest F1 score and Table 5 shows that this activity is mostly confused with *running* and vice versa. This is as expected because the activities are similar and it is often very difficult to distinguish between the two activities during labeling; the animal is often rapidly changing from *trotting* to *running* and vice versa. The activities *stationary*, *walking*, and *eating* are mostly confused with each other. This can be explained because the goat often combines these activities; e.g. during *eating* the goat can be *stationary* while chewing, and often is *walking* while *eating*. The activities in the *other activities* activity are mostly classified as either *trotting* (G4) or *walking* (G5). Tables 6 and 8 both show that the NB classifier discriminates *other activities* activity datapoints better than DT. Improved discrimination between the activities can be achieved with more tuning of the classifiers and is outside the scope of this paper.

5.3.5 Considerations for Real-Time Activity Recognition on a Collar Tag. Executing activity recognition in real-time (during the activities) and locally (on the collar tag) supports numerous applications in areas such as wildlife conservation, livestock management, and ecology research. Local processing faces the challenge of limited energy supply, processing power, and transmission bandwidth on collar tags. However, local activity recognition will significantly prolong the battery life because large amounts of data do not have to be transmitted,

Table 5. Confusion matrices of DT and NB classifiers using 3 features for goat G4. The values are summed up over 10 folds.

(a) DT classifier							(b) NB classifier						
Predict \ True	True Stationary	True Walking	True Eating	True Running	True Trotting	True Other Activities	Predict \ True	True Stationary	True Walking	True Eating	True Running	True Trotting	True Other Activities
Predict Stationary	93807	22	5471	0	0	0	Predict Stationary	113204	24	17792	0	0	0
Predict Walking	372	50773	914	22	187	654	Predict Walking	401	50038	1143	18	101	564
Predict Eating	32028	876	39000	0	0	7	Predict Eating	12581	630	26404	0	0	0
Predict Running	0	0	0	4622	463	292	Predict Running	0	0	0	4814	599	132
Predict Trotting	0	158	0	794	4684	891	Predict Trotting	0	899	0	447	4500	1039
Predict Other Activities	0	5	0	0	8	6	Predict Other Activities	21	243	46	159	143	115

Table 6. Performance of both DT and NB classifiers for goat G4

	Classifier	Stationary	Walking	Eating	Running	Trotting	Other Activities	Average Performance
Accuracy	DT	83.95%	98.64%	83.35%	99.33%	98.94%	99.21%	93.90%
	NB	86.94%	98.30%	86.36%	99.43%	98.63%	99.01%	94.78%
Recall	DT	74.33%	97.95%	85.93%	84.99%	87.68%	0.32%	71.87%
	NB	89.70%	96.54%	58.18%	88.53%	84.22%	6.22%	70.56%
Precision	DT	94.47%	95.94%	54.23%	85.96%	71.76%	31.58%	72.32%
	NB	86.40%	95.74%	66.65%	86.82%	65.36%	15.82%	69.46%
F1	DT	83.20%	96.94%	66.50%	85.47%	78.93%	0.64%	68.61%
	NB	88.02%	96.14%	62.13%	87.66%	73.60%	8.93%	69.41%

Table 7. Confusion matrices of DT and NB classifiers using 3 features for goat G5. The values are summed up over 10 folds.

(a) DT classifier							(b) NB classifier						
Predict \ True	True Stationary	True Walking	True Eating	True Running	True Trotting	True Other Activities	Predict \ True	True Stationary	True Walking	True Eating	True Running	True Trotting	True Other Activities
Predict Stationary	41528	1	2733	0	0	0	Predict Stationary	42490	28	3329	0	0	40
Predict Walking	149	21920	898	12	72	3571	Predict Walking	143	21596	915	10	28	3462
Predict Eating	6870	693	20445	0	0	341	Predict Eating	5909	770	19780	0	0	236
Predict Running	0	3	0	1923	595	401	Predict Running	0	0	0	1852	424	217
Predict Trotting	0	181	0	723	2113	410	Predict Trotting	0	332	0	572	2133	656
Predict Other Activities	1	26	5	0	21	17	Predict Other Activities	6	98	57	224	216	129

which typically consumes more energy than local data processing. Moreover, online activity recognition enables the collar tag to efficiently adapt its resource usage to a situation (e.g. the device can sleep when an animal is sedentary). A long lifetime and smaller size of a monitoring collar require an activity recognition system with minimal complexity. Three components that contribute to the energy consumption of an activity recognition system are: i. the quantity and type of sensors that are utilized, ii. the complexity of the classifier's inference phase, and iii. the complexity and quantity of the features. In the following we will discuss these three components.

Gyroscopes typically consume more energy than other sensors because they continuously vibrate at a certain frequency in order to measure the angular velocity [49]; thus, it is better to not use this sensor. The results in Tables 4a and 4b show that the optimal DT subset does not include any gyroscope features and the NB subset only one, which can be swapped out for an accelerometer feature with a minor drop in performance. Therefore, simple and orientation-independent activity recognition can be done without using a power-hungry gyroscope sensor.

During runtime, a classifier infers each new window of data. With a window size of 2 seconds and 50 % overlap there will be a new datapoint every second. For each datapoint, the CPU and memory are engaged in order

Table 8. Performance of both DT and NB classifiers for goat G5

	Classifier	Stationary	Walking	Eating	Running	Trotting	Other Activities	Average Performance
Accuracy	DT	90.77%	94.69%	89.08%	98.36%	98.11%	95.48%	94.41%
	NB	91.05%	94.52%	89.38%	98.63%	97.89%	95.07%	94.42%
Recall	DT	85.54%	96.04%	84.90%	72.35%	75.44%	0.36%	69.10%
	NB	87.52%	94.62%	82.14%	69.68%	76.15%	2.72%	68.81%
Precision	DT	93.82%	82.34%	72.12%	65.81%	61.66%	24.29%	66.67%
	NB	92.60%	82.57%	74.10%	74.29%	57.76%	17.67%	66.50%
F1	DT	89.49%	88.66%	77.99%	68.92%	67.85%	0.71%	65.60%
	NB	89.99%	88.19%	77.91%	71.91%	65.69%	4.72%	66.40%

to calculate the features and classify the datapoint; thus, minimizing the complexity of these steps will save large amounts of energy over longer runtime periods. The results in this section have shown that lightweight classifiers, such as Decision Tree and Naive Bayes [20], can obtain good performance when tested with data from diverse orientations. The monitoring system is more energy efficient with a lower quantity and complexity of features. Previous works [12, 32] have tried to obtain orientation independence through transformations of the input data. However, this requires a significant amount of additional computation for every window of data [42]. Therefore, we propose the use of only a few orientation-independent features that have low complexity.

There are two main types of features: time and frequency-domain features. Figo et al. [11] performed an extensive complexity analysis of features that are commonly used. Time domain features require less computational resources than frequency domain features. The results in Figure 14 show that most classifiers considered here have good performance with only a single feature. Figure 12 shows that using more than 3 features does generally not improve the performance of the classifiers. The optimal sets with 3 features do contain frequency domain features (frequency entropy and 6th magnitude of FFT). When energy efficiency is more important than accuracy, time domain features should be preferred, which results in a small drop in accuracy. Figure 14b shows that classifiers with a lightweight inference phase (NB, DT, and NN) obtain high accuracy (>92 %) using only the accelerometer 3D vector standard deviation, which is a time domain feature with low complexity [11].

Because we use 2-second observation windows with 50% overlap, the delay of an actual implementation would therefore be 1 second plus the time for classification. The classification time varies with classifiers and can, in theory, be done in milliseconds. Therefore, the total delay for the activity to be recognized in real time is more or less 1 second. However, transmitting an update to a central location would be really expensive in terms of energy. We propose to aggregate a log file with change points and transmit the information each time period T , or use conditional rules regarding the local context (e.g. transmit log file when $\tau > K$, where τ denotes the total duration of activity α within a time period U , and K denotes a maximum duration).

6 CONCLUSIONS

We have shown that it is possible to obtain optimal orientation-independent features by training a classifier with mixed data from various orientations. Accurate activity recognition is achieved using just one feature and can be slightly improved by using up to 3 features for the Decision Tree (DT) and Naive Bayes (NB) classifier; using a larger feature set does not improve the performance. The best scoring features are: *accel_3dvector_std*, *accel_3dvector_norm_mag_6*, *accel_3dvector_freqEntropy*, and *gyro_3dvector_twenty_fith_p*. This is a promising list because most of these features are accelerometer features. With a small drop in accuracy, the frequency-domain features can be swapped out for time-domain features. Our results have shown that the accelerometer is the

best sensor to use for activity recognition. A gyroscope will not significantly contribute to performance and consumes more energy than accelerometers. NB slightly outperforms DT and the NB feature set generalizes a little better over other classifiers. We prefer the NB classifier because of its superior accuracy, low complexity, and its ability to generalize over other goats. Our results indicate that the best location on the neck of a goat for activity recognition is at the top or bottom of the neck.

We have shown that the distribution of data over training, cross-validation, and test sets is more balanced when long segments of activities are recursively broken up into shorter segments. This method allows for improved, more balanced datasets without information leakage between training and test sets.

We argue that the activity of an animal is accurately recognized by placing only an accelerometer and a simple processing unit around its neck – in any orientation or position – with a NB classifier that only requires one feature, the standard deviation of $M(t)$ (1), over a 2 sec window of data. When using a classifier other than DT or NB, our approach can be repeated to select an optimal feature set for that classifier.

The optimal feature sets that have been revealed in this work are orientation-independent, consist of mostly time-domain features, and are used with simple classifiers; these properties support real-time, energy-efficient, and robust activity recognition for animals.

ACKNOWLEDGMENTS

The authors would like to thank the farms 'Kinderboerderij De Proosdij' in Ede and 'Kinderboerderij Enschede-Noord' for their vital support by allowing us to work with their animals. We would like to thank Vera Nauta and Romano Ferla for their help during the collection and labeling of the datasets that were used for the experiments.

REFERENCES

- [1] Louis Atallah, Benny Lo, Rachel King, and Guang Zhong Yang. 2011. Sensor positioning for activity recognition using wearable accelerometers. *IEEE Transactions on Biomedical Circuits and Systems* 5, 4 (2011), 320–329. DOI : <https://doi.org/10.1109/TBCAS.2011.2160540>
- [2] Jonghun Baek, Geehyuk Lee, Wonbae Park, and Bj Yun. 2004. Accelerometer signal processing for user activity detection. *Knowledge-Based Intelligent Information and Engineering Systems Lecture No* (2004), 610–617. DOI : https://doi.org/10.1007/978-3-540-30134-9_82
- [3] Jamali Firmat Banzi. 2014. A Sensor Based Anti-Poaching System in Tanzania. *International Journal of Scientific and Research Publications* 4, 4 (2014), 1–7.
- [4] Ling Bao and Stephen S. Intille. 2004. Activity Recognition from User-Annotated Acceleration Data. In *Pervasive Computing*. Springer Berlin Heidelberg, Berlin, Heidelberg, 1 – 17. DOI : <https://doi.org/10.1007/b96922>
- [5] Pablo Bermejo, José A. Gámez, and José M. Puerta. 2014. Speeding up incremental wrapper feature subset selection with Naive Bayes classifier. *Knowledge-Based Systems* 55 (2014), 140–147. DOI : <https://doi.org/10.1016/j.knsys.2013.10.016>
- [6] Owen R. Bidder, Hamish A. Campbell, Agustina Gómez-Laich, Patricia Urgé, James Walker, Yuzhi Cai, Lianli Gao, Flavio Quintana, and Rory P. Wilson. 2014. Love thy neighbour: Automatic animal behavioural classification of acceleration data using the k-nearest neighbour algorithm. *PLoS ONE* 9, 2 (2014), 1–7. DOI : <https://doi.org/10.1371/journal.pone.0088609>
- [7] Greg Bishop-Hurley, Dave Henry, Daniel Smith, Ritaban Dutta, James Hills, Richard Rawnsley, Andrew Hellicar, Greg Timms, Ahsan Morshed, Ashfaqur Rahman, Claire D'Este, and Yanfeng Shu. 2014. An investigation of cow feeding behavior using motion sensors. In *2014 IEEE International Instrumentation and Measurement Technology Conference (I2MTC) Proceedings*. IEEE, 1285–1290. DOI : <https://doi.org/10.1109/I2MTC.2014.6860952>
- [8] Steven J. Cooke, Scott G. Hinch, Martin Wikelski, Russel D. Andrews, Louise J. Kuchel, Thomas G. Wolcott, and Patrick J. Butler. 2004. Biotelemetry: A mechanistic approach to ecology. *Trends in Ecology and Evolution* 19, 6 (2004), 334–343. DOI : <https://doi.org/10.1016/j.tree.2004.04.003>
- [9] Ritaban Dutta, Daniel Smith, Richard Rawnsley, Greg Bishop-Hurley, James Hills, Greg Timms, and Dave Henry. 2015. Dynamic cattle behavioural classification using supervised ensemble classifiers. *Computers and Electronics in Agriculture* 111 (2015), 18–28. DOI : <https://doi.org/10.1016/j.compag.2014.12.002>
- [10] Jochen Fahrenberg, Friedrich Foerster, Manfred Smeja, and Wolfgang Müller. 1997. Assessment of posture and motion by multichannel piezoresistive accelerometer recordings. (1997). DOI : <https://doi.org/10.1111/j.1469-8986.1997.tb01747.x>
- [11] Davide Figo, Pedro C. Diniz, Diogo R. Ferreira, and João M P Cardoso. 2010. Preprocessing techniques for context recognition from accelerometer data. *Personal and Ubiquitous Computing* 14, 7 (2010), 645–662. DOI : <https://doi.org/10.1007/s00779-010-0293-9>

- [12] Blanca Florentino-Liaño, Niamh O'Mahony, and Antonio Artés-Rodríguez. 2012. Human Activity Recognition Using Inertial Sensors with Invariance to Sensor Orientation. In *2012 3rd International Workshop on Cognitive Information Processing (CIP)*. 1–6. DOI: <https://doi.org/10.1109/CIP.2012.6232914>
- [13] F Foerster and J Fahrenberg. 2000. Motion pattern and posture: correctly assessed by calibrated accelerometers. *Behavior research methods, instruments, & computers : a journal of the Psychonomic Society, Inc* 32, 3 (2000), 450–457. DOI: <https://doi.org/10.3758/BF03200815>
- [14] L A González, G J Bishop-hurley, R N Handcock, and C Crossman. 2015. Behavioral classification of data from collars containing motion sensors in grazing cattle. *Computers and Electronics in Agriculture* 110 (2015), 91–102. DOI: <https://doi.org/10.1016/j.compag.2014.10.018>
- [15] Isabelle Guyon and André Elisseeff. 2003. An Introduction to Variable and Feature Selection. *Journal of Machine Learning Research (JMLR)* 3, 3 (2003), 1157–1182. DOI: <https://doi.org/10.1016/j.aca.2011.07.027> arXiv:1111.6189v1
- [16] Tâm Huynh and Bernt Schiele. 2005. Analyzing features for activity recognition. In *Proceedings of the 2005 joint conference on Smart objects and ambient intelligence: innovative context-aware services: usages and technologies (sOc-EUSAI '05)*. 159–163. DOI: <https://doi.org/10.1145/1107548.1107591>
- [17] I Inza, P Larrañaga, R Etxebarria, and B Sierra. 1999. Feature subset selection by Bayesian networks based optimization. *Artificial Intelligence* 123 (1999), 157–184. DOI: [https://doi.org/10.1016/S0004-3702\(00\)00052-7](https://doi.org/10.1016/S0004-3702(00)00052-7)
- [18] Jacob Kamminga. 2017. Goat Orientation Data. online. (05 2017). <http://ps.ewi.utwente.nl/Datasets.php>
- [19] Dean M. Karantonis, Michael R. Narayanan, Merryn Mathie, Nigel H. Lovell, and Branko G. Celler. 2006. Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring. *IEEE Transactions on Information Technology in Biomedicine* 10, 1 (2006), 156–167. DOI: <https://doi.org/10.1109/TITB.2005.856864>
- [20] S B Kotsiantis. 2007. Supervised Machine Learning : A Review of Classification Techniques. *Informatica, An International Journal of Computing and Informatics* 3176, 31 (2007), 249–268.
- [21] Cassim Ladha, Nils Hammerla, Emma Hughes, Patrick Olivier, and Thomas Ploetz. 2013. Dog's life: Wearable Activity Recognition for Dogs. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing - UbiComp '13*. 415. DOI: <https://doi.org/10.1145/2493432.2493519>
- [22] Jonathan Lester, Tanzeem Choudhury, and Gaetano Borriello. 2006. A Practical Approach to Recognizing Physical Activities. *Pervasive Computing* 3968 (2006), 1–16. DOI: https://doi.org/10.1007/11748625_1
- [23] Yunji Liang, Kingshe Zhou, Zhiwen Yu, Bin Guo, and Yue Yang. 2012. Energy efficient activity recognition based on low resolution accelerometer in smart phones. *Advances in Grid and Pervasive Computing* (2012), 122–136.
- [24] Jacques Marais, Solomon Petrus, Le Roux, Riaan Wolhuter, and Thomas Niesler. 2014. Automatic classification of sheep behaviour using 3-axis accelerometer data. In *Proceedings of the twenty-fifth annual symposium of the Pattern Recognition Association of South Africa (PRASA)*. 1–6.
- [25] Paula Martiskainen, Mikko Järvinen, Jukka-Pekka Skön, Jarkko Tiirikainen, Mikko Kolehmainen, and Jaakko Mononen. 2009. Cow behaviour pattern recognition using a three-dimensional accelerometer and support vector machines. *Applied Animal Behaviour Science* 119, 1-2 (2009), 32–38. DOI: <https://doi.org/10.1016/j.applanim.2009.03.005>
- [26] B G Mathie, M. J., Coster, A. C., Lovell, N. H., & Celler. 2004. Accelerometry : providing an integrated , practical method for long-term , ambulatory monitoring of human movement. *Physiological measurement* 25, 2 (2004), R1. DOI: <https://doi.org/10.1088/0967-3334/25/2/R01>
- [27] MATLAB. 2015. *version 8.6.0 (R2015b)*. The MathWorks Inc., Natick, Massachusetts.
- [28] Uwe Maurer, Anthony Rowe, Asim Smailagic, and Daniel Siewiorek. 2006. Location and Activity Recognition Using eWatch: A Wearable Sensor Platform. *Ambient Intelligence in Everyday Life* 3864 (2006), 86–102. DOI: https://doi.org/10.1007/11825890_4
- [29] I Mierswa, M Wurst, R Klinkenberg, M Scholz, and T Euler. 2006. YALE: Rapid prototyping for complex data mining tasks. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 2006 (2006), 935–940. DOI: <https://doi.org/10.1145/1150402.1150531>
- [30] Ran Nathan. 2008. An emerging movement ecology paradigm. *Proceedings of the National Academy of Sciences of the United States of America* 105, 49 (2008), 19050–19051. DOI: <https://doi.org/10.1073/pnas.0808918105>
- [31] Ran Nathan, Orr Spiegel, Scott Fortmann-Roe, Roi Harel, Martin Wikelski, and Wayne M Getz. 2012. Using tri-axial acceleration data to identify behavioral modes of free-ranging animals: general concepts and tools illustrated for griffon vultures. *The Journal of experimental biology* 215, Pt 6 (2012), 986–996. DOI: <https://doi.org/10.1242/jeb.058602>
- [32] Trung Thanh Ngo, Yasushi Makihara, Hajime Nagahara, Yasuhiro Mukaigawa, and Yasushi Yagi. 2015. Similar gait action recognition using an inertial sensor. *Pattern Recognition* 48, 4 (2015), 1285–1297. DOI: <https://doi.org/10.1016/j.patcog.2014.10.012>
- [33] Juha Pärkkä, Miikka Ermes, Panu Korpipää, Jani Mäntyjärvi, Johannes Peltola, and Ilkka Korhonen. 2006. Activity classification using realistic data from wearable sensors. *IEEE transactions on information technology in biomedicine : a publication of the IEEE Engineering in Medicine and Biology Society* 10, 1 (2006), 119–128. DOI: <https://doi.org/10.1109/TITB.2005.856863>
- [34] Julie K Petersen. 2002. *Understanding Technologies Surveillance Spy Devices, Their Origins & Applications*. CRC Press, Boca Raton.
- [35] R. L. Plackett. 1983. Karl Pearson and the Chi-Squared Test. *International Statistical Review / Revue Internationale de Statistique* 51, 1 (1983), 59. DOI: <https://doi.org/10.2307/1402731>

- [36] S.J. Preece, J.Y. Goulermas, L.P.J. Kenney, and D. Howard. 2009. A Comparison of Feature Extraction Methods for the Classification of Dynamic Activities From Accelerometer Data. *IEEE Transactions on Biomedical Engineering* 56, 3 (2009), 871–879. DOI: <https://doi.org/10.1109/TBME.2008.2006190>
- [37] Nishkam Ravi, Nikhil Dandekar, Preetham Mysore, and Ml Michael L Littman. 2005. Activity Recognition from Accelerometer Data. In *Aaai*. 1541–1546. DOI: https://doi.org/10.1007/978-3-642-02481-8_120
- [38] Sasank Reddy, Min Mun, Jeff Burke, Deborah Estrin, Mark Hansen, and Mani Srivastava. 2010. Using mobile phones to determine transportation modes. *ACM Transactions on Sensor Networks* 6, 2 (2010), 1–27. DOI: <https://doi.org/10.1145/1689239.1689243>
- [39] Yasar Guneri Sahin. 2007. Animals as Mobile Biological Sensors for Forest Fire Detection. *Sensors* 7 (2007), 3084–3099. DOI: <https://doi.org/10.3390/s7123084>
- [40] Jeff Schneider. 1997. Cross Validation. online. (1997). <https://www.cs.cmu.edu/~schneide/tut5/node42.html>
- [41] Emily L C Shepard, Rory P. Wilson, Flavio Quintana, Agustina Gomez Laich, Nikolai Liebsch, Diego A. Albareda, Lewis G. Halsey, Adrian Gleiss, David T. Morgan, Andrew E. Myers, Chris Newman, and David W. Macdonald. 2010. Identification of animal movement patterns using tri-axial accelerometry. *Endangered Species Research* 10, 1 (2010), 47–60. DOI: <https://doi.org/10.3354/esr00084>
- [42] Muhammad Shoab, Stephan Bosch, Ozlem Incel, Hans Scholten, and Paul Havinga. 2015. A Survey of Online Activity Recognition Using Mobile Phones. *Sensors* 15, 1 (2015), 2059–2085. DOI: <https://doi.org/10.3390/s150102059>
- [43] Pekka Siirtola and Juha Rönning. 2012. Recognizing Human Activities User-independently on Smartphones Based on Accelerometer Data. *International Journal of Interactive Multimedia and Artificial Intelligence* 1, 5 (2012), 38. DOI: <https://doi.org/10.9781/ijimai.2012.155>
- [44] Daniel Smith, Bryce Little, Paul I. Greenwood, Philip Valencia, Ashfaqur Rahman, Aaron Ingham, Greg Bishop-Hurley, Md Sumon Shahriar, and Andrew Hellicar. 2015. A study of sensor derived features in cattle behaviour classification models. In *2015 IEEE SENSORS - Proceedings*. IEEE, Busan, South Korea, 1–4. DOI: <https://doi.org/10.1109/ICSENS.2015.7370529>
- [45] Daniel Smith, Ashfaqur Rahman, Greg J. Bishop-Hurley, James Hills, Sumon Shahriar, David Henry, and Richard Rawnsley. 2016. Behavior classification of cows fitted with motion collars: Decomposing multi-class classification into a set of binary problems. *Computers and Electronics in Agriculture* 131 (2016), 40–50. DOI: <https://doi.org/10.1016/j.compag.2016.10.006>
- [46] J Sneddon and A Mason. 2014. Automated Monitoring of Foraging Behaviour in Free Ranging Sheep Grazing a Bio-diverse Pasture using Audio and Video Information. In *8th International Conference on Sensing Technology*. 2–4. DOI: <https://doi.org/10.1109/ICSensT.2013.6727614>
- [47] Inertia Technology. 2017. ProMove mini. online. (2017). <http://inertia-technology.com/>
- [48] Jorge A. Vázquez Diosdado, Zoe E. Barker, Holly R. Hodges, Jonathan R. Amory, Darren P. Croft, Nick J. Bell, and Edward A. Codling. 2015. Classification of behaviour in housed dairy cows using an accelerometer-based activity monitoring system. *Animal Biotelemetry* 3, 1 (2015), 15. DOI: <https://doi.org/10.1186/s40317-015-0045-8>
- [49] W3C. 2017. Motion Sensors Explainer. online. (08 2017). <https://www.w3.org/TR/motion-sensors/>
- [50] Shinichi Watanabe, Masako Izawa, Akiko Kato, Yan Ropert-Coudert, and Yasuhiko Naito. 2005. A new technique for monitoring the detailed behaviour of terrestrial animals: A case study with the domestic cat. *Applied Animal Behaviour Science* 94, 1-2 (2005), 117–131. DOI: <https://doi.org/10.1016/j.applanim.2005.01.010>
- [51] Rory P. Wilson, E. L C Shepard, and N. Liebsch. 2008. Prying into the intimate details of animal lives: Use of a daily diary on animals. *Endangered Species Research* 4, 1-2 (2008), 123–137. DOI: <https://doi.org/10.3354/esr00064>
- [52] Xiuxin Yang, A Dinh, and Li Chen. 2010. Implementation of a wearable real-time system for physical activity recognition based on naive Bayes classifier. *Bioinformatics and Biomedical Technology (ICBBT), 2010 International Conference on* (2010), 101–105. DOI: <https://doi.org/10.1109/ICBBT.2010.5479000>
- [53] K. Yoda, Katsufumi Sato, Y. Niizuma, M. Kurita, Charles-André Bost, Yvon Le Maho, and Y. Naito. 1999. Precise monitoring of porpoising behaviour of Adélie penguins determined using acceleration data loggers. *The Journal of experimental biology* 202, Pt 22 (1999), 3121–3126.
- [54] Mi Zhang and Alexander a Sawchuk. 2011. A feature selection-based framework for human activity recognition using wearable multimodal sensors. In *Proceedings of the 6th International Conference on Body Area Networks*. 92–98. DOI: <https://doi.org/10.4108/icst.bodynets.2011.247018>

Received May 2017; revised November 2017; accepted January 2018