Semi-wildlife gait patterns classification using Statistical Methods and Artificial Neural Networks

D. Gutierrez-Galan, J.P. Dominguez-Morales, L. Miro-Amarante, F. Gomez-Rodriguez, M. Dominguez-Morales, M. Rivas-Perez, A. Jimenez-Fernandez, A. Linares-Barranco
Robotics and Technology of Computers Lab.
Computer Engineering School - Polytechnic Engineering School - University of Seville (Spain)
Email: dgutierrez@atc.us.es

***Abstract*—Several studies have focused on classifying behavioral patterns in wildlife and captive species to monitor their activities and so to understanding the interactions of animals and control their welfare, for biological research or commercial purposes. The use of pattern recognition techniques, statistical methods and Overall Dynamic Body Acceleration (ODBA) are well known for animal behavior recognition tasks. The reconfigurability and scalability of these methods are not trivial, since a new study has to be done when changing any of the configuration parameters. In recent years, the use of Artificial Neural Networks (ANN) has increased for this purpose due to the fact that they can be easily adapted when new animals or patterns are required. In this context, a comparative study between a theoretical research is presented, where statistical and spectral analyses were performed and an embedded implementation of an ANN on a smart collar device was placed on semi-wild animals. This system is part of a project whose main aim is to monitor wildlife in real time using a wireless sensor network infrastructure. Different classifiers were tested and compared for three different horse gaits. Experimental results in a real time scenario achieved an accuracy of up to 90.7%, proving the efficiency of the embedded ANN implementation.**

I. Introduction

Monitoring the behavior of wildlife is an important task when it comes to understand their reproduction, survival, welfare, interaction with other animals and other interesting parameters. For this reason, the study and monitoring of wildlife has always been a subject of great interest. Therefore, the design and deployment of a monitoring system capable of obtaining behavioral information from animals has been the focus of several studies [1]–[6].

Collecting and processing relevant information from wildlife is a hard technological task [7]–[9] due to several factors: (1) the development of a device that needs to be attached to the animal, (2) the deployment of a wireless network to collect and transmit the information, (3) the need to implement an accurate behavior classifier and, finally, (4) the process of storing and manipulating the data.

Animals' behavioral parameters can be measured using different types of sensors. With this data, different communication strategies can be used to send the collected information. Wireless sensor network solutions capable of measuring specific behavioral parameters and transmit them to a central base station have been designed in recent studies [3], [5], [6], [10].

The use of inertial sensors, like the accelerometer, gyroscope and magnetometer, overcomes the disadvantages of GPS (long-life batteries are needed due to the power consumption and high capacity memory to store the data) to obtain information about the acceleration and orientation of the animal [4]–[6], [11], [12]. Methods based on supervised learning algorithms in [11], [12] provide very good classification predictions with more than 90% accuracy, distinguishing between two categories: active and inactive.

Regardless of the type of sensors used to monitor animals, large amounts of data are needed when studying their behavioral patterns, implying important analysis and interpretation steps of the information. Currently, many studies apply Wilson's metric Overall Dynamic Body Acceleration (ODBA) [13] to classify animal behavior. Wilson checked the relationship between body acceleration and the oxygen consumed (energy expenditure). He proved that body acceleration can quantify the amount of mechanical work performed by the body. Hence, these metrics are able to distinguish between active and inactive behaviors. On the other hand, the most common classifier systems are Neural Networks (NN), Support Vector Machines (SVM) [8] or even complex statistical methods [14], which can detect specific behaviors such as sleeping, foraging, rumiating, etc. The use of NNs has increased due to the fact that their architecture can be easily adapted when new animals or behaviors need to be introduced in the system, which is not supported by other classifiers (a new statistical test needs to be done in order to set the new configuration parameters). This scalability issue can be seen in works like [14], in which the authors present a collar device provided with inertial sensors to obtain information from the animal that is later classified using a supervised threshold-based technique and a finite-state machine, achieving a good accuracy result, but lacking on the reconfigurability that NNs support.

In [15] a decision-tree algorithm that used tri-axial accelerometer data from a neck-mounted sensor is studied to classify three types of biologically relevant behaviors in dairy cows. Although the system achieves a good accuracy, it has the same disadvantage as [14] or as any classification method based on thresholds or statistical methods: the system cannot be adapted to new behaviors in an online way (i.e. without collecting the sensors from the animal and then placing them back on the neck of the animal).

In [5], a premilinar study on the application of NNs in animal behavior classification was performed. Several tests were carried out offline using different ANN architectures and input datasets. The results concluded the viability of using ANN for this purpose and the need of processing the collected data before being used as input for the NN.

MINERVA is a research project whose main aim is to study and classify wildlife behavior inside Doñana National Park [16]. For this purpose, a collar device that contains a set of sensors has been developed in order to be placed on animals and collect behavioral information from them. This information is sent using a wireless modem, through a ZigBee Personal-Area-Network (PAN) distributed along the study area, to a remote database server for further research by Doñana's biologist staff.

Currently, the tracking and classification systems for wildlife that are deployed and being used in Doñana National Park obtain their position using a GPS and transmit it via GSM (by SMS). To reduce power consumption, the position is obtained between two and five times a day. These solutions are not enough for biologists' interests: they need more information to monitor animals and recognize behaviors. To solve this lack of information, two solutions can be implemented: the system could be adapted to transmit information more regularly (since communications consume in average more than 80% of battery life, this option is inefficient); or, on the other hand, this information could be processed in a local way in order to classify the animal behavior and transmit only the result of the classification. This fact increases battery life: instead of sending the information after every sensor reading, the communication to the network only occurs after several sensor readings. Viability and power consumption tests for these two approaches have been carried out by the authors in [17]. This project has the additional aim of developing a communication infrastructure for collecting this information and make it accessible on the Internet.

In the network topology implemented for the MINERVA project there are three different device types: **base station** (coordinator), **motes** (routers) and **collars** (end devices). Routers must join the base station before transmitting and receiving data. This type of devices allow others to join the network. Finally, the collars must also join the PAN, but they do not allow other devices to join the network. However, in this work we will obtain the dataset using a point-to-point connection between the collar and the computer and, then, we will focus on the collar device alone.

To recognize the animal's behavior, several classifiers could be used. In this project, there are two essential requirements for the classifier system: good accuracy and low computational load. Therefore, two parallel tests have been conducted on the dataset: a full spectral and statistical analysis (in a desktop computer) and an ANN implementation (inside the collar embedded system). The aim of this work is to obtain an efficient and accurate classification mechanism with a high hit rate that can be embedded into the collar.

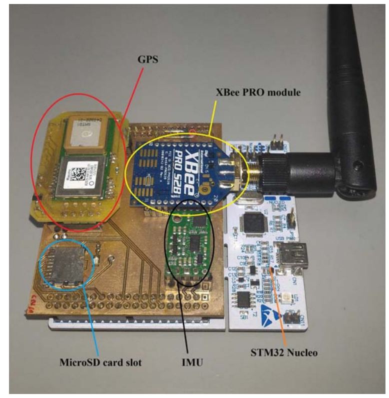

Fig. 1. Collar Prototype.

The paper is structured as follows: Section 2 describes the collar device that is placed on the animals to monitor them. In section 3, spectral and statistical analyses are performed with horse gait information obtained from inertial sensors. Then, section 4 presents the ANN implementation on the collar and the experiments that have been performed in order to test its performance. In Section 5, the authors discuss the results obtained and compare the statistical and spectral analyses with the horse gait accuracy achieved in the real testing scenario. Finally, Section 6 presents the conclusions of this work.

II. Collar

The collar collects information from the animal on which it is placed by using different sensors. It has a MinIMU-9V2 inertial measurement unit (IMU), which consists of a LSM303DLHC 3-axis accelerometer, a L3GD20 3-axis gyroscope and a 3-axis magnetometer. An I2C interface accesses nine independent rotation, acceleration, and magnetic measurements that can be used to calculate the sensor's absolute orientation. Each of these sensors have 12-bit resolution for a more precise data acquisition. The IMU is used in addition to a GPS, which provides location and time information in all weather conditions.

The main aim of the collar is to classify the animal behavior (between three different gait patterns) using the information obtained from the IMU as an input for a feed-forward neural network implemented on the collar's microcontroller unit (MCU). The periodic measurements of each sensor are carried out using a low power microcontroller (STM32L152 [18]) with a real-time operating system (RTOS).

The collar prototype (see Fig. 1) has an XBee module (XBee PRO S2B [19]) that can transmit data through a wireless network (standard ZigBee 2.4GHz at 250kbps). XBee modules are integrated solutions based on ZigBee, which is an open global standard of the IEEE 802.15.4 MAC/PHY [20], [21]. This device family allows to implement a mesh network of motes (or routers) where collars (or devices) send information, and these elements redirect the packets to a base station (or coordinator) where the information is uploaded to a database server. In the case that the signal cannot reach a valid point to transmit, i.e. the animal is out of the network coverage, the collar carries a microSD card where the information is stored; thus, the animal behavioral information can be accessed later or offline, avoiding data loss. Having this network architecture, the system allows to obtain the information in real time without the need of capturing the animal. Due to the fact that capturing a semi-wild animal is very expensive and complex, the microcontroller is able to switch to sleep mode if there are no routers in the network coverage capable of receiving this collar's information, which increases battery life. Moreover, the measurements are transmitted periodically according to a frequency value that is established and that can be modified, reducing radio transmissions and, thus, reducing power consumption [17].

To obtain the dataset for the training process, the collar was placed on six semi-wild horses. For each horse, the collar was configured to transmit the accelerometer data in a continuous way every 33 milliseconds (sample rate of 30.3 Hz), while a computer stored the information received. During these sessions, the horses performed various tests in order to obtain data for the different study behaviors: standing, walking and trotting; and these gaits were then labeled manually.. A dataset consisting of 180,000 samples was obtained.

Using the data stored in those tests, two different tests were conducted: a spectral and statistical analysis, and an ANN implementation. Both tests are described in sections 3 and 4, respectively.

III. Data Analysis

Acceleration data represent a useful and reliable measure for accurately recording the activities and detailed behaviors of horses. As pointed out in [22], accelerometers are generally useful for determining different categories of animal behavior. Fig. 2 shows the acceleration data obtained from each axis of the accelerometer (X, Y, Z) for three horse gaits (Standing, Walking and Trotting). These plots show a subset of 2,200 samples. As it is shown, acceleration values depend on the activity level of the horse.

In this study, the following parameters were used:

- Accelerometer axes (X, Y, Z): X, Y, Z axes.
- Magnitude (M): Square root of the sums of squares of the acceleration in the X, Y, Z axes.
- Magnitude Fast Fourier Transform (MFFT): Fast Fourier Transform of the magnitude of the acceleration power spectrum.
- Dominant frequency of Magnitude, in Hz (DFM): Frequency at dominant power spectrum in 330 sample windows.
- Standard Deviation of dynamic acceleration (SDX, SDY, SDZ): Standard deviation of dynamic acceleration.
- Overall Dynamic Body Acceleration (ODBA): Sum of the mean of dynamic acceleration values along X, Y, and Z.
- Overall Dynamic Body Acceleration Mean (ODBA Mean): Mean of the ODBA.

A. Spectral analysis

The Fast Fourier Transform (FFT) is one of the most popular methods in accelerometer data analysis. Fourier Transformations identify the individual frequencies that are present in the raw acceleration waveform and determine the power spectral densities of those frequencies. The periodic properties of the acceleration signals recordings during dynamic gaits of the horse (standing, walking and trotting) allowed us to apply a FFT in order to determine the frequency of a particular movement.

Fig. 3 summarizes the spectral analysis. From the data shown in the top row, the FFT was computed. The duration of each acceleration window was fixed to 10s (303 samples). The central row of Fig. 3 shows the power spectrum ranged from 0 to 16.5 H, where some peaks can be observed; these peaks correspond to the dominant frequencies of the horse's movements.

The dominant frequency spectra show one to three marked peaks occurring, on average, at 1.097 Hz, 1.541 Hz and 3.064 Hz, during standing, walking and trotting, respectively. In the case of trotting, more peaks can be observed, which correspond to secondary movements of the horse.

This spectral analysis proves the dominant frequency can be used to distinguish the gait (standing, walking or trotting). Two thresholds are defined: the UpperThreshold is set to 2.7 Hz and the LowerThreshold to 1.3 Hz (horizontal black lines in the bottom row of Fig. 3). The spectral classifier uses these thresholds. When the dominant frequency is higher than the UpperThreshold, it is considered that the horse is trotting; when dominant frequency is lower than the LowerThreshold, it indicates that the horse is standing. If the dominant frequency is between the two thresholds, it means that the horse is walking.

- *Results of the Spectral Analysis*

Table I shows the horse gait pattern classification results, where each row shows the results of the spectral classifier when the horse is standing (top row), walking (central row) and trotting (bottom row), and the columns show the behavioral prediction. The spectral classifier obtains an average hit rate of 84.63%. The results indicate that the DFM metric is a useful method for horse gait classification.

TABLE I
Hit and Miss rate of the horse gait prediction using the Spectral Classifier (%).

Gait	Prediction: Standing	Prediction: Walking	Prediction: Trotting	Average hit rate
Standing	***70.98%***	28.92%	0.10%	
Walking	5.41%	***82.23%***	11.36%	***84.63%***
Trotting	0.32%	0.00%	***99.68%***	

B. Statistical analysis

In addition to the spectral analysis, the authors performed a simpler statistical analysis based on the standard deviations

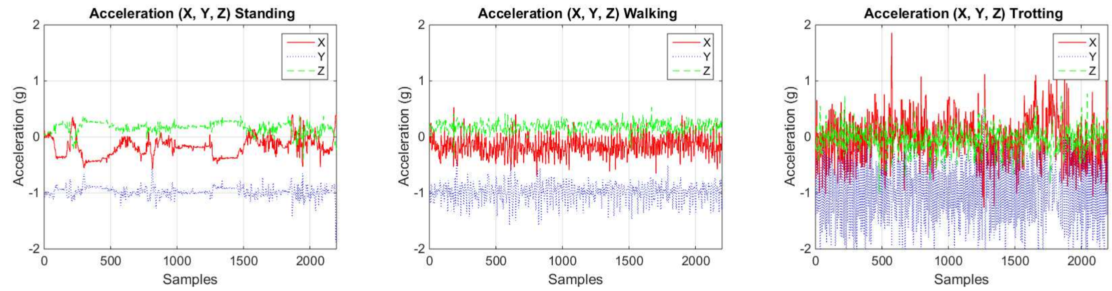

Fig. 2. Accelerometer Data.

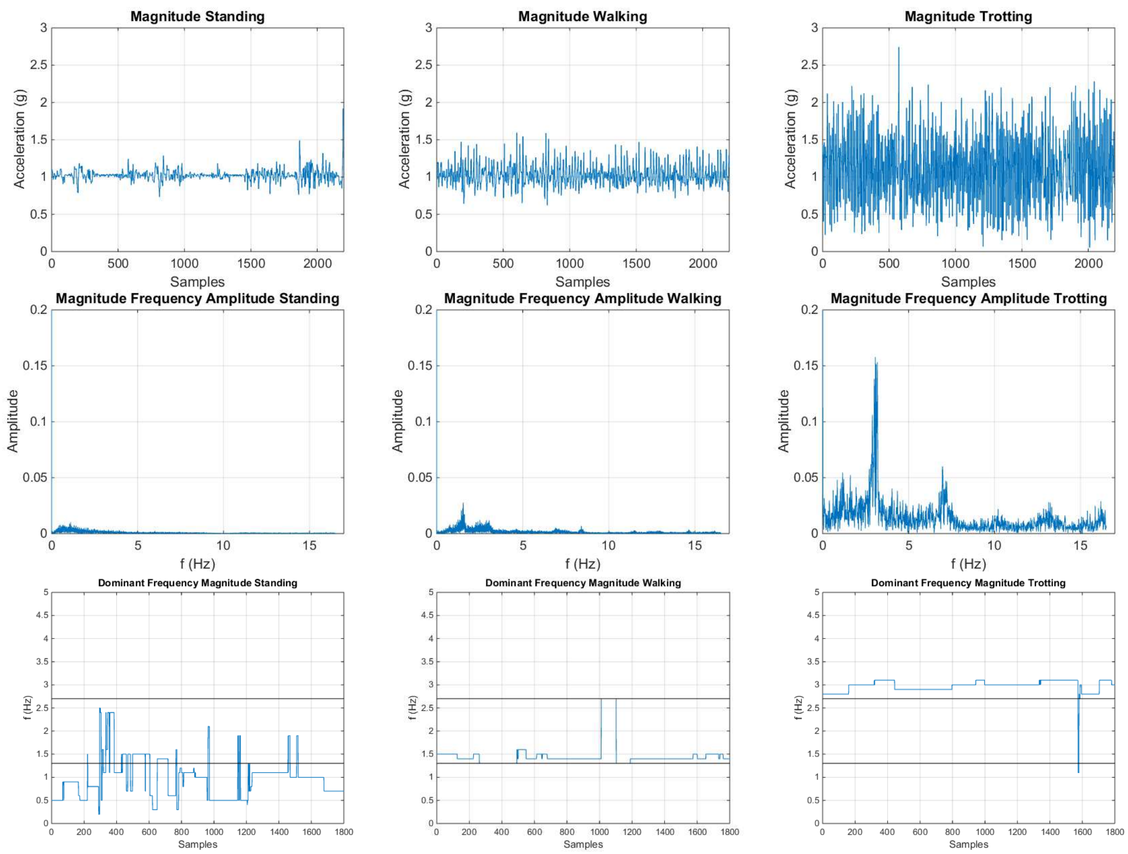

Fig. 3. Spectral Analysis. Top row (M): magnitude of the acceleration; central row (MFFT): fast Fourier Transform of magnitude power spectrum; bottom row (DFM): dominant frequency of Magnitude. Each column corresponds to one gait.

of the accelerometer axes (SDX, SDY and SDZ) and the Overall Dynamic Body Acceleration Mean (ODBA Mean). In this case, the accelerometer data is processed in slices of 1.21 seconds (40 samples). Fig. 4 shows the results of these four tests. Each plot presents one of these analyses applied on the three gaits (standing, walking and trotting). These plots show a subset of 2,200 samples. Depending on the gait of the horse, the values of standard deviations and ODBA Mean are slightly different. The results obtained with the SDY and ODBA Mean metrics present a greater difference and the values do not overlap.

In this work, four statistical classifiers based on the four metrics (SDX, SDY, SDZ and ODBA Mean) were used. For each of them, the upper and lower thresholds in which there is a minor overlap of the different gaits (horizontal black lines) were set. The thresholds are obtained manually taking into account the minimum, maximum and mean values of the standard deviation of the acceleration axes, and the ODBA mean. For SDX, the lower threshold is set to 0.1 and the upper threshold to 0.35; for SDY, 0.1 and 0.35; for SDZ, 0.06 and 0.2; and, for ODBA mean, 0.25 and 0.6 (horizontal black lines in Fig. 4). When a standard deviation value or ODBA Mean value is higher than its upper threshold, it is considered that the horse is trotting; when a value is lower than its lower

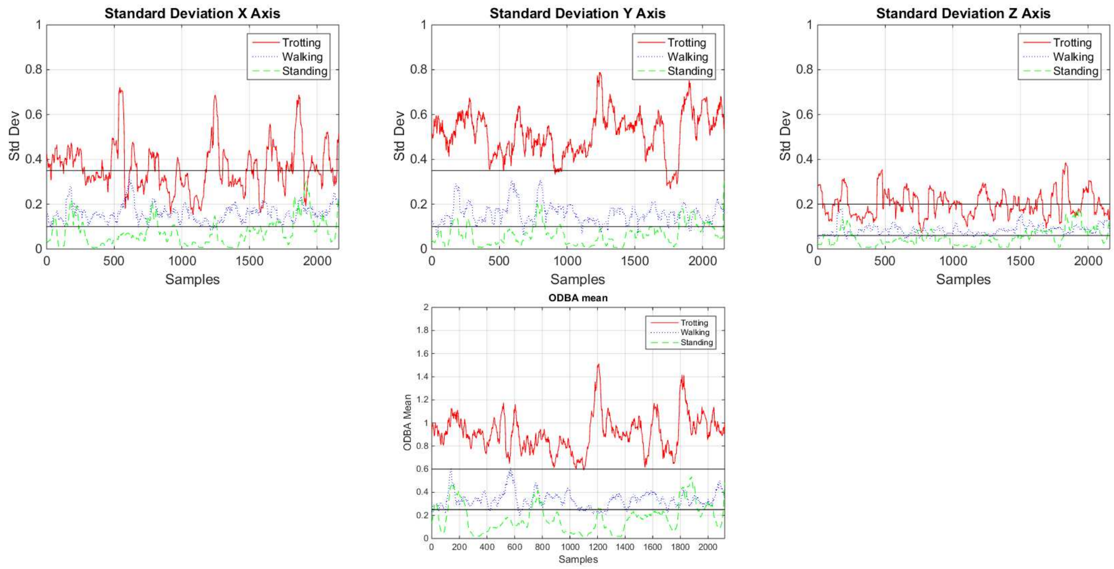

Fig. 4. Statistical Data. Three first rows (SDX, SDY and SDZ): standard deviation of the acceleration axes; and bottom row (ODBA Mean): Overall Dynamic Body Acceleration Mean.

threshold, it indicates that the horse is standing. If the value is between the two thresholds, it means that the horse is walking.

- *Results of the Statistical Analysis*

Table II summarizes the hit and miss rates of the horse gait classification from the statistical analysis. The standing and walking gaits are correctly identified using the standard deviations of the accelerometer axes (over 77%) and the ODBA mean (over 64%). However, the trotting gait can be only identified with the standard deviation of the Y accelerometer axis (over 94%) and the ODBA mean (over 99%); the other metrics achieve a success below 52%.

TABLE II
HIT AND MISS RATE OF THE HORSE GAIT PREDICTION USING THE STATISTICAL CLASSIFIER (%).

	Gait	Prediction			Average hit rate
		Standing	Walking	Trotting	
SDX	Standing	***77.65%***	21.85%	0.50%	***72.89%***
	Walking	10.55%	***89.45%***	0.00%	
	Trotting	0.00%	48.43%	***51.57%***	
SDY	Standing	***85.78%***	14.08%	0.14%	***90.39%***
	Walking	8.49%	***90.74%***	0.77%	
	Trotting	0.00%	5.35%	***94.65%***	
SDZ	Standing	***90.04%***	9.67%	0.29%	***73.92%***
	Walking	15.64%	***84.36%***	0.00%	
	Trotting	0.00%	52.63%	***47.37%***	
ODBA Mean	Standing	***64.02%***	35.68%	0.29%	***81.36%***
	Walking	18.98%	***80.45%***	0.57%	
	Trotting	0.00%	0.37%	***99.63%***	

The SDY statistical classifier gets an average hit rate of 90.39% and the ODBA Mean statistical classifier gets a 81.36%. From the results, we conclude that SDY and ODBA Mean are the most reliable parameters for this classification.

IV. EMBEDDED ARTIFICIAL NEURAL NETWORK

We have explored the use of ANNs in this work by embedding and deploying them in an embedded system (collar's microcontroller) using a free open source neural network library called Fast Artificial Neural Network (FANN*) [23], which implements multilayer ANNs in C programming language. The FANN library has bindings with more than 20 programming languages and several graphical user interfaces (GUIs), although in this paper the standard C library with no wrappers or GUIs is used due to the fact that the ANN is being implemented on a microcontroller. It is easy to use, well documented, versatile and allows to use both floating point and fixed point numbers.

Since the MINERVA project the application needs to be focused on low-power consumption devices (capturing an animal to replace their collar is very expensive and difficult), a microcontroller with no Floating Point Unit (FPU) is used. Therefore, the fixed point number classes of the FANN library are used. By having the ANN implemented on the collar, only the classification results are transmitted to the PAN, which is much less information than the data that should be transmitted if the classification step was done in the server instead of being done in the collar (a lot of information is needed to predict a behavior from the accelerometer data). As it has been proved in [17], this approach highly reduces power consumption, making it the best solution for this project.

In this work, two different versions of the library are used. The first one is the full fixed point FANN library that is available in GitHub. This version is used to train the neural network and, after this step, to test the dataset with the configuration that has been obtained in the training phase.

*Link GitHub FANN: https://github.com/libfann/fann

This whole simulation process is done in a PC to test how good the classication results would be before deploying the ANN configuration on the microcontroller. The second version of the FANN library used corresponds to a modified version that we have developed to be implemented on the collar. In this case, the performance has been improved and some parts of the code have been removed for power saving, such as: the training phase, floating point operations and other classes and functionalities that are not necessary when it comes to obtaining the classication results. This second version, embedded in the collar, does not need a large amount of resources. The reasons why the training functionality has been removed from the release version of the FANN library are: (1) the microcontroller has limited processing capabilities and memory; (2) the computational cost of the training step (which significantly reduces battery life); and (3) the training process could be done in the PC instead of in the collar, due to the fact that, for this purpose and work, there is no need to train the network in real time on the embedded system.

The FANN library implements several activation functions that the users may choose to be used in their model: linear, threshold, threshold symmetric, sigmoid stepwise, sigmoid symmetric stepwise, linear piece and linear symmetric piece. The sigmoid symmetric stepwise was used because the authors used it in the experiments that were carried out in [6], so that the results can be compared using the same network architecture. FANN supports a set of training algorithms, but the default and most used one is the backpropagation algorithm. Some parameters like the mean squared error (MSE), the number of epochs and the activation function of neurons in both the hidden and the output layer can be configured to improve the resulting accuracy.

The training process consists of two different phases. First, we need to generate the input data files (training and testing files) to feed the ANN from the obtained sensors information, using both raw or processed data, with the correct format that the library supports. After this, we need to configure the parameters that were previously mentioned, i.e. MSE, the number of neurons, activation functions, etc., and generate a file that contains the full architecture of the ANN.

For the first part, the IMU data need to be extracted from the sensors data packet. After this, two approaches could be considered: to feed the ANN using raw data, or to process the IMU data, obtaining a new input dataset. In [24], the authors calculated variance values from raw data obtained from the axes of the accelerometer (x, y and z axes) taking sets of 30 samples and using them as input to a Spiking Neural Network (SNN), obtaining an average accuracy of 83.33%. For the present work, the variance is calculated using slices with a width of 40 samples of accelerometer data. A 40-sample frame corresponds to 1.2 seconds of the movement performed by the horse (according to the sample rate), which is enough for monitoring a complete single period of any of its gaits. The reason why the accelerometer data were used is because it is the sensor that provides the most important information about the animal's movement, according to [5]. In that work, a premilinary study of the application of ANNs in animal monitoring was performed by the authors. Several tests were carried out using different ANN architectures and input datasets. The results concluded the viability of using ANNs for this purpose and the need of processing the collected data being used as input to the ANN.

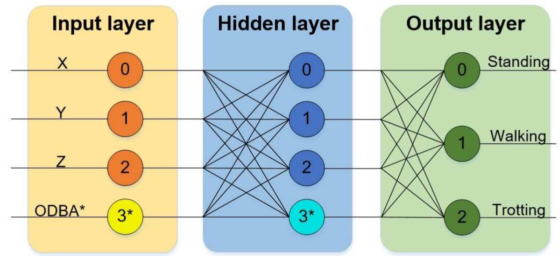

Fig. 5. ANN architecture implemented. Neurons with (*) only works when ODBA is used as input.

For the second part, several experimental tests were performed in order to obtain the best parameter configuration, which led to achieve better results. The number of input and output neurons were fixed in the first part, so the value of parameters like the number of hidden layers or neurons per layer, among others, were changed and tested to determine which configuration obtains the best classification accuracy. Depending on the study case (horse breed, dataset, etc) the MSE value and the number of epochs could be changed to match the system requirements. All of these configuration parameters are tested using the release version of the FANN library, which allows to simulate the ANN and obtain the results in terms of accuracy ratio. When good results are obtained using a specific ANN architecture, a formatted text file is generated containing the necessary information about this architecture, which can be loaded directly into the embedded version of the FANN library.

As our study is focused on horse gaits (standing, walking and trotting), and only accelerometer data are used (3 axes: x, y and z), the ANN has four neurons in the input layer, one hidden layer with three neurons, and three neurons in the output layer, as can be seen in Fig. 5. As it was described before, the input data of the ANN can be both raw or processed data so, in any case, the number of inputs of the ANN are three (one per axis). However, in section 3, the results show that the classifier system using ODBA obtains a good hit average; therefore, one input neuron (and another one in the hidden layer) could be added to the ANN architecture in order to provide it with more information about horse gaits and try to improve the hit rate.

Several tests were performed with the aim of obtaining the average accuracy of each ANN architecture by using different input datasets in the PC running the full fixed point FANN library. For this purpose, ten simulations per architecture and the mean hit rate per gait were calculated. In the training process, a MSE of 0.02 and 1500 epochs was established and the backpropagation algorithm was used. The results of these tests are shown in Table III.

TABLE III
ACCURACY OF THE SIMULATION OF HORSE GAIT CLASSIFICATION USING ANN.

Using ODBA	DATA	Standing	Walking	Trotting	Average hit rate
NO	RAW	88.1%	75.9%	75.4%	***79.8%***
	Processed	85.4%	87.9%	97.5%	***90.3%***
YES	RAW	89.6%	78.1%	80.3%	***82.6%***
	Processed	84.7%	88.8%	98.6%	***90.7%***

As can be seen, while 90% of hit average was achieved when the processed dataset was used, 82% was achieved using RAW dataset. On the one hand, trotting was the best classified gait when processed data was used. However, on the other hand, standing was better when raw data was used. These results also show that the accuracy is not improved when the ODBA value is used as an ANN input.

After these tests were performed, a test in a real scenario was carried out in order to validate the results obtained in the simulations. In this test, the collar implementing the embedded version of FANN was placed on six different horses: three using processed data without ODBA and three using processed data and ODBA. 5 minutes per gait were considered, calculating the average accuracy obtained for each one at the end of it. Tests using raw data were not performed in this case because in simulations their results were not as good as those of the processed data.

TABLE IV
ACCURACY OF THE HORSE GAIT CLASSIFICATION TEST USING ANN IN A REAL SCENARIO.

	Using ODBA	Standing	Walking	Trotting	Hit average
Processed data	NO	85.1%	88.1%	97.3%	***90.2%***
	YES	86.1%	89.2%	95.3%	***90.3%***

Table IV show the results obtained in this test. Like in the previous test, the hit rate percentage obtained in a real scenario was also around 90% in both cases, with or without ODBA. Therefore, the authors concluded that to use the ODBA value as an ANN input is not relevant in this case of study.

V. COMPARATIVE STUDY

After the experiments were carried out using different methods for horse gait classification, as described in previous sections, a comparative study between them is presented in order to determine which one is the best in terms of average accuracy (Fig. 6) and which one meets the project requirements more adequately.

As can be seen from the sensor data analysis approach, classifier systems using DFM and SDY mechanisms obtained the best accuracy results. DFM achieved a mean of 84.63% hit rate, with a standard deviation of 14.40. Although the obtained results are better than the ones that are implemented at the moment in Doñana National Park, this solution cannot be deployed on the collar device, in order to calculate DFM, complex operations need to be computed and the MCU does not

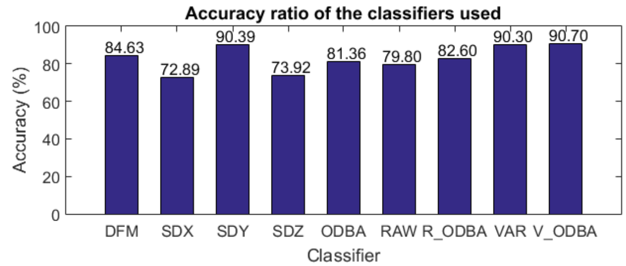

Fig. 6. Average accuracy of each of the classification systems studied.

have FPU to perform them. On the other hand, SDY achieves a 90.39% mean accuracy ratio (4.45 standard deviation), which is even better than of DFM. In addition to the need of FPU, the disadvantage is that, to obtain these classification results, two thresholds need to be established manually, which could differ depending on the dataset and the species. Furthermore, these two classifiers are usually used to study the animal's behavior after the whole dataset is gathered, thus, they cannot be supported in a real-time embedded device.

Using the ANN implemented on the collar, the system achieves a classification result around 90% which is similar to the ones obtained from DFM and SDY. In this case, the system runs on real time instead of being studied offline with no real time classification, as in [5], where the samples were processed by complex operations (like FFT and Kalman filter, among others), using them as inputs in the MATLAB Neural Network Toolbox. This is an advantage when it comes to monitoring wildlife, because the Doñana National Park's research staff are able to get information about the animals' activity at any time without the need of performing a classification analysis of the data on their own. However, it is important to note that, in order to train the ANN a huge amount of data is needed. On the other hand, the power consumption of this system is definitely lower than that of those that were tested in the data analysis section, because those tests involve the use of a computer, while this system runs on a low-power MCU. The average accuracy achieved in the best case in this work (90.3%) is slightly lower than the best case in [5] (95.1%). However, the achieved improvements are very important due to the fact that, by applying only simple operations we obtained 5% less than when complex mathematical algorithms are applied, and less types of datainputs were needed (since in order to apply Kalman filter, all IMU sensor samples are needed, while the variance is calculated using accelerometer values only).

There are several requirements established by the MINERVA project, which must be satisfied. Among them, real time classication is the most important. Therefore, an exhaustive study would be performed by researchers. Proposed ANN provides results without any further pre-processing, what reduces computational resources. However, some important things in embedded devices, like memory, have also been taken into account. A qualitative comparison between methods used in this work is shown in Table V:

The ANN is, in this case, the solution that best satisfies the

TABLE V
QUALITATIVE COMPARISON.

Method	Real time	Simple operations	Memory optimization	Adaptable
ANN	Yes	Yes	No	Yes
Spectral analysis	No	No	No	No
Statistical analysis	No	Yes	Yes	No

requirements. The main advantage of this method is that it is able to recognize features and learn them, being automatically adaptive.

VI. CONCLUSIONS

In the present paper, the authors studied the use of several systems for horse gait classification. The results of these classifiers were compared with the ones obtained from an embedded implementation of an ANN on the collar that animals carried using the FANN library. The accuracy ratio from both approaches are very similar (90.3%), proving that the expected results obtained after the spectral and statistical analyses of the dataset can also be achieved in real time in a real scenario using a different classifier, which is optimized for low-power MCUs, and similar preprocessing algorithms.

The system was tested in Doñana National Park with six different horses, satisfying both the requirements of the MINERVA project and the needs of the biologists that work in the park, since this classifier allows to monitor horse gaits and upload the information to a remote database server in real time. Currently, the device with an embedded ANN is being used in the park with several horses to test the behavior of the whole system for a long time period in terms of reliability and resistance, obtaining very good results up to date.

Future work will be focused on improving gait classification by testing new classifiers such as Self-Organizing Maps, adding more gaits like galloping or behaviors like eating and sleeping, implementing an over-the-air programming mechanism to change the ANN parameters and configuration remotely, and finally, making a smaller version of the collar for different species other than horses.

ACKNOWLEDGMENT

This work is supported by the excellence project from the Andalusiam Council MINERVA (P12-TIC-1300), and also by the Spanish government grant (with support from the European Regional Development Fund) COFNET (TEC2016-77785-P). The authors would like to thank Ramon C. Soriguer, Francisco Carro, Francisco Quirós and the EBD-CSIC for their support on the tests that were done in Doñana National Park.

REFERENCES

[1] E. Schlecht *et al.*, "The use of differentially corrected global positioning system to monitor activities of cattle at pasture," *Applied Animal Behaviour Science*, vol. 85, no. 3, pp. 185–202, 2004.

[2] E. D. Ungar *et al.*, "Inference of animal activity from GPS collar data on free-ranging cattle," *Rangeland Ecology & Management*, vol. 58, no. 3, pp. 256–266, 2005.

[3] M. Schwager *et al.*, "Robust classification of animal tracking data," *Computers and Electronics in Agriculture*, vol. 56, no. 1, pp. 46–59, 2007.

[4] E. S. Nadimi *et al.*, "Monitoring and classifying the behavior of a herd of sheep using ad hoc wireless sensor networks and artificial intelligence," in *7th World Congress on Computers in Agriculture Conference Proceedings, 22-24 June 2009, Reno, Nevada*. American Society of Agricultural and Biological Engineers, 2009, p. 1.

[5] E. Cerezuela-Escudero *et al.*, "Performance evaluation of neural networks for animal behaviors classification: Horse gaits case study," in *Distributed Computing and Artificial Intelligence, 13th International Conference*. Springer, 2016, pp. 377–385.

[6] D. Gutierrez-Galan *et al.*, "Embedded neural network for real-time animal behavior classification," *Neurocomputing*, Accepted for publication.

[7] M. Dominguez-Morales *et al.*, "Technical viability study for behavioral monitoring of wildlife animals in Doñana: An 802.15.4 coverage study in a Natural Park," in *Data Communication Networking (DCNET), 2011 Proceedings of the International Conference on*. IEEE, 2011, pp. 1–4.

[8] E. S. Nadimi *et al.*, "ZigBee-based wireless sensor networks for classifying the behaviour of a herd of animals using classification trees," *Biosystems engineering*, vol. 100, no. 2, pp. 167–176, 2008.

[9] D. Anderson, "Virtual fencing – a concept into reality," in *Proceedings of the Spatial Grazing Behavior Workshop.(Ed. GJ Bishop-Hurley.) pp*, 2006, pp. 61–91.

[10] E. S. Nadimi *et al.*, "ZigBee-based wireless sensor networks for monitoring animal presence and pasture time in a strip of new grass," *Computers and electronics in agriculture*, vol. 61, no. 2, pp. 79–87, 2008.

[11] H. Fourati *et al.*, "Posture and body acceleration tracking by inertial and magnetic sensing: Application in behavioral analysis of free-ranging animals," *Biomedical signal processing and control*, vol. 6, no. 1, pp. 94–104, 2011.

[12] C. Umstatter *et al.*, "An automated method of simple behavior classification as a tool for management improvement in extensive systems," in *Fourth International Workshop on Smart Sensors in Livestock Monitoring*, 2006, pp. 57–58.

[13] R. P. Wilson *et al.*, "Moving towards acceleration for estimates of activity-specific metabolic rate in free-living animals: the case of the cormorant," *Journal of Animal Ecology*, vol. 75, no. 5, pp. 1081–1090, 2006.

[14] L. González, G. Bishop-Hurley, R. N. Handcock, and C. Crossman, "Behavioral classification of data from collars containing motion sensors in grazing cattle," *Computers and Electronics in Agriculture*, vol. 110, pp. 91–102, 2015.

[15] J. A. V. Diosdado, Z. E. Barker, H. R. Hodges, J. R. Amory, D. P. Croft, N. J. Bell, and E. A. Codling, "Classification of behaviour in housed dairy cows using an accelerometer-based activity monitoring system," *Animal Biotelemetry*, vol. 3, no. 1, p. 15, 2015.

[16] "Doñana National Park," http://whc.unesco.org/en/list/685, Accessed: 2017-06-04.

[17] J. P. Dominguez-Morales *et al.*, "Wireless sensor network for wildlife tracking and behavior classification of animals in Doñana," *IEEE Communications Letters*, vol. 20, no. 12, pp. 2534–2537, 2016.

[18] "STM32L152," http://www.st.com/en/microcontrollers/stm32l152vb.html, Accessed: 2017-06-04.

[19] "XBee PRO S2B," https://www.digi.com/support/productdetail?pid=4549, Accessed: 2017-06-04.

[20] IEEE 802 Working Group *et al.*, ""IEEE Standard for Local and metropolitan area networks - Part 15.4: Low-Rate Wireless Personal Area Networks. Amendment 1: MAC sublayer"," *IEEE Std*, vol. 802, pp. 4–2011, 2011.

[21] IEEE 802 Working Group and others, ""IEEE Standard for Local and metropolitan area networks - Part 15.4: Low-Rate Wireless Personal Area Networks. Amendment 2: Active Radio Frequency Identification (RFID) System Physical Layer (PHY)"," *IEEE Std*, vol. 802, pp. 4–2011, 2011.

[22] D. D. Brown *et al.*, "Observing the unwatchable through acceleration logging of animal behavior," *Animal Biotelemetry*, vol. 1, no. 1, p. 20, 2013.

[23] S. Nissen, "Implementation of a Fast Artificial Neural Network library (FANN)," *Report, Department of Computer Science University of Copenhagen (DIKU)*, vol. 31, p. 29, 2003.

[24] A. Rios-Navarro *et al.*, "A sensor fusion horse gait classification by a spiking neural network on spinnaker," in *International Conference on Artificial Neural Networks*. Springer, 2016, pp. 36–44.