Automatic Recognition of Safety and Performance Related Activities in Motocross

Bas Breider

A thesis presented for the degree of
Master of Science in Mechanical Engineering

Student number:	4087461
Thesis committee:	Prof. dr. F.C.T. van der Helm
	Dr. J.C. van Gemert
	Prof. dr. H.E.J. Veeger
	Dr. ir. J.C.F. de Winter
	H.J. Ober

University of Technology Delft
August 2017

Preface

This report is the result of exploratory research done on the topic of activity recognition in motocross. The main content of this report is the scientific paper. This paper describes and validates an approach for the recognition of safety and performance related activities in motocross. The rest of the work consists of appendices where explanatory work can be found. The last appendix consists of the conducted literature study on activity recognition and the specific challenges in motocross. The research was done in collaboration with MYLAPS Sports Timing.

Before wrapping up I would like to acknowledge the valuable contribution of a number of people. First of all, I would like to thank dr. Jan van Gemert. He consistently allowed this paper to be my own work, but steered me in the right the direction whenever I needed it. Furthermore, my daily supervisor at MYLAPS Henk Jan Ober, who helped me focusing the problem and facilitated the project when needed. I would also like to acknowledge prof. dr. Frans van der Helm as the head of my graduation committee, and I am gratefully for his valuable comments on this thesis. Finally, I would like to thank the participants for volunteering.

Bas Breider,
Haarlem, August 2017

Contents

1 | Scientific Paper

Automatic Recognition of Safety and Performance Related Activities in Motocross

Bas Breider

Abstract—

Motocross is a popular, but dangerous sport: improvements in performance and safety should be made to make it more attractive and less dangerous. By automatically recognizing activities of the rider on the track, riders can be informed about dangerous situations, and fans can be provided with insights into the performance of the riders. The goal of this study is to develop and validate an automatic activity recognition methodology that can determine safety and performance related activities in motocross. A 3D accelerometer and gyroscope were used to collect movement data of the rider and motorcycle. Time and frequency domain features were extracted and used to evaluate several machine-learning classifiers: decision tree, k-nearest neighbor model, support vector machine, and multilayer perceptron neural network. These classifiers were evaluated based on accuracy, precision, recall, and speed to show overall classifier performance in real time, and to identify classification patterns for individual activities. The results were validated for multiple riders at different types of motocross tracks to test generalizability of the approach. Overall accuracy showed no large differences between the individual classifiers (74%-78% $\pm$ 6.8%). Similar results were found when the approach was validated with new riders and tracks (73%-79% and 68%-72%). The neural network classifier showed the highest precision for the safety related activities: *stopping* and *falling* (82%-95%). However, low precision was found for the performance related activities: jumping, *turning* and driving straight (20%-78%). To conclude, the neural network approach can be used for the detection of safety related activities, but more data of different riders is needed to confirm the proposed approach.

Keywords: Activity recognition, motocross, safety, performance, machine learning, classification.

I. Introduction

Motocross is a high-risk sport and is currently the most widely practiced motorcycle sport, both in Europe and the USA [1]. The popularity can be ascribed to the extreme physical and psychological demanding elements of motocross, which captivate the fans and media, but also expose riders to dangerous situations. Motocross racing is inherently risky due to the high speeds attained and the unpredictable nature of the tracks. Furthermore, the demanding competition, and in some cases inadequate training have increased the frequency and the severity of accidents [2]. Despite the increasing popularity of motocross racing and the associated harm it brings, little has been done to improve safety and performance in motocross. In this thesis, the improvement of safety and performance in motocross is addressed.

A potential technique for improving safety and performance that is already used in other related fields as sports [3] and automotive [4], is automatic activity recognition. One of the advantages of automatic activity recognition is that it could determine dangerous situations. Currently in motocross, marshals alarm upcoming riders with a flag signal when they detect a dangerous situation. Hence, the upcoming rider can slow down and avoid the dangerous situation the rider in front of them is causing. However, lack of volunteers and their limited experience [5] make it difficult to ensure safety on the track. Automatic recognition of dangerous activities such as *falling* and *standing still*, could assist or eventually replace marshals. In addition, by recognizing other motocross techniques as *jumping*, *turning* and *driving*, valuable information on the rider's behavior can be provided. For instance, the users can easily retrieve specific movement data for performance analysis [6] and provide insights to increase fan experience [7]. All of these potential applications benefit from accurate automatic recognition of motocross activities. Therefore, accurate automatic recognition of safety and performance related activities is explored.

To the best of our abilities, we have found no scientific literature on activity recognition in motocross. In non-racing motorcycling, several studies have been conducted to recognize activities thereby improve knowledge on driver behavior and safety [8], [9]. In both studies, street motorcycles were equipped with lightweight accelerometer and gyroscope sensors that measured three-dimensional (3D) movements of the rider. These small sensors were used because they are easily attached to the motorcycle and have low power consumption in comparison with another frequently used activity recognition sensor: the camera. The recorded accelerometer and gyroscope data was then used to evaluate multiple classifiers: decision trees, support vector machines, and k-nearest neighbor models. These classifiers were able to recognize activities like *turning*, *stopping*, and *driving straight* correctly with an average accuracy of 84%. Besides the usual driving activities, undesired situations like *falling*, could be distinguished from natural driving activities when using accelerometer and gyroscope data as well [10]. However, the threshold-based classifier used in this study is tuned on specific street motorcycle domain knowledge and therefore difficult to transfer with the limited scientific knowledge of motocross available. As alternative, a neural network classifier is proposed: the multilayer perceptron. This classifier seems suitable to motocross, because of its capability to identify complex underlying characteristics of nonlinear data [11]. In this thesis, accelerometer and gyroscope motocross activity data is acquired and used to evaluate multiple machine learning classifiers: decision trees, support vector machines, k-nearest neighbor, and multilayer perceptron neural network.

Not all activity recognition solutions adopted in non-racing motorcycling can be directly applied to motocross. To customize the activity recognition process to motocross, some extra challenges arise, predominantly caused by the motocross bike's dynamical capabilities and the challenging conditions of the motocross track. In the data acquisition process a large variety of activity execution is expected, because of the bumpy and unique track composition [12]. Therefore, the accelerometer and gyroscope sensors' ranges must be large enough to capture all movements and withstand extreme shocks and vibrations. The algorithm design process is also effected extreme dynamic demands of motocross. The vibrations caused by the engine create unwanted noise in the data and decreases the classifier's capabilities to predict activities correctly [10]. For the evaluation of the adopted motocross approach and its possible applications, this paper does take into account the real time aspect of the designed setup. The effectiveness of such a system is dependent on the ability to process the data quickly. The proposed approach is adapted to the extreme conditions during motocross racing and challenges for the recognition of the motocross activities.

The goal of this study is to develop and validate an activity recognition methodology that can recognize safety and performance related activities in motocross. This paper proposes to use machine learning algorithms to classify accelerometer and gyroscope data into six different motocross activities: *left turn*, *right turn*, *drive straight*, *jump*, *fall*, and *stop*. The choice of these selected activities was made to represent the most relevant activities related to safety and performance in motocross racing. The focus of the evaluation is to compare the effectiveness of the proposed classifiers on a real motocross dataset in terms of accuracy and potential real time use. Four machine learning classifiers are compared: decision tree (DT), support vector machine (SVM), k-nearest neighbor model (k-NN), and multilayer perceptron neural network (NN). Additionally, this study is used to evaluate how well the proposed approach can be generalized across different riders and tracks. One of the other facets of the work presented in this paper is the creation of the first database in which motocross movement data is included. This dataset can be used in future studies that focus on analyzing motocross rider's behavior.

The remainder of this paper is organized as follows. The data collection process, including the sensor and dataset description, is presented in Section II. Section III includes the proposed methodology for analyzing the data in terms of preprocessing, feature extraction, model learning, and evaluation. The performance of the different machine learning approaches is presented in section IV. These results are discussed in section V. The final section concludes the study and suggests future directions for this work.

II. Data Collection

In this section, we introduce the sensors, activities, data acquisition process, and resulting datasets.

A. Sensors

In this study, the dynamics of the rider and the motorcycle were recorded using a tri-axial accelerometer (range: ± 160 m/s^2, resolution: 0.05 m/s^2) and a tri-axial gyroscope (range: ± 2000 $°/s$, resolution: 0.08 $°/s$). The accelerometer and gyroscope were set to a sampling rate of 1 kHz, which is needed in highly dynamic driving situations [12]. In this paper, only the data collected with the accelerometer and gyroscope were used for classification. Therefore, the observations from the sensors are defined as:

$$\mathbf{x} = (a_x, a_y, a_z, r_x, r_y, r_z) \tag{1}$$

where a_x, a_y, and a_z are the longitudinal, vertical, and lateral accelerations respectively: r_x, r_y, and r_z are the roll, yaw, and pitch angular velocities respectively.

Fig. 1. Placement of the sensor on the motorcycle including the references axis orientation.

The accelerometer and gyroscope have to operate wireless and without constraining motion of the rider. Therefore, the accelerometer and gyroscope were combined into a standalone wearable sensor device. The sensor functionalities were implemented inside a 100 mm x 60 mm x 30 mm aluminum case and mounted on the left front of the motorcycle (Figure 1). Aluminum housing was used to make the sensor device strong enough to withstand impacts during the race, but at the same time make it lightweight so it could be easily attached to the motorcycle. A 2000 MB memory drive was placed inside the housing to save the data locally. By placing the sensor device on the motorcycle, the movements were measured without intervention of the researchers.

A ground truth is needed to evaluate the performance of the approach. The ground truth is typically defined through video analysis. For this purpose, one action camera was embedded on the motorcycle helmet (Figure 2). The camera recorded activities with a frame rate of 50 fps and a resolution of 720p.

Fig. 2. Placement of the action camera on top of the helmet. The top mount provides a complete view of the rider's movements.

The ground truth was then used to provide the accelerometer and gyroscope data with an activity label. Labeling is achieved by linking the sensor measurements with the ground truth. Global positioning system (GPS) was used as tool to provide time synchronization between the accelerometer, gyroscope, and video data. Both the accelerometer/gyroscope sensor device and the action camera were equipped with GPS antenna/module. The GPS adds a standardized global timestamp to sensor and video data. This allowed a comparison between the timestamp of ground truth activities set by the camera and the timestamp of the sensor data. With the GPS time synchronization, the data of the accelerometer and gyroscope sensors were provided with an activity label. (See Appendix A for more information about the labeling process.)

B. Activities

There are two categories for the target activities: safety related activities and performance related activities. Safety related activities such as *stopping* and *falling* are in contradiction with the dynamic nature of motocross racing and therefore often indicate danger. Performance related activities involve motocross movements techniques such as *jumping*, *turning*, and *driving*. These activities are related to performance, because these activities are essential for riding fast. The activities used in this study and their descriptions are given in Table I.

TABLE I
LIST OF SELECTED ACTIVITIES. ACTIVITIES ARE CATEGORIZED IN SAFETY AND PERFORMANCE RELATED ACTIVITIES, WITH NAME AND DESCRIPTION OF THE RECORDED ACTIVITIES.

State	Activity	Description
Safety Activities	Fall	Rider hits the ground and lies horizontally
	Stop	Motorcycle is not moving while standing upright
Performance Activities	Drive straight	Rider is moving in a straight line
	Turn left	Rider is turning the motorcycle to the left
	Turn right	Rider is turning the motorcycle to the right
	Jump	Motorcycle comes clear of the ground with both wheels

C. Data acquisition

One of the evaluation criteria is to see if the proposed approach is generalizable across riders and tracks. Therefore a requirement for data acquisition was to recruit different motocross riders with different driving experience, thus depicting different behaviors faced with the same driving event. Furthermore, the conditions of the track should be different to get a complete view of the conditions during motocross.

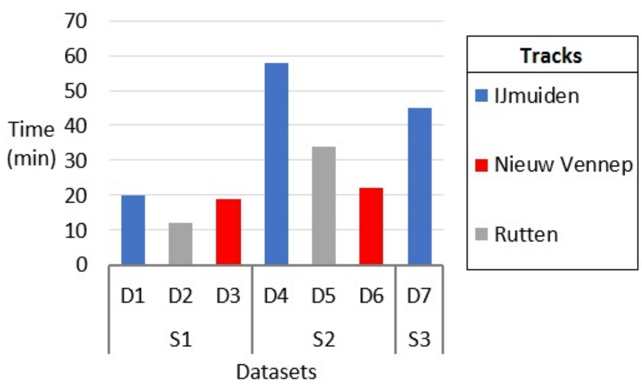

Fig. 3. Duration of each of the 7 datasets in minutes. Datasets ($D1$, $D2$, ..., $D7$) are presented with the corresponding rider ($S1$, $S2$, $S3$).

Data were collected at three different motocross tracks in the Netherlands: Rutten, IJmuiden, and Nieuw-Vennep. The tracks are a combination of sand and clay tracks, which are the most common surface types of motocross tracks [13]. Three subjects ($S1$, $S2$, and $S3$) with different profiles (age: 16-52 years) and motocross riding experiences (motocross experience: 4-35 years) participated in the data acquisition process. Subjects gave informed consent to participate. The riders used a motorcycle that fitted their physical capabilities: KTM 125SX, KTM 450EXC, and Kawasaki KX450F. No instructions were given on how and when to perform certain activities to ensure a naturalistic approach of motocross riding. For two of the riders ($S1$ and $S2$) data were collected on all three motocross tracks; the third rider ($S3$) was detained by unavoidable circumstances during two sessions and therefore data were only collected on one track (IJmuiden) for this rider. Figure 3 shows the duration of each dataset sorted by riders and tracks. Thus, the total dataset contains data from 3 different riders and 3 different tracks.

D. Datasets

For this study, the total dataset is composed of 7 individual datasets ($D1$, $D2$, ..., $D7$), where each one was performed according to the sequential activities uniquely corresponding to the track composition. Figure 4 shows the percentage of samples of each activity for each dataset. We can see that the different classes are not equally distributed. The activities *falling* and *jumping* are poorly represented compared to other activities. We also note that the majority of the sequences are composed of *driving straight* or *stopping* activities. To limit the influence of the imbalanced dataset on the performance, special attention must go out to evaluation measures address the performance for each individual activity.

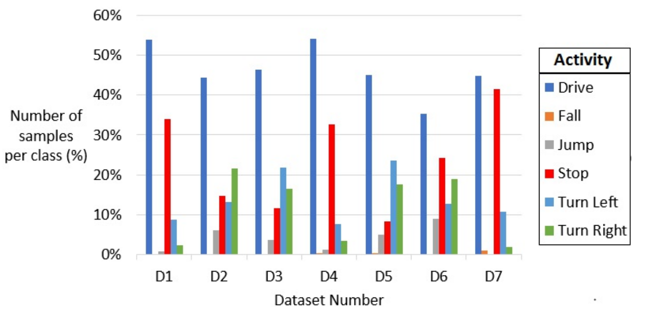

Fig. 4. Percentage samples per class for each dataset. Note that *fall* and *jump* are poorly represented in the datasets.

III. Approach

In this section, we present the proposed methodology including preprocessing, feature extraction, classification and performance evaluation. Figure 5 summarizes the different steps of the adopted approach.

A. Data preprocessing

The filtering step is important for separating meaningful data and noise. The accelerometer and gyroscope are sensitive to high frequency vibrations. Because of the noise and the variance in the collected data, some preprocessing tasks were required. Initially, a frequency analysis by power spectral was conducted to estimate the vibration characteristics of the motorcycle. The power spectrum of the vertical acceleration, i.e., a_y, shown in Figure 6, indicates that the frequency of the motorcycle noise in the static state (stop) is approximately 30 Hz. The high frequency power spectral peaks represent the noise caused by the vibration of the engine and therefore do not contain important characteristics of the activity movements. To remove the noise, data was filtered.

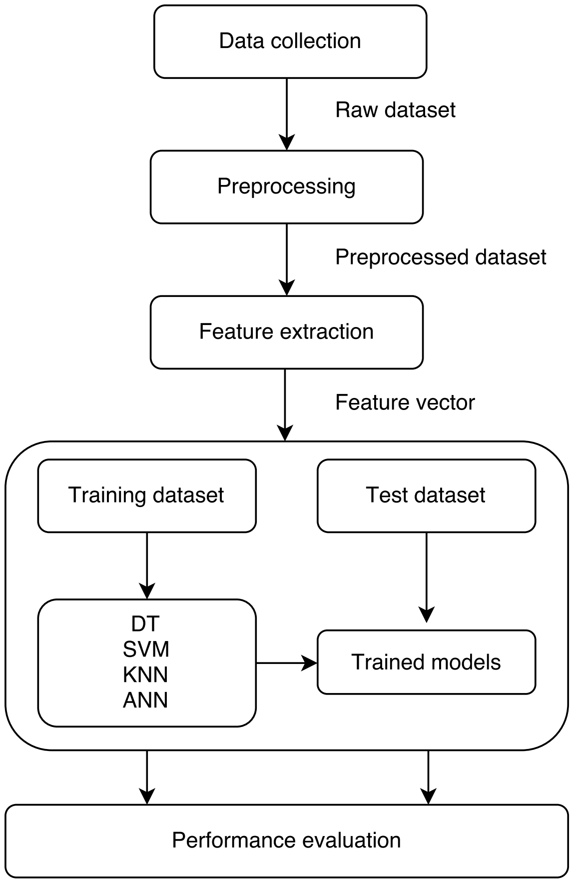

Fig. 5. Representation of data flow for testing and training activity recognition algorithm.

To remove high frequency noise, a low pass filtering technique was applied to the accelerometer and gyroscope data. For the design of the algorithm and the potential applications, it is important to operate in real time. Therefore, the choice for which low pass filtering technique to use, is based on the ability to remove noise and the execution time. The low-pass Butterworth filter is the best suited for our algorithm, because the execution time is shorter than other commonly used low pass techniques as wavelet and median filters [10]. A third-order low pass Butterworth filter with a 20-Hz cutoff frequency was used to remove the noise.

B. Feature extraction

Before the data was used as input for the classifier, feature extraction was done to extract informative insights of the data for each activity. Table II presents the features used in this study, sorted by domain.

1) Sliding window

To calculate the features from the activity data, the dataset first had to be segmented into windows. In this study, the fixed sliding window approach is used for segmentation. Since it does not require any preprocessing treatments, the sliding window approach is well-suited to real-time applications [14]. The windows are calculated with a certain overlap to handle transitions more accurately, and reduce misclassification due

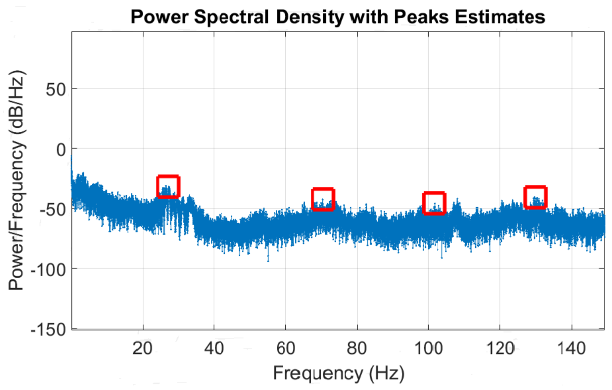

Fig. 6. Power spectral density for vertical acceleration (a_y) measured for stop activity. Power peaks indicated by red box. Note that first set of peaks is at approximately 30 Hz.

to transitions [15]. For each sliding window with a size of 0.5 s (500 samples each), a 50% (250 ms) overlap was used. The window size was set to be 0.5 seconds to ensure that short duration activities like jumping could be captured fully without including other activities. However, no smaller window size than 0.5 seconds was used to limit the effect of only capturing transitions between activities and not the specified activities.

TABLE II
LIST OF SELECTED FEATURES. THE FEATURES ARE CATEGORIZED INTO TWO GROUPS: SAFETY RELATED ACTIVITIES AND PERFORMANCE RELATED ACTIVITIES.

Domain	Feature
Time Domain	Mean Root mean square Maximum value Minimum value Range
Frequency Domain	Peak frequency Peak magnitude

2) Features

The accelerometer and gyroscope both have 3 dimensions in which data are collected. From these individual axis, characteristic information about the orientation can be extracted [15]. It is also possible to infer characteristics of an activity by using the vector norm of the 3D accelerometer and gyroscope signal. If the norm is used, the features calculated are independent of the orientation of the sensors. This will increase robustness to small placement errors. Therefore, in addition to the x, y, and z components, the Euclidean norm of these three axis was also used for feature extraction (Equation (2). This means that the accelerometer and gyroscope each have four components (x, y, z, and Euclidean norm) that are used for calculating features. Mean and root mean square have been used in previous activity recognition studies to characterize accelerometer and gyroscope data in the time domain [16]. As an extension of these features, we included the minimum and maximum value, and the range within the data window.

$$x = \sqrt{x^2 + y^2 + z^2} \tag{2}$$

Frequency domain features are used for discriminating static and dynamic activities in previous activity recognition studies [17]. The dynamic and static activity division is also seen in motocross activities: stopping and falling as (partly) static activities and driving and the related maneuvers as dynamic activities. To derive frequency domain features, a power spectrum was performed on each window of activity data. The peak frequency and magnitude of the first dominant peak in power spectrum of accelerometer and gyroscope data are used as frequency domain features. As with the time domain features, the frequency domain features were derived for each axis. A total of 56 (8 components x 7 features) features were calculated from each window of activity data.

3) Normalization

Since the ranges of raw data vary widely, in some machine learning algorithms, classifier functions will not work properly without normalization. For example, the majority of classifiers calculate the distance between two points by the Euclidean distance. If one of the features has a broad range of values, the distance will be governed by this particular feature [18]. Therefore, the range of all features should be normalized so that each feature contributes approximately proportionately to the final distance. So to avoid the domination of one signal over the others in the classification step, the data sets were normalized to a zero mean and a unit variance in all dimensions. The normalization equation is listed below:

$$x' = \frac{x - \bar{x}}{\sigma} \tag{3}$$

where x is the original feature vector, $\bar{x}$ is the mean of that feature vector, and σ is the standard deviation.

C. Classifiers

The aim of the work is to automatically recognize motocross activities. To evaluate the approach, we used four machine-learning classifiers: decision tree, support vector machine, k-nearest neighbor model, and multilayer perceptron neural network.

1) Decision trees

A decision tree is an oriented graph formed by a finite number of nodes departing from the root node. In binary trees, each parent node is linked to only two lower level nodes. A branch of the tree is a sub-tree obtained by pruning the tree at a given internal node [19]. In the tree growing, predictors generate candidate splits at each internal node of the tree, so that a suitable criterion needs to be defined to choose the best split of the features. In the case of the decision tree algorithm, we used Gini diversity index as split criterion (Equation (4)). Gini index is a measure of how often a randomly chosen element from the set would be incorrectly labeled if it was randomly labeled according to the distribution of labels in the subset. Gini index can be computed by summing the probability p_j of an activity with label i being chosen times the probability $1-p_j$ of a mistake in categorizing that activity. The split selected minimizes the value of the difference between Gini index before splitting and Gini index after splitting. It reaches its minimum (zero) when all cases in the node fall into a single target category.

$$GiniIndex = 1 - \sum_{j=1}^{i} p_j^2 \quad (4)$$

The only parameter to tune is the number of splits, which is determined by varying the number of trees from to 1 to 100. We picked the number of splits that provided the best accuracy rate when cross-validated on the datasets. The number of splits used in the template is 30.

2) Support vector machine

The support vector machine is based on the idea of calculating maximal margin hyperplanes that separate the data in a way that the margin between data points of each activity is maximized. For nonlinear accelerometer and gyroscope data, the data is mapped to a higher dimensional space using a kernel function [20]. Correctly classified points lying outside the margin boundaries of the support vectors are not penalized, whereas points within the margin boundaries or on the wrong side of the hyperplane are penalized in a linear fashion compared to their distance from the correct boundary by the hinge loss function (Equation (5)). In this hinge loss function L, y is the output of the classifier, t is the intended output. It can be seen that when t and y have the same sign, meaning that the support vector machine predicts the right class, the hinge loss is zero. The support vector machine performs well on data sets that have a large amount of attributes. However, it also is able to classify activities when data sets contain very few cases on which to train the model, in our case the *fall* activity.

$$L = max(0, 1 - y \cdot t) \quad (5)$$

For the support vector machine classifier with the Gaussian radial basis function kernel, the two hyper parameters to be tuned are C and γ. These parameters were estimated with Bayesian optimization. This algorithm uses previous observations of the loss to determine the next (optimal) point to sample for. The optimal values are C = 7 and $\gamma = 1$.

3) k-Nearest neighbor

The objective of the k-nearest neighbor model is to classify new data points based on the similarities that they exhibit with examples of the learning database using distance functions. The new observation is assigned to the most common activity through a majority vote of its k nearest neighbors. To classify unknown data points, the Euclidean distance of the new data point p to the known data points q is computed (6). Its k-nearest neighbours are determined, and the activity labels of these neighbours are then used to identify the activity label of the unknown object. The k-nearest neighbor model can handle missing values and is resistant to outlying data points, which are often seen in accelerometer and gyroscope sensors.

$$D(p, q) = \sqrt{\sum_i (p_i - q_i)^2} \quad (6)$$

In the case of kNN method, the best choice of k was tuned by varying k from 1 to 20. The optimal value of $k = 5$ led to the best combination of high accuracy and lowest computing time with cross-validation for the datasets.

4) Neural network

The multilayer perceptron is an artificial neural network with multilayer feed-forward architecture and is in general based on non-linear activations for the hidden units [21]. This neural network minimizes the error function between the estimated y and the desired network outputs d (Equations (7) (8)) which represent the activity labels in the classification context. The neural network uses backpropagation for training the network, which allows the network to converge on a satisfactory feature weighting and flow. For this study, a two-layer multilayer perceptron neural network with sigmoid hidden and softmax output neurons, is constructed.

$$e_j(n) = d_j(n) - y_j(n) \quad (7)$$

$$E(n) = \frac{1}{2} \sum_j e_j^2(n) \quad (8)$$

The amount of neurons in the hidden layer was varied from 10 to 30. The number of hidden neurons was chosen to be 15 based on optimal combination of accuracy and training time. In total, this led to a neural net with 56 input neurons (equal to the number of features), 15 hidden neurons, and 6 output neurons (equal to the number of activity classes).

D. Evaluation

1) Performance measures

The performance of the proposed method must be evaluated to gain insights on the classifier performance. A general measure used to evaluate the classifiers performances was accuracy, which measures the proportion of correctly classified examples. The accuracy can be expressed as follows:

$$Accuracy = \frac{T_p + T_n}{T_p + F_p + T_n + F_n} \quad (9)$$

where T_p (true positives) represent the correct classifications of positive examples and T_n (true negatives) represent the correct classifications of negative examples. F_n (false negatives) and F_p (false positives) represent, respectively the positive examples incorrectly classified into the negative classes and the negative examples incorrectly classified into the positive classes. However, the accuracy measure does not give a complete view of the classifier performance when unbalanced datasets are used, because the accuracy is particularly biased to favor the majority classes.

In this study, the class proportions are not well balanced since the number of *falling* and *jumping* activity samples is small compared to *driving straight* and *stopping* samples (Figure 4). To avoid the influence of the class imbalance, other metrics such as precision, recall, and F-measure are calculated. The precision is the number of correctly classified positive instances to the total number of instances classified as positive. Alternatively, the recall is the ratio of correctly classified positive instances to the total number of positive instances. So high precision will lead to more relevant results than irrelevant ones, while high recall implies that the most of the relevant results are classified correctly [22]. The F-measure is defined as the combination of two criteria, the precision and the recall, which are defined as follows:

$$Precision = \frac{T_p}{T_p + F_p} \tag{10}$$

$$Recall = \frac{T_p}{T_p + F_n} \tag{11}$$

$$F - measure = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \tag{12}$$

To summarize, the following evaluation criteria are considered: the average of the accuracy rate and its standard deviation (std), recall, precision, and f-measure. The combination of these measures gives an indication of the overall performance, and performance on the individual activities.

2) Validation

The dataset acquired during the motocross tests are used for validation. First, the 7 datasets were checked for abnormalities by training a classifier for each dataset and validate this classifier by a k-fold cross validation [23]. The data set is randomly partitioned into k subsets, and the holdout method is repeated *k* times. Each time, one of the *k* subsets is used as the test set and the remaining *k-1* subsets are put together to form a training set. The cross-validation process is repeated *k* times, with each of the k subsets used exactly once as the validation data. Then the average error across all *k* trials is computed. Using this method each dataset can be evaluated individually.

For a further analysis of the dataset the cross-validation method was used. This technique was used for assessing the average results of the classifiers for the datasets. One dataset was used for testing $(D = 1)$ and the remaining datasets were used for learning $(D = 6)$. This procedure was repeated 7 times (i.e., once for each sequence) and the average of the performance was calculated. After the overall analysis, the classifier that performs best in terms of statistical measures like accuracy, F-measure per activity, training time, and prediction time, was picked to be analyzed on its individual performance on safety and performance related activities.

Another specific contribution of this paper is to determine if the classifiers are able to generalize across riders and tracks. These two paths were investigated by using leave-one-rider-out and leave-one-track-out procedure respectively. By leaving one rider out as dataset, and in the second experiment leaving one track out. This hold-out set should be as large as possible, to accurately represent the activity variation that may be expected. However, keeping a large part of the data from the training set gives the classifier less data to train on. Hence, a balance between the size of the training set and the size of the holdout set must be struck. To provide a split between test and training data, rider 2 is used as test in the leave-one-rider-out validation and track Nieuw-Vennep is left out as training set to test the capabilities to generalize across tracks. Using this split, all activities are present in training and test sets and therefore all activities can be evaluated individually.

All algorithms were built with a custom made code run by a commercial software package (MATLAB Release 2016b, The MathWorks, Inc., Natick, Massachusetts, United States. Academic licence provided by Delft University of Technology).

IV. Results

In this section, we review and compare the performances of the machine learning approaches to recognize motocross activities.

A. Dataset investigation

The results obtained with the different classification approaches for each individual dataset are given in Table III. The data show that the average classifier results were above 80%, except for dataset 5. The classification accuracy for dataset 5 reported an average accuracy of approximately 60% for all classifiers. Dataset 5 differed significantly from all other datasets ($p<0.05$). These results indicate that there is difference in performance, which must be analyzed before using the dataset for experimental results.

TABLE III
ACCURACY OF DIFFERENT CLASSIFIER TECHNIQUES FOR EACH DATASET. FOR EACH DATASET, THE BEST CLASSIFIER IS IN BOLD. NOTE THAT THE ACCURACY OF DATASET 5 (66.7%) IS SIGNIFICANTLY LOWER COMPARED TO THE OTHER DATASETS ($p<0.05$).

	D1	D2	D3	D4	D5	D6	D7
Decision Tree (%)	85.0	80.1	82.0	84.7	59.7	79.8	89.2
k-Nearest Neighbor (%)	87.0	85.7	85.0	86.2	**66.7**	**82.6**	**91.0**
Support Vector machine (%)	**87.7**	**86.9**	**87.1**	**87.3**	64.9	82.5	90.9
Neural Network (%)	86.3	85.4	85.1	85.9	57.2	81.1	86.6

An analysis of the video footage in combination with the sensor data of dataset 5 showed that there was a failure in the experimental setup. The sensor attachment on the front fork came loose and as a result the sensor was able to rotate independently of the front fork. Thus, the accelerometer and gyroscope orientation changed during the experiment, which resulted in erratic sensor values. As this error in sensor placement does not confirm the defined experimental setup, dataset 5 was excluded from further analysis. From this point on, the total number of datasets used for classification results was reduced to $D = 6$.

B. Cross-validation

One of the goals of this study is to compare the performance of the machine-learning classifiers. The results of the classifier comparison using cross-validation are listed in Table IV. The table shows the average accuracy, F-measures, and speed of the classifiers. All correct classification rates exceeded 74%. The SVMs resulted in the highest overall classification rate at 78.1%. In addition to the overall accuracy, the SVMs also obtained the highest average F-measure (0.60). However, the other classifiers performed similarly on both the overall accuracy and average F-measure: the accuracy and F-measure of the NNs were 1% lower than SVM, the kNNs were 2% lower, and the DTs were 4% lower. The combination

TABLE IV
CROSS-VALIDATION RESULTS IN TERMS OF AVERAGE ACCURACY, F-MEASURE PER ACTIVITY, AVERAGE F-MEASURE, TRAINING TIME, AND PREDICTION TIME.

	Accuracy (%) ± std	F-measure per activity						F-measure	Training Time (sec)	Testing Time (sec)
		Stop	Fall	Jump	Turn Right	Turn Left	Drive Straight			
Decision Tree	74.4 ± 7.4	0.92	0.35	0.02	0.34	0.50	0.78	0.59	3.9	**0.1**
k-Nearest Neighbor	76.0 ± 6.0	0.92	**0.48**	**0.14**	0.44	0.58	0.78	0.58	**0.4**	13.6
Support Vector Machine	**78.1** ± 7.1	**0.94**	0.31	0.11	0.44	**0.61**	**0.81**	**0.60**	123.5	14.0
Neural Network	77.8 ± 6.7	0.92	0.37	0.05	**0.47**	0.60	0.80	0.59	12.8	**0.1**

of small differences in overall accuracy and large standard deviation of the classifiers, do not indicate any differences in performance patterns. In line with these findings, we found that the differences in overall accuracy between the classifiers were not significant (p >0.05).

The F-measures in Table IV show the performance of the classifiers for each activity. The safety related activities *stop* and *fall* are on average best classified by SVM and kNN: 0.94 and 0.48 respectively. For the *stop* activity, the F-measure for all classifiers exceeded 0.90, however no classifier exceeded 0.50 for *falling*. Out of all performance related activities *jumping* was classified worst, reporting F-measures ranging from 0.02 - 0.14. The SVM obtained the best classification rates for *driving straight* (0.81), but the other classifiers scored similarly for this activity (0.78 - 0.80). These measures show that no classifier is generally superior in terms of overall classification and classification of each activity.

The classifier's capability to predict in real time is essential for usage in safety or performance applications. Two measures that show the speed of the classifier: training and testing time, are reported in Table IV. The SVM classifier, which reported the highest overall classification rate, needed the longest time to classify the activities in the test set. The NN and DT classifiers were less accurate, but had a prediction time of approximately 1% compared to the SVM and kNN classifier. Therefore, the DT and NN have a higher execution speed than the SVM and kNN.

C. Leave-one-rider-out validation

The classifiers ability to generalize across new riders is tested by leaving the dataset of rider 2 as test set and using the datasets of riders 1 and 3 as training sets.

The results of the classifiers for the leave-one-rider-out validation are presented in Table V. We can see that SVM classifier reports the highest accuracy of 79.0% correct classification. This is 0.9% higher than the accuracy found with cross-validation (78.1%). The other classifiers show a slight decline in accuracy relative to the repeated cross-validation (-1 - -4%). However in general, the overall performances did not differ significantly when validated on a new rider (p>0.05).

To identify the underlying classification patterns that are difficult to recognize from statistical means only, a confusion matrix is constructed for the NN classifier. The NN classifier is chosen above the other classifiers, because this classifier showed the combination of best overall F-measure (0.51) and lowest prediction speed (0.1 s). Above all, the NN classifier showed best performance for classifying the essential safety related activity: *fall*. The confusion matrix with the corresponding recall and precision for each activity is given in Table VI. (See Appendix B for the confusion matrices of the other classifiers.)

For the safety related activity *stop*, the NN classifier obtained a precision of 82%, and a recall of 95%. The other safety related activity *fall*, was predicted with a precision of 92%, but a recall of only 27%. The most of confusion occurred between the the performance related activities *driving straight*, *turning right*, *turning left*, and *jumping*. As can be seen, there was a large influence of *driving straight* on the performance of the other activities, predominantly due to predicting one of the other activities as *driving straight*. So regarding the individual activities, the safety related activities were recognized with a higher precision than the performance related activities.

TABLE VI
CONFUSION MATRIX OBTAINED WITH THE NEURAL NETWORK CLASSIFIER FOR LEAVE-ONE-RIDER-OUT VALIDATION.

NN		**Predicted activities**						
		Stop	Fall	Jump	Turn Right	Turn Left	Drive Straight	**Recall**
True activities	Stop (%)	4760	1	0	6	15	216	0.95
	Fall (%)	36	24	0	3	14	10	0.27
	Jump (%)	22	0	8	0	20	413	0.01
	Turn Right (%)	85	0	0	533	54	769	0.37
	Turn Left (%)	39	0	0	42	1029	1235	0.43
	Drive Straight (%)	808	1	32	159	872	8069	0.81
	Precision	0.82	0.92	0.2	0.71	0.51	0.75	

D. Leave-one-track-out validation

Another variable in the motocross datasets are the tracks. The generalization across tracks is tested by leaving the datasets recorded on motocross track in Nieuw-Vennep as test set and using the remaining datasets as training sets. The results of the leave-one-track-out validation method are shown in Table VII. The SVM obtained the highest overall accuracy of 71.4%, which is within one standard deviation from the SVM results of the cross-validation (78.1% ± 7.1%). The overall accuracy of NN and DT classifiers also laid within one standard deviation from the repeated cross-validation mean; the kNN reported an accuracy that was slightly more than one standard deviation lower. Although the performance of the classifiers in leave-one-track-out validation was worse than in repeated cross-validation, no classifier differed significantly (p>0.05).

In Table VIII the confusion matrix for neural net with leave-one-track-out validation is shown. We see most prediction errors occurred for the activities *turning*, and *jumping*, resulting

TABLE V

LEAVE-ONE-RIDER-OUT RESULTS IN TERMS OF ACCURACY, F-MEASURE PER ACTIVITY, AVERAGE F-MEASURE, TRAINING TIME, AND PREDICTION TIME. NOTE THE DIFFERENCES OF THE F-MEASURE BETWEEN THE CLASSIFIERS FOR THE *fall* ACTIVITY.

	Accuracy (%)	F-measure per activity						F-measure	Training Time (sec)	Testing Time (sec)
		Stop	Fall	Jump	Turn Right	Turn Left	Drive Straight			
Decision Tree	73.6	0.90	0.34	0.05	0.35	0.38	0.79	0.47	2.6	**0.1**
k-Nearest Neighbor	76.9	0.91	0.09	**0.12**	0.52	**0.47**	0.81	0.48	**0.3**	17.6
Support Vector Machine	**79.0**	**0.92**	0.09	0.08	**0.57**	0.46	**0.82**	0.49	59.4	9.2
Neural Network	74.8	0.89	**0.42**	0.03	0.49	**0.47**	0.78	**0.51**	8.8	**0.1**

TABLE VII

LEAVE-ONE-TRACK-OUT RESULTS IN TERMS OF ACCURACY, F-MEASURE PER ACTIVITI, AVERAGE F-MEASURE, TRAINING TIME, AND PREDICTION TIME.

	Accuracy (%)	F-measure per activiti						F-measure	Training Time (sec)	Testing Time (sec)
		Stop	Fall	Jump	Turn Right	Turn Left	Drive Straight			
Decision Tree	68.9	0.86	0.54	-	0.60	0.62	0.73	-	4.2	**0.1**
k-Nearest Neighbor	68.0	0.89	0.56	-	0.59	0.55	0.73	-	**0.3**	7.9
Support Vector Machine	**71.4**	**0.93**	0.69	0.05	0.61	**0.64**	**0.75**	0.61	123.5	4.5
Neural Network	70.1	0.90	**0.76**	**0.16**	0.61	**0.64**	0.73	**0.63**	9.9	**0.1**

in low recall values for these activities. In most cases, these activities are predicted as *driving straight*. Furthermore, this classifier was able to predict a *fall* with 95% precision and 64% recall. Out of the 9 classification errors, 7 times a *fall* was classified as *stop*. So the prediction errors of a *fall* were mainly due to the confusion with the other safety related activity *stop*. Overall the same patterns are seen: safety related activities are better classified than the performance related activities.

TABLE VIII

CONFUSION MATRIX OBTAINED WITH NEURAL NETWORK FOR THE LEAVE-ONE-TRACK-OUT VALIDATION.

NN		Predicted activities						
		Stop	Fall	Jump	Turn Right	Turn Left	Drive Straight	**Recall**
True activities	Stop (%)	914	1	0	8	5	55	0.93
	Fall (%)	7	18	0	1	0	2	0.64
	Jump (%)	0	0	40	0	1	399	0.10
	Turn Right (%)	2	0	0	836	22	861	0.48
	Turn Left (%)	1	0	4	15	1206	1049	0.53
	Drive Straight (%)	132	0	14	148	258	3993	0.88
	Precision	0.86	0.95	0.69	0.83	0.80	0.63	

V. DISCUSSION

In this section, we will discuss the results based on the goals of this study presented in the introduction.

A. Safety related activities

The goal of this study was to develop and validate an automatic activity recognition methodology that could recognize safety and performance related activities in motocross. First, we discuss the performance of the approach for safety related activities. The combined F-measure for *stop* and *fall* for the different classifiers suggest that the multilayer perceptron neural network is the best classifier for classifying safety related activities. The high precision and lower recall indicate that the classifier is careful when predicting a *fall*, and therefore sometimes misses a *fall*. For the application of a safety detection device in motocross, the high precision - low recall trade off is acceptable. Having irrelevant signals indicating a dangerous situation will lead to reduced trust in the system. As riders want to be as quick as possible to win the race, unnecessary stops will cause important loss of time. Eventually, the rider will not obey the system if it will lead to unneeded losses.

Ideally, the false negatives seen when predicting a *fall* are reduced, because missing a *fall* may lead to dangerous situations, especially in the case when such a safety alarm system would be used as replacement of the marshals that are responsible for safety. The confusion matrices show that the false negatives most of the time appear because a *fall* is classified as *stop*. Explanation of these results could be that after a *fall* a transition to a *stop* takes place. Due to the sliding window approach, it is likely that such a transition is labelled with either *stop* or *fall*, but includes characteristic features of the other activity as well. For the potential alarm system the confusion between a *fall* and *stop* would not give a problem as these activities are both considered unwanted during a race. Therefore this approach could only be used as assistance for the marshals on the track.

B. Performance related activities

The confusion matrices shown in Table VI and Table VIII show that the activities *jumping*, *turning*, and *driving straight* are often mixed up. The confusion between *driving*, *turning*, and *jumping* can be explained by the similarities between these activities. Firstly, labeling the data is dependent on the definitions the researchers give to an activity. For example, a turn can either be long or, in the most extreme case, be a hairpin turn. Where a long turn only requires the rider to slightly lean to the side, a hairpin turn can only be performed when the rider abruptly rotates his steer and leans horizontally. The long turn will have huge resemblance with *driving straight* and that might be the reason for the confusion between *turning* and *driving straight*. Secondly, motocross activities are often short in duration and alternate frequently. Therefore, many transitions are included in the data. These transitions have shown to give problems in classifying activities, because features of transitional data are less characteristic [24]. Another factor that may cause the confusion of the performance related activities, is the track composition [12]. The bumpy track could lead to a jump like motion during normal driving, causing a confusion between *driving* straight and *jumping*. To summarize, the confusion between the performance related activities indicates a difficulty in recognizing these activities, which can either be assigned to similarity due to dynamics of the motorcycle, variety of movement executions or subjective labeling.

C. Performance of the different classifiers

The focus of the evaluation is to compare the effectiveness of the proposed classifiers on a real motocross dataset in terms of accuracy and potential real time use. The results indicate that there is no big difference in accuracy between the four different classifiers. These findings are supported by [8], who similarly did not find any big differences between kNN, SVM, and DT classifiers accuracy in motorcycle activity recognition. However classifiers do differ in execution speed. The short prediction time of decision tree and neural network makes these classifier better suited for real time applications. Unlike decision trees, neural network were not expected to be as quick in predicting, because they use all features to classify, were the decision tree only evaluates one feature per node. One possible explanation and limitation of the speed measures is that is predicted on a laptop, which is known to have large computational power. Any application with wearable sensors would have less processing power of a computer. Therefore in future research, online classification on the wearable device must be evaluated to test the classifiers capability to operate in real time.

D. Generalizability

Additionally, this study is used to evaluate how well the proposed approach can generalize across different riders and tracks. The similarities in terms of overall performance between cross-validation and leave-one-out validation methods indicate that the classifiers can generalize to new motocross tracks and different riders. Although the differences between the repeated cross-validation and the leave-one-track-out were larger than the-leave-one-rider out, performance is similar when multiple riders and tracks are included in the training set.

We expected the neural network to outperform the other classifiers, based on its ability to model non-linear patterns [25] seen in accelerometer and gyroscope data. For generalization, the use of a neural network was also expected to have a higher accuracy than the other classifiers, because it should be able to infer unseen relationships in new complex data [26]. Still, the neural network did not outperform other classifiers when validated for new riders and tracks. A possible reason for this could be the imbalanced dataset (figure 4). By normalization and using class specific evaluation measures like precision and recall, we tried to reduce the effect of class imbalance on the classifiers. However, the weights of the neurons are already adapted to the skewness of the data in the training phase [27]. The neural networks may likely ignore the features representing the classes that have small number of examples in the training set. To overcome the problem of a imbalanced dataset, more samples of the underrepresented activities (*fall* and *jump*) should be recorded. When more training data is added, we assume that the neural network classifier will outperform the other classifiers.

E. Dataset

One of the other facets of the work presented in this paper is the creation of the first database in which motocross movement data is included. The accelerometer and gyroscope sensors were able to measure all motocross movements. The measurements were recorded with sampling rate of 1 kHz, which is recommended when recognizing the impact of *falling* [12]. However for a potential real time use, we would suggest to evaluate lower sampling rates to reduce computational load. This exploratory research showed that ground truth can be measured with an action camera, but the labeling is still subjective. For future motocross studies, we would suggest to adapt the labeling process to the requirements of the particular research topic. However, this complete dataset can be used in future studies that focus on analyzing motocross rider's behavior. (See Appendix C for examples of the database.)

VI. Conclusion & Future Work

In this paper, we presented an automatic activity recognition approach with the goal to detect dangerous and performance related activities in motocross. This is done by measuring 3D accelerometer and gyroscope motocross activity data to evaluate multiple machine learning classifiers. No large overall accuracy differences were found between the classifier performances. However, the proposed multilayer perceptron neural network approach shows the highest predictive power for the safety related activities: *stopping* and *falling*. In addition, the neural network approach had the lowest prediction time of the classifiers tested. The combination of high predictive power and good real time capabilities indicates that the neural network approach could be used as safety assistance device to assist marshals.

No satisfying performances of the approach were found for the recognition of the performance related activities: *jumping*, *turning*, and *driving straight*. Confusion between the activities was often encountered and therefore no good classification is guaranteed. The confusion is most likely caused by the similarities between the performance related activities and by the variety of different ways one activity can be executed. In conclusion, this approach is not suitable for any application that includes the recognition of the performance related activities.

We validated the generalizability of our method by testing the trained models with datasets of unseen riders and tracks. No differences in performance were noticed when we tested the approach on new riders and tracks. Therefore, the adopted approach should be able to generalize when used in a real life application. However in this study, only a small group of 3 riders is used to validate the approach. Future validation step of this approach should involve expansion of the dataset by adding more riders. This will also benefit the expansion of the first motocross database constructed in this study.

VII. Acknowledgements

This research was supported by MYLAPS Sports Timing.

References

[1] Alberto Gobbi. Motocross injuries: Incidence, management, and prevention. In *Sports Injuries*, pages 2841–2852. Springer, 2015.

[2] A Noelle Larson, Anthony A Stans, William J Shaughnessy, Mark B Dekutoski, Michael J Quinn, and Amy L McIntosh. Motocross morbidity: economic cost and injury distribution in children. *Journal of Pediatric Orthopaedics*, 29(8):847–850, 2009.

[3] Eduardo Velloso, Andreas Bulling, Hans Gellersen, Wallace Ugulino, and Hugo Fuks. Qualitative activity recognition of weight lifting exercises. In *Proceedings of the 4th Augmented Human International Conference*, pages 116–123. ACM, 2013.

[4] Donald Selmanaj, Matteo Corno, and Sergio M Savaresi. Hazard detection for motorcycles via accelerometers: A self-organizing map approach. 2016.

[5] Knmv motorsport organisatie vrijwilligers [online]. available: http://www.knmv.nl/motorsport/knmv/organisatie/vrijwilligers/.

[6] Le Nguyen Ngu Nguyen, Daniel Rodríguez-Martín, Andreu Català, Carlos Pérez-López, Albert Samà, and Andrea Cavallaro. Basketball activity recognition using wearable inertial measurement units. In *Proceedings of the XVI International Conference on Human Computer Interaction*, page 60. ACM, 2015.

[7] Tsutomu Terada, Masakazu Miyamae, Yasue Kishino, Takahito Fukuda, and Masahiko Tsukamoto. An event-driven wearable system for supporting pit-crew and audiences on motorbike races. *J. Mobile Multimedia*, 5(2):140–157, 2009.

[8] Ferhat Attal, Abderrahmane Boubezoul, Latifa Oukhellou, and Stéphane Espié. Powered two-wheeler riding pattern recognition using a machine-learning framework. *IEEE Transactions on Intelligent Transportation Systems*, 16(1):475–487, 2015.

[9] Ferhat Attal, Abderrahmane Boubezoul, Allou Samé, and Latifa Oukhellou. Powered-two-wheeler safety critical events recognition using a mixture model with quadratic logistic proportions. In *ESANN 2015-23rd European Symposium on Artificial Neural Networks*, pages pp–421, 2015.

[10] Abderrahmane Boubezoul, Stéphane Espié, Bruno Larnaudie, and Samir Bouaziz. A simple fall detection algorithm for powered two wheelers. *Control Engineering Practice*, 21(3):286–297, 2013.

[11] Leopoldo Marchiori Rodrigues and Mário Mestria. Classification methods based on bayes and neural networks for human activity recognition. In *Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), 2016 12th International Conference on*, pages 1141–1146. IEEE, 2016.

[12] Jussi Parviainen, Jussi Colliri, Timo Pihlström, Jarmo Takala, Kari Hanski, and Aki Lumiaho. Automatic crash detection for motor cycles. In *IECON 2014-40th Annual Conference of the IEEE Industrial Electronics Society*, pages 3409–3413. IEEE, 2014.

[13] Motocross and speedway tracks in europe [online]. available: http://www.mxbrothers.com/mx-tracks/europe.

[14] Stephen J Preece, John Yannis Goulermas, Laurence PJ Kenney, and David Howard. A comparison of feature extraction methods for the classification of dynamic activities from accelerometer data. *IEEE Transactions on Biomedical Engineering*, 56(3):871–879, 2009.

[15] Oscar D Lara and Miguel A Labrador. A survey on human activity recognition using wearable sensors. *IEEE Communications Surveys and Tutorials*, 15(3):1192–1209, 2013.

[16] Andreas Bulling, Ulf Blanke, and Bernt Schiele. A tutorial on human activity recognition using body-worn inertial sensors. *ACM Computing Surveys (CSUR)*, 46(3):33, 2014.

[17] Stephen J Preece, John Y Goulermas, Laurence PJ Kenney, Dave Howard, Kenneth Meijer, and Robin Crompton. Activity identification using body-mounted sensorsa review of classification techniques. *Physiological measurement*, 30(4):R1, 2009.

[18] Jhun-Ying Yang, Jeen-Shing Wang, and Yen-Ping Chen. Using acceleration measurements for activity recognition: An effective learning algorithm for constructing neural classifiers. *Pattern recognition letters*, 29(16):2213–2220, 2008.

[19] S Rasoul Safavian and David Landgrebe. A survey of decision tree classifier methodology. *IEEE transactions on systems, man, and cybernetics*, 21(3):660–674, 1991.

[20] Johan AK Suykens and Joos Vandewalle. Least squares support vector machine classifiers. *Neural processing letters*, 9(3):293–300, 1999.

[21] Jiexiong Tang, Chenwei Deng, and Guang-Bin Huang. Extreme learning machine for multilayer perceptron. *IEEE transactions on neural networks and learning systems*, 27(4):809–821, 2016.

[22] Jamie A Ward, Paul Lukowicz, and Gerhard Tröster. Evaluating performance in continuous context recognition using event-driven error characterisation. In *International Symposium on Location-and Context-Awareness*, pages 239–255. Springer, 2006.

[23] Payam Refaeilzadeh, Lei Tang, and Huan Liu. Cross-validation. In *Encyclopedia of database systems*, pages 532–538. Springer, 2009.

[24] Jorge-L Reyes-Ortiz, Luca Oneto, Albert Sama, Xavier Parra, and Davide Anguita. Transition-aware human activity recognition using smartphones. *Neurocomputing*, 171:754–767, 2016.

[25] Akram Bayat, Marc Pomplun, and Duc A Tran. A study on human activity recognition using accelerometer data from smartphones. *Procedia Computer Science*, 34:450–457, 2014.

[26] Ebrahim Al Safadi, Fahim Mohammad, Darshan Iyer, Benjamin J Smiley, and Nilesh K Jain. Generalized activity recognition using accelerometer in wearable devices for iot applications. In *Advanced Video and Signal Based Surveillance (AVSS), 2016 13th IEEE International Conference on*, pages 73–79. IEEE, 2016.

[27] Zongyuan Zhao, Shuxiang Xu, Byeong Ho Kang, Mir Md Jahangir Kabir, and Yunling Liu. Investigation of multilayer perceptron and class imbalance problems for credit rating. *International Journal of Computer and Information Technology*, 3(4):805–812, 2014.

Appendices

A | Labeling

A challenge for activity recognition tasks is the collection of annotated or ground truth labeled training data. Ground truth annotation is a time-consuming task, as the annotator has to go through the data and manually label all activity instances post hoc. In addition, motion data recorded from an accelerometer or gyroscope is often more difficult to interpret than data from other sensors, such as cameras. Therefore in this paper, the ground truth is established by a camera and linked to the sensor data of the accelerometer and gyroscope based on time. The process of time synchronization is shown in Figure A.1. In this appendix, we will describe the labeling process.

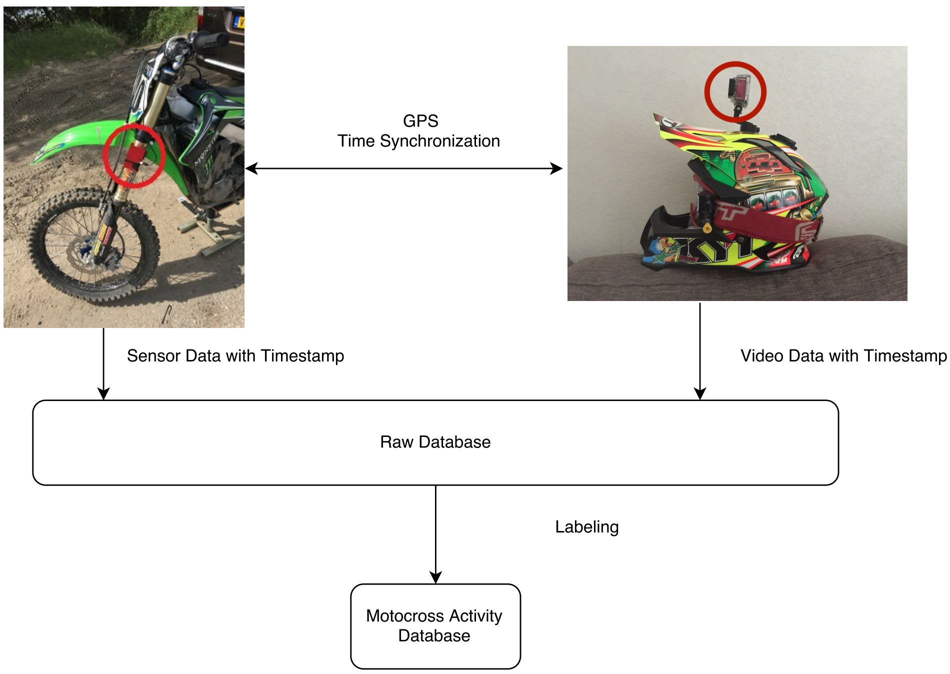

Figure A.1: GPS time synchronization between the sensor data and video data.

To label the accelerometer and gyroscope sensor data, the ground truth and sensor data must be synchronized. The first step in this process is to establish a ground truth. For the establishment of the ground truth, an action camera was used. This camera was fixed on top of the helmet to get a wide view of the rider and track. In the datasets, six different motocross activities are represented: *drive straight*, *jump*, *turn right*, *turn left*, *stop*, and *fall*. In Figure A, the six motocross activities are shown as seen in the raw video data.

Inside the action camera, a GPS module was built. The GPS provides timestamps to the video data, because each GPS satellite contains multiple atomic clocks that send time global signals to the device. The update rate of the GPS embedded in the action camera was 10 Hz. The time between the update points was calculated based on the frames per second. This means that with a frame rate of 50 fps, the maximum accuracy for video labeling is 0.02 s.

The next step is to label the activities seen in the data with their corresponding timestamp. The activities were manually labeled according to the activities seen in video data. The current time of each frame was shown in the video data. The start and stop time of each activity were determined by skipping through the video data frame by frame. This results in a large file in which for the total length of the video data, the activities were written down together with their start and stop time. Now the ground truth is established: all activities are provided with the corresponding timestamp.

To link the activity timestamps to the accelerometer and gyroscope sensor data, these sensor datapoints also needed a timestamp. Therefore, the accelerometer and gyroscope was attached to a GPS antenna. The GPS antenna is able to decode the satellite signals and therefore can add timestamps to the data of the accelerometer and gyroscope. The update rate of the GPS sensor was 12,5 Hz; the sampling rate of the sensor was 1000 Hz. To solve this problem, the timestamps for the sensor datapoints between the two GPS update points were interpolated.

The last step is to add the activity labels to the sensor data points based on their common timestamps. After this is done, each sensor data point has a corresponding activity label. The sensor data en activity labels can then be further processed and used to train the classifiers.

Figure A.2: Video data of the six motocross activities. Labels were added according to the activities seen in video data

B | Confusion matrices

In the paper provided in the first part of thesis, we analyzed the confusion matrices of the neural network. In this appendix, we provide the confusion matrices of the decision tree, support vector machine and k-nearest neighbor classifiers for both the leave-one-rider-out validation and the leave-one-track-out validation. The confusion matrices are shown in the tables below. Tables B.1, B.2, and B.3 are the confusion matrices validated with the leave-one-rider-out method for the k-nearest neighbor, support vector machine and decision tree classifier respectively. Tables B.4, B.5, and B.6 are the confusion matrices validated with the leave-one-track-out method for the k-nearest neighbor, support vector machine and decision tree classifier respectively.

Table B.1: Confusion matrix obtained with k-nearest neighbor model, validated with leave-one-rider-out method

NN

		Predicted activities						
		Stop	Fall	Jump	Turn Right	Turn Left	Drive Straight	**Recall**
True activities	Stop (%)	4539	2	1	7	8	441	0.91
	Fall (%)	55	4	0	5	10	13	0.05
	Jump (%)	5	0	34	8	8	408	0.07
	Turn Right (%)	14	0	0	616	120	691	0.43
	Turn Left (%)	24	1	6	43	975	1296	0.42
	Drive Straight (%)	329	0	71	214	672	8655	0.87
	Precision	0.91	0.57	0.30	0.69	0.54	0.75	

Table B.2: Confusion matrix obtained with decision tree, validated with leave-one-rider-out method

NN

		Predicted activities						
		Stop	Fall	Jump	Turn Right	Turn Left	Drive Straight	**Recall**
True activities	Stop (%)	4305	2	0	0	7	684	0.86
	Fall (%)	4	21	0	0	8	54	0.24
	Jump (%)	0	0	16	0	21	426	0.03
	Turn Right (%)	6	0	2	339	289	805	0.24
	Turn Left (%)	11	2	1	15	826	1490	0.35
	Drive Straight (%)	240	11	77	131	795	8687	0.87
	Precision	0.94	0.58	0.17	0.70	0.42	0.72	

Table B.3: Confusion matrix obtained with support vector machine, validated with leave-one-rider-out method

NN

		Predicted activities						
		Stop	Fall	Jump	Turn Right	Turn Left	Drive Straight	**Recall**
True activities	Stop (%)	4625	3	0	0	6	364	0.93
	Fall (%)	52	4	0	2	10	13	0.04
	Jump (%)	0	0	23	0	3	428	0.05
	Turn Right (%)	19	0	0	643	100	679	0.45
	Turn Left (%)	12	0	5	36	860	1432	0.37
	Drive Straight (%)	295	0	54	141	374	9077	0.91
	Precision	0.92	0.57	0.28	0.78	0.64	0.76	

Table B.4: Confusion matrix obtained with k-nearest neighbor model, validated with leave-one-track-out method

NN

		Predicted activities						
		Stop	Fall	Jump	Turn Right	Turn Left	Drive Straight	**Recall**
True activities	Stop (%)	929	6	0	0	0	48	0.95
	Fall (%)	5	14	0	1	0	8	0.50
	Jump (%)	3	1	0	0	0	436	-
	Turn Right (%)	8	0	0	774	28	911	0.45
	Turn Left (%)	2	0	0	13	1125	1135	0.49
	Drive Straight (%)	220	2	0	77	196	4050	0.89
	Precision	0.79	0.61	-	0.89	0.83	0.61	

Table B.5: Confusion matrix obtained with decision tree, validated with leave-one-track-out method

NN

		Predicted activities						
		Stop	Fall	Jump	Turn Right	Turn Left	Drive Straight	**Recall**
True activities	Stop (%)	891	0	0	0	1	91	0.91
	Fall (%)	5	11	0	0	0	12	0.39
	Jump (%)	13	0	0	1	0	426	-
	Turn Right (%)	10	0	0	776	38	897	0.45
	Turn Left (%)	19	0	0	4	936	1316	0.41
	Drive Straight (%)	88	0	0	115	153	4189	0.92
	Precision	0.86	1.00	-	0.87	0.83	0.60	

Table B.6: Confusion matrix obtained with support vector machine, validated with leave-one-track-out method

NN

		Predicted activities						
		Stop	Fall	Jump	Turn Right	Turn Left	Drive Straight	**Recall**
True activities	Stop (%)	932	0	0	0	3	48	0.95
	Fall (%)	5	15	0	0	1	7	0.54
	Jump (%)	1	0	11	0	0	428	0.03
	Turn Right (%)	1	0	0	781	22	917	0.45
	Turn Left (%)	2	0	0	14	1153	1106	0.51
	Drive Straight (%)	60	0	0	56	184	4245	0.93
	Precision	0.93	1.00	1	0.92	0.85	0.62	

C | Database

One of the contributions of this work is the creation of motocross activity database. The raw data from the motocross tests was stored into a .txt file to get a clear structured dataset. Two examples of the database are presented below. Figure C shows the raw data acquired directly from the accelerometer, gyroscope and GPS. Figure C shows the data after it is labeled. This data was used as input for the feature extraction process. The following parameters can be found in the excel sheets, in sequence of display:

Miscellaneous
Time: time in seconds from the moment the device is turned on
Dist: distance in meters calculated by the GPS

Accelerometer
ACCX: Accelerations in m/s^2 in x-direction
ACCY: Accelerations in m/s^2 in y-direction
ACCZ: Accelerations in m/s^2 in z-direction

Gyroscope
GYROX: Angular velocity in $^{\circ}/s$ in x-direction
GYROY: Angular velocity in $^{\circ}/s$ in y-direction
GYROZ: Angular velocity in $^{\circ}/s$ in z-direction

Global Positioning System
MMDD: First 2 digits represent months, last 2 digits represent days
HHMM: First 2 digits represent hours, last 2 digits represent minutes
SSHH: First 2 digits represent seconds, last 2 digits represent hundredths
GPSVALID: 1 if GPS is connected to satelites, 0 if no connection is available

Label
ActID: Activity label

A	B	C	D	E	F	G	H	L	M	N	O
Time	Dist	ACC_X	ACC_Y	ACC_Z	GYRO_X	GYRO_Y	GYRO_Z	MMDD	HHMM	SSHH	GPSValid
NUMBER	NUMBER	NUMBER	NUMBER	NUMBER	NUMBER	NUMBER	NUMBER	NUMBER	NUMBER	NUMBER	NUMBER
Time	Dist	ACC_X	ACC_Y	ACC_Z	GYRO_X	GYRO_Y	GYRO_Z	MMDD	HHMM	SSHH	GPSValid
s	m	m/s^2	m/s^2	m/s^2	deg/s	deg/s	deg/s				°
0,000	0	-9,89	-1,55	1,60	-2,32	-0,32	0,40	6,16	10,25	42,70	1,0
0,001	0	-9,89	-1,55	1,58	-2,24	-0,40	0,24	6,16	10,25	42,70	1,0
0,002	0	-9,89	-1,56	1,57	-2,16	-0,48	0,08	6,16	10,25	42,70	1,0
0.003	0	-9,89	-1,57	1,57	-2,08	-0,64	-0,00	6,16	10,25	42,70	1,0
0,004	0	-9,89	-1,57	1,55	-2,08	-0,72	-0,16	6,16	10,25	42,70	1,0
0,005	0	-9,89	-1,58	1,54	-1,92	-0,88	-0,24	6,16	10,25	42,70	1,0
0,006	0	-9,88	-1,57	1,53	-1,84	-0,96	-0,32	6,16	10,25	42,70	1,0
0,007	0	-9,87	-1,57	1,52	-1,76	-1,12	-0,48	6,16	10,25	42,70	1,0
0,008	0	-9,87	-1,58	1,52	-1,60	-1,20	-0,48	6,16	10,25	42,70	1,0
0,009	0	-9,87	-1,57	1,51	-1,52	-1,28	-0,64	6,16	10,25	42,70	1,0
0,010	0	-9,87	-1,57	1,50	-1,28	-1,36	-0,72	6,16	10,25	42,70	1,0
0,011	0	-9,86	-1,57	1,50	-1,12	-1,44	-0,80	6,16	10,25	42,70	1,0
0,012	0	-9,85	-1,57	1,50	-0,96	-1,44	-0,96	6,16	10,25	42,70	1,0
0,013	0	-9,84	-1,57	1,50	-0,80	-1,44	-1,04	6,16	10,25	42,70	1,0
0,014	0	-9,83	-1,56	1,49	-0,56	-1,52	-1,04	6,16	10,25	42,70	1,0
0,015	0	-9,83	-1,55	1,49	-0,40	-1,52	-1,12	6,16	10,25	42,70	1,0
0,016	0	-9,82	-1,54	1,49	-0,16	-1,60	-1,12	6,16	10,25	42,70	1,0
0,017	0	-9,81	-1,54	1,50	-0,00	-1,60	-1,28	6,16	10,25	42,70	1,0
0,018	0	-9,80	-1,54	1,51	0,24	-1,68	-1,28	6,16	10,25	42,70	1,0
0,019	0	-9,80	-1,53	1,51	0,48	-1,68	-1,36	6,16	10,25	42,70	1,0
0,020	0	-9,79	-1,53	1,51	0,64	-1,68	-1,36	6,16	10,25	42,70	1,0
0,021	0	-9,78	-1,52	1,52	0,88	-1,68	-1,44	6,16	10,25	42,70	1,0
0,022	0	-9,77	-1,52	1,52	1,12	-1,68	-1,52	6,16	10,25	42,70	1,0
0,023	0	-9,76	-1,52	1,53	1,28	-1,68	-1,52	6,16	10,25	42,70	1,0
0,024	0	-9,75	-1,52	1,54	1,52	-1,68	-1,60	6,16	10,25	42,70	1,0
0,025	0	-9,74	-1,52	1,55	1,68	-1,68	-1,68	6,16	10,25	42,70	1,0
0,026	0	-9,74	-1,52	1,56	1,84	-1,76	-1,76	6,16	10,25	42,70	1,0
0,027	0	-9,73	-1,52	1,57	2,00	-1,76	-1,76	6,16	10,25	42,70	1,0
0,028	0	-9,72	-1,52	1,58	2,08	-1,84	-1,84	6,16	10,25	42,70	1,0
0,029	0	-9,71	-1,51	1,59	2,24	-1,84	-1,92	6,16	10,25	42,70	1,0
0,030	0	-9,71	-1,51	1,60	2,32	-1,92	-2,00	6,16	10,25	42,70	1,0
0,031	0	-9,71	-1,52	1,61	2,40	-1,92	-2,00	6,16	10,25	42,70	1,0
0,032	0	-9,70	-1,51	1,62	2,56	-2,00	-2,08	6,16	10,25	42,70	1,0
0,033	0	-9,69	-1,51	1,62	2,64	-2,00	-2,08	6,16	10,25	42,70	1,0
0,034	0	-9,68	-1,51	1,63	2,80	-2,08	-2,08	6,16	10,25	42,70	1,0

Figure C.1: Representation of the raw data acquired by the accelerometer, gyroscope, and GPS before labeling.

1 ACC_X	2 ACC_Y	3 ACC_Z	4 GYRO_X	5 GYRO_Y	6 GYRO_Z	7 ActID
3.1200	-1.5600	1.8500	-7.9200	52.0800	7.5200	Turn Left
11.5900	3.6800	1.2600	-8.3200	56.0800	6.4800	Turn Left
17.6500	11.0300	0.8000	-8.8000	57.6000	5.7600	Turn Left
16.1300	13.8100	1.2800	-9.1200	55.7600	6.2400	Turn Left
8.3600	10.5000	0.2800	-9.3600	53.4400	8	Turn Left
1.8400	7.6400	-0.9800	-9.5200	52.4800	9.6000	Turn Left
3.1800	10.0400	0.1200	-9.5200	55.1200	9.6800	Turn Left
9.3600	13.5400	1.0300	-9.2000	58.6400	8.4800	Turn Left
14.5100	10.7300	1.5000	-8.8000	60.2400	7.4400	Turn Left
15.8800	7.6300	0.7200	-8.4800	59.8400	7.5200	Turn Left
15.3300	10.5600	0.5200	-8.0800	58.8800	7.7600	Turn Left
14.9000	14.9800	2.1300	-7.6800	58.4800	7.5200	Turn Left
13.2200	17.1300	3.3000	-7.2000	58.1600	6.9600	Turn Left
10.6400	16.6300	2.0500	-6.9600	57.3600	7.2000	Turn Left
9.8000	14.0300	0.2200	-6.8000	57.5200	7.8400	Turn Left
9.0600	8.2800	-1.4600	-6.6400	57.9200	8.2400	Turn Left
9.8000	4.5100	-2.1500	-6.4000	57.3600	8	Turn Left
14.0900	6.0400	-0.1400	-6.4800	57.4400	7.1200	Turn Left
17	9.6400	1.2400	-6.8800	56.8000	6.4800	Turn Left
16.6500	11.5700	2.1700	-7.4400	54.9600	6.5600	Turn Left
13.7500	13.4100	1.5400	-7.7600	53.6800	7.2000	Turn Left
8.1100	10.2500	0.3900	-7.8400	52.4800	8	Turn Left
3.8300	-1.4800	1.0800	-7.4400	51.7600	8.1600	Turn Left
6.7300	-5.0700	1.8300	-7.0400	52.4800	7.5200	Drive
12.8100	2.5500	1.3800	-7.3600	54.4000	6.4800	Drive
14.6000	8.6500	-0.6500	-7.7600	54	5.9200	Drive
10.8200	9.2100	-1.0600	-7.8400	52	6.1600	Drive
4.0900	3.9100	-0.3000	-7.3600	50.3200	7.2800	Drive
0.9100	0.3900	0.7600	-6.8800	50.6400	8.4800	Drive
3.8400	3.6700	2.1200	-6.7200	53.2800	8.4800	Drive
9.0900	6.3800	1.7300	-6.8000	56.0800	7.2800	Drive
13.4200	4	1.5200	-6.4800	57.1200	6.1600	Drive
14.7600	2.2300	2.0800	-5.9200	57.2000	6.0800	Drive
12.5100	4.3700	2.2800	-5.5200	56.4000	6.2400	Drive
9.3400	6.9500	3.0100	-5.1200	55.8400	6.0800	Drive

Figure C.2: Representation of the raw data acquired by the accelerometer and gyroscope after labeling.

D | Literature Review

Review of Activity Recognition and the Challenges Related to Off-road Motorcycle Riding

Bas Breider

Abstract—**Activity recognition has expanded its influence by providing information on people's behavior in numerous applications. With activity recognition having advanced substantially, so has the number of challenges in designing, implementing, and evaluating activity recognition systems. This review specifically focuses on the challenges that arise when designing an activity recognition system for motocross applications. In this paper, different approaches to activity recognition are analyzed to examine the current status. First, the various activities and sensors used in activity recognition are discussed with respect to their challenges that arise in motocross. In the next section, the related work is discussed according to the main steps of activity recognition: preprocessing, feature extraction, classification, evaluation. Finally, various trends and ideas are presented to address in future research.**

I. Introduction

Physical activity has been defined as any bodily movement produced by skeletal muscles, which can be categorized into occupational, sports, conditioning, household, or other activities [1]. Automatic recognition of these activities has emerged as a key research area in pattern recognition, especially in the last couple of decades. In general, activity recognition is seen as a broad concept involving the use of technology to automatically recognize different activities and then apply this information for various applications. To increase the lifestyle quality of people, the need for the automated activity classification applications has been identified in various fields, ranging from health care to entertainment industries [2].

A potential application domain for activity analysis is sports. Especially the analysis of movement is considered as a growing research area within the field of sport monitoring. The aim of sports monitoring is to enhance the performance of athletes, prevent injuries, or optimize training programs. Considering the ongoing research on the development of activity recognition, the application areas of movement analysis in sports will also benefit from this trend. The activity recognition used for movement analysis can for example provide valuable information to better understand the physical demands of numerous sports [3], which subsequently can be used to provide feedback to athletes [4] and design training programs to improve performance [5]. Activity recognition is also increasingly used in sports broadcasting allowing analysts to identify interesting events of the game and provide better viewer insights [6]. These various examples indicate the significance of activity recognition in sports.

In addition to sports monitoring and sports entertainment industries, safety related applications could benefit from the development of activity recognition. Health benefits of activity recognition include the reduction of injury risk in sports [7]. One of those high risks sport that could benefit from increased safety is motocross. This extreme sport, mainly popular in America and Europe [8], has one of the highest injury rates among all sports [9]. This work focuses on off road motorcycle riding, in particular motocross.

The automated recognition of activities is a challenging area of work, where the challenges are related to both technical and human factors in terms of design and implementation. In addition to the common issues in activity recognition like variability within and between persons, sport activity recognition has its own challenges [10]. The variety of sport activities can be highly diverse and its recognition therefore requires careful selection of several various sensors that differ in their capabilities and characteristics. Furthermore practical limitations on placement and number of sensors have to be taken into account, together with sensor characteristics as power consumption, privacy, accuracy, and noise [11]. To successfully recognize multiple activities in motocross, an algorithm must overcome these challenging factors.

The aim of this review is to present the current state in human activity recognition making use of the current challenges and their relation to activity recognition in motocross. As the main challenges in activity recognition are briefly introduced before, their elaboration and application to motocross will be discussed in section 2. Now that every activity recognition is a combination of methods to overcome these challenges, the activity recognition process follows a general framework. Solutions for the main steps of the framework are required to get a functional activity recognition method of motocross activities. This subdivision of methods is used to organize section 3 of literature study. Furthermore, this review introduces some of the most relevant open problems in the field providing directions for future research in section 4 and ends with a conclusion in section 5.

II. Challenges

There is no uniform solution to the problem of activity recognition due to the abundance in applications. Most of the challenges that appear when designing an activity recognition system are related to variance since there are countless possibilities depending on the application goal. The wide range of activities, the different ways of data acquisition, the environmental changes, and the surplus of variance in appearance of the user and performance of the action, cause that the activity recognition process has to be adapted to each individual system. To customize the activity recognition process to motocross, some extra challenges arise, predominantly caused by the extreme conditions during motocross races. The notion that action recognition is still an unsolved problem brings up the following question: what are the common challenges in the field of action recognition and how are these challenges related to motocross? This section provides background information on the available sensors and the activities that can be detected, pinpoints its problems, and gives a comparison of the state of the art to tackle the challenges in motocross.

Activity Challenges

1) Specification of activities

Activity recognition is a broad concept within the current research literature. To gain insight in the fundamental processes, the first step is to define what is exactly an 'activity'. In brief, activity can be described as any bodily movement produced by skeletal muscles, which can be categorized into occupational, sports, conditioning, household, or other activities [1]. Therefore the goal of the application determines which activity is useful to recognize.

Daily life activities Current literature shows that the applications of activity recognition in health care are abundant [12]–[14]. In many of the health care applications the user is monitored to recognize daily activities like sitting, standing, jumping, and walking [15]. These activities are characterized by their slow sequence of movements or in fact even a static posture. As motocross is a sport in which high speeds are needed to achieve a good performance, the natures of the activities are mainly dynamic. These slowly executed or even static activities are thus not relevant activities in motocross to be recognized.

Falls among physically active people remain an essential public health issue. So can falls lead to serious damage to the body like contusion, fractures, and concussions [9]. Even when no physical damage is done to the body, the resultant fear of falling may lead to reduced freedom of movement. With regards to health care applications, falling is detected and used to alert caregivers that can then provide help to the injured person. Falling can also be detected before the impact actually occurs to deploy an inflatable jacket and prevent serious injury [16]. This inflatable jacket concept is also tested in motor riding [17]. A fall detection system could potentially reduce the high injury rates in motocross [9] by preventing the impact of the fall. Besides the application of fall recognition for these kinds of safety devices, fall detection could also improve safety in motocross after the fall has already occurred. During motocross races and practices, volunteers stand alongside the track to increase safety for the riders. These volunteers will alarm race control when an accident has happened, provide the rider with medical care, and are held responsible for creating safe racing environment for all riders [18]. Considering all these different tasks, some help in creating a safe situation could benefit the volunteers as well as the riders. With a fall detection system, the task load of the volunteers would be reduced and safety would increase during motocross racing.

Sport activities Besides that daily life events can be related to motocross, there are also some sport specific activity recognition application that could be related to motocross. The sports activity recognition applications range from daily sports activity measurement and performance classification [19] to sports entertainment [6]. Although the applications of activity recognition in sports are large, yet it has to be applied to motocross. Usually thousands of motocross fans watch the races live on television or go to circuit, which indicates the high entertainment value of motocross for their fans. When watching a race, fans can gain a sense of togetherness with riders and other fans. However, they cannot obtain sufficient information to understand the current race situation at the circuit, because it only provides very little information about timing of the riders. Therefore, there is an increasing demand for information systems to provide the information that the fans want to know. To solve this problem, activity recognition is used for the detection of high entertainment valued activities like turning and crashes in motorcycle racing [6]. In the case of recognizing such an event, the information can transferred to the television control room from which these events can be broadcast with relevant added information. Besides the direct link between activity recognition and valuable information, there are more sources that could provide interesting information about the performance of the rider. With the use of Intel's micro sensors [20] the crowd and racer can be provided with valuable information about the performance of the jump in terms of airtime, height, and speed during the jump. The recognition of high entertainment valued activities could also be applied to motocross. So can jumping and turning be seen as key activities for the spectators as well as the riders and the team, which after the recognition can be broadcast and shown with additional performance metrics.

Motor riding For this review about activity recognition there is no directly applied relevant work when it comes to motocross. The closest what has been done compared to activity recognition in motocross is fall detection [21], impact detection [22] and safety critical event recognition of powered two wheelers on the road [23], [24]. In particular, the riders safety in terms of falling has been researched [21]. The falls were specified to falling in a turn caused by a slippery surface. As smooth falling styles are not the only falling patterns, [22] presented a concept for crash detection in general. [23] and [24] took it one step further and identified dangerous riding patterns, like swaying and extreme accelerations and decelerations. However, this is to a limited extend comparable to motocross. So can changes in surface cause the motor and motor rider to perform maneuvers that are normal during motocross racing, but would indicate a hazardous situation on the road [22]. Furthermore, jumping while riding a motorcycle would in most cases mark a dangerous situation, however this is a frequently occurring activity in motocross. Yet the safety reasons and their related activities like for example jumping and falling are relevant for motocross.

2) Composite and Concurrent activities

To discuss qualitative activity recognition, relevant activities have to be separated from the other activities that occur during motocross. The challenge of refining an activity becomes much harder when we take activities into account that are composed of multiple other activities. Composed activities like sewing and drilling are composed of several instances of lifting hands, pushing, and pulling [25]. Consider the problem of automatically recognizing a jump during motocross. With the evolution of the jumping technique for advanced riders, the jump evolved into what is called a 'scrub' [26]. This 'scrub' includes, beside the vertical movement of the rider, a horizontal positioning of the motorcycle while in the air. So this activity is composed by several activities, jumping and falling, which makes the differentiation in recognition process harder.

During the execution of a sequence of movements, it is possible that a certain activity begins when the previous activity is not yet finished: concurrent activities. The assumption that an individual only performs one activity at a time is true for basic activities. However in general, activities are rather overlapping and concurrent. This also introduces the problem of defining which activity is performed when. So can a person for example be falling while running or riding a bike [27]. This challenge is also present in motocross, as falling is an unintentional and unwanted activity that occurs mainly during jumping and turning during the race as is shown in figure 1.

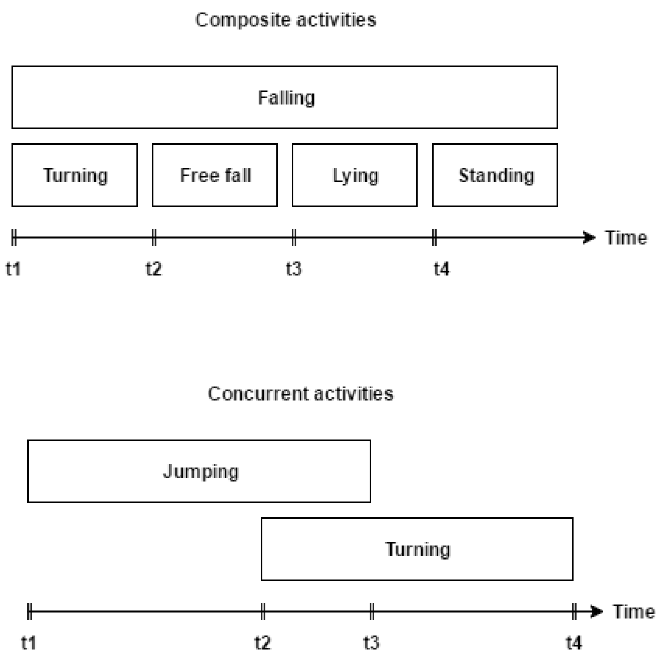

Fig. 1. Representation of composite and concurrent activities in motocross

3) Variance in activities

So far the different activities are discussed, but there are some more challenges that arise when executing these activities. Namely, the activity recognition system is the best when it works for a large number of users instead of tailored to a specific use, because differences between individuals result in variations in the way that people perform tasks. Several factors can affect the performance of the activity, such as stress, fatigue, or the emotional or environmental state in which the activity is performed [28]. As this effect is shown in many different sport activities [29], this could also be related to motocross. Especially when you take in mind the high physical activity demand on the riders. At the beginning of a race, riders are less fatigued and therefore more powerful to execute a jump higher and further than at the end of a tiring race. This means that for classification of the activities, the difference between and within the racers have to be taken into account.

The set of activities that are being recognized, play an important role in the design decisions of an activity recognition system. From what is discussed so far, there are important activities like falling, jumping, turning, and starting/stopping that are useful for activity recognition in motocross for different reasons like safety, entertainment, and performance. In addition to the application, the activity specification can help developers or researchers to make different design choices, such as which sensors to use. This sensor selection and other sensor challenges are discussed next.

Sensor Challenges

4) Sensor selection

A practical challenge for implementing activity recognition in applications is caused by the sensing equipment, more specifically the variability in sensors and their characteristics. The sensors used in activity recognition can be roughly divided into two categories: (1) camera sensor based methods, which can be either fixed or mobile; (2) micro physical sensor based methods, which utilize motion sensors and environmental cues to get information of the user's state.

Camera

Camera sensor based methods make use of a camera to record the activity sequence and recognize the activities using computer vision algorithms. Within the activity recognition field, cameras are used in two different setups: fixed camera in the environment of the user and a wearable camera attached to the user.

Fixed camera Fixed camera setup for activity recognition has been used in health care applications to capture activity or fall in household settings [30]. In an entertainment environment, the activity recognition can improve the human computer interaction, such as the automatic recognition of different players actions during a sports game [31]. Using cameras fixed to the environment may be acceptable when activities are confined to certain parts of a constrained environment, like indoor sports. When activities are performed in a changing environment and involve going from place to place fixed camera systems are not very practical because acquiring video data is difficult for long-term human motion analysis in such unconstrained environments. In motocross, the variety in track compositions does not make a static camera setup suitable for activity recognition.

Wearable camera Recently, wearable camera systems have been proposed to overcome the problems of a fixed camera. From a first person view, a wearable camera mounted on the chest was able to capture which activities were performed with the users hands [32]). Another application of activity recognition with a wearable camera is to recognize activity of others [33]. Besides the continuous capture of relevant activities with a wearable camera, the other disadvantages, such as occlusion effects, the high cost of processing and storing images, the need for using multiple camera projections from 3D to 2D, and cameras privacy issues do still exist.

When it comes to activity recognition in motocross, using a wearable camera would not be sufficient, nevertheless first person viewpoint could be used for entertainment purposes.

Capturing physical activity with any optical system can be challenging, even impractical, within large volumes of data due to the cost and difficulty capturing specific activities with enough cameras to ensure enough coverage. Additional complications can also arise during outdoor data capture sessions where there is little control over lighting conditions or occlusions when motion from high velocity movements is required and tracking can easily be lost. Efforts to recognize activities in unconstrained settings have caused a shift from camera usage toward using inertial sensors.

Microelectromechanical systems

Alternative sensors that can overcome the drawbacks of the camera are microelectromechanical systems (MEMS). MEMS are miniature devices comprising of integrated mechanical and electrical components designed to work together to sense and report on the physical properties of environment. The vast majority of activity recognition systems make use of MEMS such as inertial sensors, accelerometers, and gyroscopes. These sensors are used either alone, in combination with each other, or in combination with other sensors like GPS. In the following part of this review the most used MEMS in activity recognition will be elaborated.

Accelerometer Accelerometers are sensors which measure the accelerations of objects in motion along reference axes. Measuring activity using accelerometers is preferred because acceleration is proportional to external force and thus can reflect intensity and frequency of movement. Research has shown that accelerometers can be used to identify human activity for high energy actions such as walking, running, and jumping [34]–[36]. Measured accelerations can also be due to gravity, and therefore accelerometers can be used to calculate tilt angle. The resulting inclination data can be used to classify body postures. Fall detection makes use of these features to classify if a person is either lying or standing [37]. In sports, accelerometers have been used to monitor elite athletes in competition or training environments. For example in swimming, accelerometers have allowed the comparison of stroke characteristics for a variety of training strokes and therefore have helped to improve swimming technique [38]. With these characteristics, accelerometers are capable of providing sufficient information for measuring a wide range of human activities. Especially the good results in unconstrained and varying environments make the accelerometer suitable to motocross.

Gyroscope Gyroscopes provide angular rate information around an axis of sensitivity. By integrating the gyroscope signal, change in orientation can be obtained. The orientation can then be used to identify postures and change in position. This is done in [39] to identify static and dynamic activities like standing, running, walking upstairs and downstairs. The gyroscope produces better recognition accuracy for activities that are characterized by rotations around the reference axis of the gyroscope, like walking upstairs and downstairs with gyroscope placed on the leg. Activities which are characterized by change in orientation can thus be identified from a gyroscope signal. However, the combination of an accelerometer and gyroscopes performs better for activities that exhibit unique patterns of orientation changes. So did [39] and [40] report that activity recognition improved when the accelerometer signals where complemented with gyroscopic signals. In combination with an accelerometer, the gyroscope performs better than when used individually. This is an important result and supports the idea of using both sensors in combination with each to better recognize activities. To see if this also holds for other activities, [24] did use a gyroscope in combination with an accelerometer to detect dangerous maneuvers during motor riding. By combining the accelerometer and gyroscope data, all changes in the dynamic behavior of the motorcycle can be detected. For the activities that occur during motocross, positioning of the bike and rotational changes are characteristic. A gyroscope would be a good sensor to provide information about the rotations that the rider and the bike experience/endure during a race.

Magnetometer Sensor units that contain accelerometers and gyroscopes can be enhanced by magnetometers. Combining magnetometers with accelerometer and gyroscopes in a single inertial measurement unit (IMU), allows the sensor heading to be determined and can help increase robustness against high acceleration motions [41]. The magnetometer, accelerometer, and gyroscope were evaluated individually and together based on the activity recognition accuracy. Comparing the classification results based on the different sensor combinations indicates that if only a single sensor type is used, the best correct differentiation rates are obtained by magnetometers, followed by accelerometers and gyroscopes. However, combining all three sensors lead to the best recognition accuracy [39]. Especially in motocross, a sport in which the posture is of main importance to be able to perform under hard conditions. Turning is an obvious activity in which the posture of the rider is used to keep balance. So any sensor that could give information about the heading of the rider, could improve the characterisation of movements. It should be kept in mind, however, that magnetometer outputs can be easily distorted by metal surfaces. The deflection in magnetometer readings can lead to misleading information and wrong interpretations. So the usability for activity recognition for motocross is reduced as the motor itself is partially made out of steel components.

GPS The Global Positioning System (GPS) enables all sort of location based services that can complement inertial sensors to recognize activities based on velocity and position. In [42], the authors aimed to recognize several sports activities based on accelerometer and GPS data. The location of the user can be helpful to infer their activity using logical reasoning. As an example, if a rider's location is off the track, there might be a problem with the safety of the rider. In addition, the location of the wearer is also useful in a risk-associated sport like motocross to detect the position of the rider and pass this through to emergency services in case the driver was unobserved by others. This application has already been introduced for motorcycles and other vehicles on the road [22], [43]. GPS would not be helpful for activity recognition in motocross, because of the lack of information from the coordinates on the track. However, GPS could help to increase safety by providing volunteers and emergency services the exact location of the injured rider.

5) Sensor Placement

One of the challenges in activity recognition using sensors is the variability in placement. The placement of sensor systems refers to the locations where the sensors are attached and the positions in which the sensor is placed. Fixed sensor systems have to be placed in a specified position in which every activity can be recorded, while being stationary. This non-intrusive way of motion recording is ideally used in situation in which the user is constantly moving within the range of the sensor, as the inference of activity entirely depends on the interaction of the users with the fixed sensors. In motocross this is not feasible as the area in which the motions are performed are large, this would require many fixed sensors and therefore increase cost of installation and maintenance.

Naturally, wearable sensors give much more variability in placement for motion detection. Most of the measured attributes of wearable sensors are related to the users movement, environmental variables, or physiological stimuli. To measure the changes in these signals, the miniature inertial sensors have to be attached to the human body instead of employing sensor systems fixed to the environment. Problems are caused by the wrong placement or orientation of the sensors and changes in the position of sensors during motion.

For some wearable sensors such as the accelerometer and gyroscope, activity recognition results depend on orientation changes during the motion. When attached to the body, the axis are aligned and calibrated to the user. Consequently, a change of orientation of the inertial sensors influences the robustness of the measurements. So did [44] report that accuracy decreased when sensor displacement was detected. Consistent orientation can be achieved by a fixed position, which is needed to have a practical activity recognition solution.

So with the lack of orientation independence, users are required to place the sensors in a specific orientation to increase robustness. When riding a motorcycle, the sensors can be placed on the frame of the motorcycle instead of the rider. The sensor devices can then be easily attached to or detached from a motorcycle before and after the race. Therefore, motor-placement causes no constraint in body movement and discomfort can be minimized as well [24].

6) Sensor Consumption

Power consumption is the main factor affecting the design of activity recognition system. The size of the sensor network constraints the power consumption that is related to the battery size, furthermore sensor systems should have enough memory space both for computational power and data that needs to be stored. The requirements of the data storage and computational power depends on the response time of the sensors, which are set by the application of the activity recognition system. The sensor system can either provide *online* or *offline* feedback about the activity recognition. Online feedback involves that the information on the performed activities is directly fed back to the user. Offline feedback includes delayed information, because either more time is needed to recognize activities due to high computational demands, or are intended for applications that do not require real-time feedback.

Fixed sensor systems can be designed in a way that the energy providing source can be connected to the sensor system in multiple ways. The sensors are not explicitly limited by a power supply, as a sensory system can be connected to a stationary power source. The extra volume and weight of the battery is needed to provide, for example a stationary camera, with enough power to match the requirements of the high computational cost when high resolution images have to be processed [45]. However, the maintenance cost and deploying difficulties of the power supply make the system hard to use in situations that need constant adaptations to a changing environment like a motocross track.

As for wearable devices, which are powered by small batteries, size constraints battery capacity. The miniature size is essential for wearable systems, leaving little space to accommodate sufficient energy storage. In most cases, existing wearables store the data in local memory storage and do not make the data available in real-time [46]. The data is stored onboard, and thus data has to be first downloaded from the multiple devices and synchronized offline. Naturally, if data cannot be received from wearable devices in real time, applications that use this data cannot be implemented. So offline recognition, unlike online, is not greatly affected by processing and storage issues, because the required computations could be executed in an external server that has large computational power and data storage capabilities. If feedback is provided online, the power consumption of wearable devices increases [47]–[49]. Despite the increased power consumption with direct feedback, inertial measurement units can reduce the size of batteries, enhance sensor lifetime, and enable long-term activity monitoring.

One way to reduce energy consumption is to only use the sensors that are needed. Although some of the sensors perform better in certain situations than others, the best recognition accuracy is achieved when various sensors are included in the design. Consequently, the energy consumption increases when all sensors are turned on. Since all sensors may not be necessary simultaneously, dynamic and adaptive sensor selection can be used to improve battery life. This means that the sensors are turned on and off in real time in an adaptive way for energy-efficient activity recognition. Such an adaptive sensor selection did not result in a lower accuracy, while the sensor consumption was reduced [48], [50]. See table 1 for a summary of the advantages and disadvantages of the various sensor approaches.

III. Activity Recognition

There are many different methods for interpreting activity information from raw sensor data. However, all activity recognition processes can be summarized as a sequence of signal processing, pattern recognition, and machine learning techniques to classify activities based on raw sensor data. These techniques are based on multiple general steps that define an activity recognition process. The incoming sensor data is first processed to optimally characterize meaningful features of the activities. When characteristic features of the

TABLE I
COMPARISON OF SENSOR APPROACHES

Category	Advantages	Disadvantages
Vision based	• Fixed camera at place with high activity • No need for multiple physical sensory devices • Intuitive to operate	• Capture only race track specific parts • High processing time and power consumption • Camera are comparatively expensive • Sensitive to occlusion and light • High amount of computational power needed
Physical sensor based	• Precise information of body movements during racing • Able to capture the movements of the rider continuously • Low power consumption and computational power • Relatively low cost	• Many sensors needed for complete information • Sensitive to noise • Discomfort of wearing sensors

data are determined, these features and corresponding ground truth class labels are used as input to train a classifier model, which is the training phase. In classification phase, new sensor data is fed into the trained classifier to test the performance in new situations. The final performance evaluation phase allows the assessment of the performance of classifier by setting a score for each activity class and map these scores into a single class label (figure 2). In the following section the most widely used algorithms and methods for each of these steps together with the challenges that arise during the process design are presented.

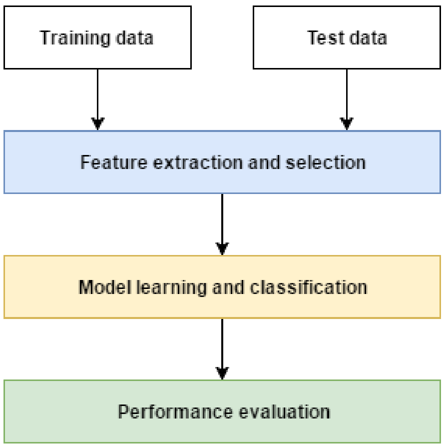

Fig. 2. Representation of data flow for testing and training activity recognition system

A. Feature extraction

In the first phase raw data is acquired using several sensors placed either on the body or in the environment of the user. Before using the streams of data as input to the activity recognition chain, characteristic features have to be determined from the data. In general, these features can be defined as the abstractions of raw data and the purpose of feature extraction is to find the main characteristics of sensor data that precisely represent the original data [51]. Different features have been considered in the literature ranging from raw vision and inertial data signals to high-level descriptors [52]. The feature extraction and selection stage reduces the data into quantitative measures that are discriminative for the different activities. In the following section the most common feature extraction and selection techniques for each of the previously described sensors will be discussed (table 2).

Microelectromechanical systems

The incoming raw inertial sensor data can have unwanted noise caused by a variety of sources like environmental factors, sensor malfunction or activity execution, which interfere with the activity patterns. Therefore the feature extraction phase includes filtering of the raw data. There are different types of filters used for noise removal: low-pass, high-pass, or band-pass filter [21]. Apart from using low-pass filters for the removal of high-frequency noise [24], low-pass filters can also be used to remove the dynamic acceleration, and thus the direction of the gravity vector is found during quasi-static activities. On the other hand, gravitational acceleration can be extracted with a high-pass filter to analyze the dynamic acceleration often characterizing the movement [53].

These filtering techniques have in common that their purpose is to retain useful information in a signal while rejecting unwanted information. When it comes to motocross, the engine of the bike and the bumpy track may cause unwanted vibrations. Therefore the accelerometer and gyroscope data become very noisy and require filtering. Initially, a static frequency analysis can conducted to estimate the spectral signature of the data, which will show the frequency of the engine noise. The engine noise can then be removed by a low-pass or band-pass filter [24]. It is important to note that this preprocessing is supposed to be generic; that is, it should not depend on anything but the data itself. It should not, for example, be specific to any particular person or bike. After the preprocessing stage, characteristics can be extracted from the inertial data. The features are divided here into four categories: heuristic features, time domain features, frequency domain features and timefrequency domain features. The heuristic features are typically related to specific aspects of the movements that occur during the activity and therefore are only useful in certain applications. For motocross this mainly includes the postural changes and execution speed of movements in time that happen when riding a bike on a bumpy track. To extract these features, the designer of the features is required to have expert knowledge about the application so that they know which features will best serve their purpose. For example during a fall, the rider is first subject to a free fall period, which is followed by a rapid deceleration when he hits the ground [22]. In addition there is also a measurable change in orientation during a fall [54]. For every other activity that happens during motocross such features have to be engineered to work well.

In contrast to the heuristic approach, the time, frequency and time-frequency domain features are not typically related to specific aspects of individual movements or postures. Instead they simply represent different ways of characterizing the information within the time varying. The time domain features are the general statistical measurements that can represent the generalization of the data [55]. The frequency domain features analyze the frequency performance of the sensor signals, which is usually the periodicity of the signal over a long duration. Therefore the frequency domain features are good at analyzing stationary signals that contain frequency patterns specific for each activity [56]. The timefrequency domain features refer to features that captures both time and frequency information simultaneously with different timefrequency representations that are useful for non stationary signals in which the frequency component changes in time. They are mainly used to detect transition between different activities [57].

Camera

In vision based activity recognition, a video object is first segmented from its background to extract the required object from a sequence of images. Video object segmentation methods have been categorized into background construction-based methods, and foreground extraction-based methods. In background construction-based methods, the camera is static, hence the background information is obtained in advance and the model is build up for object segmentation. On the contrary, foreground extraction is better used when the activity is captured by the moving camera. Due to the moving object and background,the background model cannot be built in advance. Therefore, the model is obtained in an online setting [58]. After the segmentation, important characteristics of the silhouettes are extracted and presented as a set of features. Ideally, the features should be able to cope with small variations in the images, such as viewpoint, user appearance and action execution. Meanwhile, they should be sufficient enough to support robust recognition of activities.

Features in video based activity recognition can be divided into two categories: Global features and local features: global features follow top-down procedure, while local features proceed bottom-up. Extracting global features needs localisation of the user and specification of a region of interest. The general use of global features is powerful since they encode rich information. However, it requires precise procedures for determining important parts in the image. On the other hand, local features describe the image as a collection of independent blocks of data. They detect the essential spatio-temporal points and calculate the local correlation around these points. At last, the blocks around these points are unified into a final representation. Compared with global features, local features are less sensitive to noises and do not require high quality background subtraction. However, they require an intensive preprocessing step to extract a sufficient amount of relevant points [59].

Feature selection For a given classification problem it is often difficult to identify optimal features. Therefore, methods for selecting optimal features from a larger set and methods for reducing dimensionality of features can be used as preprocessing stage before algorithm learning. The goal of the feature selection is to identify the features with relatively small intra-class and large inter-class variations. Since such features are more discriminative, they result in more accurate classification. Furthermore, the often unnecessarily large number of features can be reduced, which decreases the computation requirements during the activity recognition process [60]. To reduce the feature dimension, multiple approaches can be applied to adapt the feature extraction to the dataset. A well-known method to reduce dimensionality of the existing feature space is principal component analysis (PCA) [61]. The PCA feature extraction method finds the direction of maximal variance in the sensor data and then the data is projected onto those optimal directions. After running the PCA, the most discriminating features are maintained by removal of highly interacting features. As a consequence of the rejection of the features with smaller variance, a reduction of features can be achieved. In addition, neural networks have been proposed to extract more meaningful features, because neural networks can identify local dependencies in sensor data without specific domain knowledge [62]. Another key advantage of a feature extraction done by neural networks is its representation of input features that can create a more generalized learning methods for feature extraction for different activities [63]. This will potentially increase performance in the varying environments and circumstances that are inherent to the nature of motocross.

TABLE II
SUMMARY OF FEATURE EXTRACTION METHODS

Category	**Methods**	**References**
Heuristic domain	Body angle, body velocity, body acceleration	[22], [43], [54], [64]
Time domain	Mean, axis crossings, standard deviation, variance	[13], [19], [34], [55]
Frequency domain	Fast Fourier transform, discrete cosine transform	[15], [48], [56]
Frequency-time domain	Wavelet transform	[57]
Global features	Discrete Fourier transform	[32]
Local features	Scale invariant feature transform	[65]

B. Training & Classification

After feature extraction from the sensor data, the next step in the activity recognition process is to make use of classification algorithms. These algorithms are helpful to describe processes, recognize patterns, and predict sensor data. The degree of complexity of these different classification algorithms varies from simple self-defined threshold-based schemes to more advanced machine learning algorithms like neural networks. The choice for one of these classifiers will be determined by a number of considerations. As well as accuracy, factors like ease of development and speed of real-time execution will influence the final choice. The following paragraphs briefly summarize the different techniques, giving an overview of the potential advantages and disadvantages of each method (table 3).

Threshold-based methods Naturally, the threshold-based classification scheme is the most intuitive classifier known for activity recognition. A predetermined characteristic feature is simply compared to its threshold value to determine whether a particular activity is being performed [66]. For example, threshold-based classification has been successfully applied to the detection of falls when riding a motorcycle. In this case, the extreme postural change that happens when falling with a motorcycle was identified as a characterizing movement. As a result, the rapid deceleration which occurs when the user comes in contact with the ground exceeded the threshold value and therefore was identified as a fall [22].

Threshold-based hierarchical classification schemes The threshold based classifier can be extended by adding multiple features and corresponding thresholds. The range of different characteristics that have been used to develop heuristic features are then used in threshold-based classification schemes. Accordingly, a number of studies have demonstrated improved fall detection accuracy when a number of different threshold rules are combined together [65], [67]. Setting the rules manually is labor-intensive and requires knowledge about the movement of the activity, however the execution of the scheme is executed with minimal computational power [42].

Decision trees The decision tree approach is similar to hierarchical classification as it flows through a decision scheme in which a set of rules is created. However, rather than the decision structure being build up manually by the user, the decision rules are formed by the algorithm. The classifier is therefore fast to develop without considerable user intervention. It is normally executed with minimal computational power and is therefore well suited to real-time applications [68].

K-nearest neighbor Nearest neighbor algorithms are used for classification of activities based on the closest training examples. With a k-nearest neighbor (kNN) classification scheme, each feature makes up a dimension and thus a multi-dimensional feature space is constructed. To classify a new data set, the most common label of the *k* closest training data labels is chosen as activity class [69]. The kNN classifier has proved to identify activities riding a bike and falls during a prolonged period of time [24], [70].

In terms of feasibility, classification algorithm using the kNN method can be developed rapidly, are highly versatile and can be used to classify a large range of different activities. However, for a large dataset each new sample will be used to calculate the distances and therefore it can be computationally expensive. The kNN algorithm requires all the training data to be kept in memory for comparison between each new data point and the entire training datasets. The computation burden of this approach constraints the energy resource capabilities and real time response requirements of the activity recognition system. Hence, there is need to make kNN more suitable for online recognition of activities [71].

Support vector machine The support vector machine (SVM) is a prominent classifier for activity recognition. This machine learning method is based on finding optimal hyperplanes between classes in a way that the margin between patterns of each class is maximized. To do so, an optimization technique is used to find the corresponding separating hyperplanes that perform the required classifications [61]. The SVM method is based on kernel functions that enables classification on higher dimension, but therefore may also be very slow to train with large datasets. When trying to maximize the performance of the SVM method difficulties arise with setting the kernel parameters and type. Because the knowledge is hidden within the model, this may hinder the analysis and incorporation of the algorithm by the creator of the recognition system [72].

Hidden Markov model Markov chains describe system which contain information on the probability of transition between different activity states. In particular, hidden Markov models are a subset of the Markov chain and represent a hierarchical approach for identifying a sequence of activities from a sequence of measured features. At each step, a human is considered to be in one state, which generates an observation. Then a system transits into another state and the transition probabilities between the states are calculated. With this technique, classification of sensor data depends not only on the observed features but also on the likelihood of a transition from a previous activity [73].

Neural network Artificial neural networks are algorithms inspired by biological neural existing in the human brain. It consists of a large number of nodes acting as neurons in a network and the weighted connections between different neurons representing the relationship between its inputs and outputs [74]. In activity recognition, neural networks capture local dependencies of activity signals. In image recognition tasks, the nearby pixels typically have strong relationship with each other. Similarly, with wearable sensors given an activity the nearby inertial sensor readings are likely to be correlated [75]. With a large enough set of training data and parameter tuning, a neural network can provide high classification performance in complex settings [76]. One of the factors that make a neural network well suited for activity recognition is that it preserves feature scale invariances. In image recognition and wearable sensor integration, the recognition system is presented with different scales in images and sensor amplitude readings. Neural networks are capable of classifying activities in these circumstances due to their ability to model local dependencies [75]. However, neural network classifiers have several limitations as well. One of them is the large data set that is often required for training, usually such a dataset is not easily available for wearable based applications. Moreover, trained networks are not interpretable for users as the underlying algorithm weights and links are formed without human interference. And although they have demonstrated high levels of accuracy for a number of classification problems, they can be slow to train and some types of networks difficult to implement [62].

Supervised vs. Unsupervised Two different approaches can be applied to the learning process of classifiers, namely supervised and unsupervised learning. For supervised learning the labelled activity data is required to train the classification algorithm. Once the training phase is complete, the classifier is able to assign an activity label to unknown sensor data. In comparison, unsupervised classification does not require activity labels but directly clusters the sensor data into possible

TABLE III
SUMMARY OF CLASSIFICATION MODELS

Classification model	Advantages	Limitations	References
Threshold-based	• Understandable set of rules • Minimal computational power	• Requires knowledge about movement	[22], [65]–[67]
Decision trees	• Fast to develop set of rules	• Binary classification	[68]
k-Nearest neighbor	• Few parameters tuning • Can be used for real-time applications	• Parameter selection essential	[24], [70]
Support vector machine	• Good performance for multiple activities	• Knowledge hidden in model	[61], [72]
Hidden Markov model	• Likelihood transition used for prediction • Sequence of movement recognition	• Weak performance for time-independent activities	[73]
Neural network	• Can cope with raw data as input • Superior performance in complex settings	• Large dataset required • Chances of overfitting • Slow to train	[62], [75], [76]

activities. The use of an unsupervised approach may be needed in such a context of activity recognition when it is difficult to have labels for the activity data. In general, the user is expected to execute an activity as natural as possible to simulate real life as close as possible. For this reason it is recommended for users to perform activities without the participation of many sensors and researchers interferences. Under these circumstances the chances of missing training data classification labels is increased [77]. In addition, large datasets are often costly to process, because the time needed for annotation of the training data depends on the quality and amount of data. Especially for systems detecting adverse events, like falls, datasets are large and the relevant activities infrequent. And due to the low frequency of these adverse events and the long time needed for annotation, unsupervised learning is a suitable solution for speeding up classifying [78].

Combining different classifiers Combining classifiers is another very promising approach to increase performance of activity recognition systems. The fusion of multiple classifiers leads to the combination of results outputs into a single decision. As these classifiers are all using the same data set to base their decision on, the difference in classifying boundaries is determined by their sensitivity for different patterns. Fusion techniques commonly used in activity recognition research are majority voting, where the majority class is chosen [79], stacking which trains the classifiers to use as a new training stage [80], and boosting which assigns weights to the training patterns to combine weak classifiers [81]. In comparison with single classifiers, majority voting provided better accuracy for daily life activities [55]. Similarly, adaptive boosting improved the performance of other learning algorithms when applied to both artificial and real world datasets [82], [83]. Overall ensemble classifiers can provide complementary decisions and improve the overall accuracy. The increased performance can be ascribed to a few different aspects such as; increased robustness to variability in sensor characteristics and reduced complexity through use of classifiers generated for a specific activity. Despite the increased performance, combining classifiers is clearly more expensive, computationally speaking, as they require several models to be trained and evaluated [84].

C. Evaluation

In order to evaluate an activity recognition system, there is need for performance evaluation. In general, performance evaluation is based on the detection of the activity, which can either be correctly detected, confused or falsely detected by the system. The vast majority of activity recognition systems are using *true positives* and *true negatives* for correct classification. On the other hand wrong classification will lead to *false positives* and *false negatives*. From these classifications multiple metrics can be defined to access the performance of the activity recognition system. First is the accuracy, the most standard metric used for overall performance. The instances that are either correctly detected or rejected are divided by these correct detection and rejection plus the false detection and rejection. Two other measures of relevance are precision and recall. The precision is the number of correctly classified positive instances to the total number of instances classified as positive. Alternatively, the recall is the ratio of correctly classified positive instances to the total number of positive instances. So high precision will lead to more relevant results than irrelevant ones, while high recall implies that the most of the relevant results are classified correctly [85].

Confusion matrix One way to visualize these metrics is to use a so called confusion matrix. The confusion matrix displays the ground truth activities on the rows of the matrix, while the type of activity that is detected is shown on the columns. In this way the matrix shows the number of correct classifications on the diagonal and the number of misclassified activities off diagonal. By comparing the ground truth with the detected activities, the previous described metrics can be determined for overall performance. However, the overall accuracy is not representative of the true performance of a classifier. The number can be strongly biased by dominant classes, usually the less relevant background class. For example a fall is relatively infrequent activity, where jumping and turning are activities that happen more often. However riding between these events are considered to be a large part of the dataset with no relevant information. To address this class skew problem, normalized confusion matrices should be used to allow for objective comparison between the different activity classes [86]. Instead of absolute counts of instances, a normalized confusion matrix shows the confusion as a percentage of the total number of ground truth activity instances. By using the normalized confusion matrix, a clear visual interpretation of the performance of the classifier for each activity can be made.

ROC curve One of the main challenges that arise when evaluating the performance is to set the optimal decision threshold for the classifier. Therefore, a specific approach is to tune the performance for each individual activity and analyze the behavior in so-called Receiver Operating Characteristic (ROC) curve. The ROC curve plots the true positive rate versus the false positive rate [87] and therefore the area underneath the ROC curve represents the performance of the system. A problem that arises with the ROC curve is the sensitivity to class imbalance, which lead to incomplete view as the assumption of balanced classes needed to calculate the ROC is violated.

Note that while designing an activity recognition system, the performance metrics and the corresponding confusion matrix and ROC area are dependent on the overall performance. An application can choose to operate at any point that satisfies the needs in performance. For example, if false positives are more costly than false negatives, an operating point toward the left hand side of the ROC curve would be chosen [88]. However this introduces a new problem that if the characteristics of one activity differ largely from the other, the optimization problem will have to compromise between the accuracy of both activities to get the highest performance measure.

Cost sensitive classification Apart from the optimization challenges introduced by class imbalance effecting the performance metrics, the effect of the performance evaluation on the practical application is of great importance. The optimization objective may be to maximize a single performance metric or several at the same time, however the translation of these metrics into real life situations have to be taken into account. Often it is favorable to reduce false negatives at the price of false positives. This is for example the case in fall detection for elderly [88]. The cost of missing a fall outweighs the cost of indicating a fall when no fall has occurred. The cost of faults can be significant, especially for individuals who are vulnerable, and those that are in need of care. In other cases, a high false positive rate can make people ignore the systems notifications and eventually abandon the system [89].

Time based vs Event based To determine if the activity is correctly classified, the predicted activity has to be compared to the ground truth. However, there are different approaches to evaluate the classification when it comes to evaluation criteria. Classification is typically evaluated with respect to time, which is based on the comparison of each frame of the classified activity and the ground truth. The alternative is to evaluate the classifiers performance in detecting activity based on events rather than time segments. The classified events segments are correct if the activity label has the same label as ground truth for a predefined overlap threshold [80]. In figure 3 the example of the overlap threshold is shown, where the overlap between the ground truth (GT) and classification (C) can be expressed as:

$$O(GT, C) = \frac{GT \cap C}{GT \cup C}$$

Where *O* is the overlap percentage of the activity classification and ground truth compared to the total amount of data of the particular grounth truth and classification.

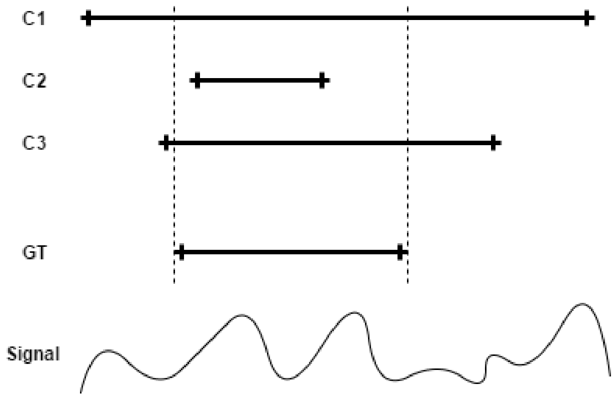

Fig. 3. Representation the ground truth and classification for event based evaluation. The signal is plotted with the ground truth activity (GT) and the three classifications (C).

IV. Discussion & Future work

This review presented aspects of activity recognition and their challenges with respect to the applications in motocross. Despite the extensive literature, it is difficult to compare all of these studies due to their different experimental setups, classifiers, evaluation methods, and their various applications. On the whole, the previously described articles give rise to new challenges, here, an overview of some general trends in activity recognition is given and the aspects that can be improved in future research are highlighted.

Most previously published activity monitoring studies vary considerably in the choice of sensor placements and in the range of activities analyzed, which means that comparisons of the results from different studies should be treated with caution. However, a clear trend show that when tracking activity takes place in an unconstrained environment, the use of wearable sensors is preferred over camera. Although almost all of the relevant studies make use of inertial sensors, there is less consensus when it comes to hazardous activity definition. Falls, near falls, crashes and unexpected driving behavior are all considered dangerous driving but no clear definition for each one is given in the existing literature. Moreover, the detection of dangerous riding patterns or falls are often simulated or executed in controlled environments. For future research it is good to implement a more naturalistic data acquisition process in which all the relevant activities will be defined beforehand.

In the third section a range of different classification techniques has been reviewed. Although a small number of studies do compare the individual performance of different classifiers, there is suggested that either decision trees or artificial neural networks may give the highest classification accuracy in motorcycle hazard detection. Furthermore, there are other methods that showed high potential in detection adverse events such as unsupervised learning, but the application is still in its infancy. In addition, the classifiers are not only chosen based on performance, but on their practical benefits in terms of computational cost and power consumption. This aspect is underexposed in the reviewed literature and therefore further work is required to establish the suitability of the different techniques for specific activity recognition problems this review deals with.

V. Conclusion

In this paper, different approaches to activity recognition are surveyed and geared towards the applications in motocross racing. The information has been organized into three sections. First the background of the activities and the related challenges were discussed, next the different kind of sensors characteristics and challenges were analyzed. In the third section, the related work was discussed according to the fundamentals of activity recognition: preprocessing, feature extraction, classification, evaluation. Finally, various trends and ideas were presented to address in future research due to their high relevance for motocross applications.

By providing background information on the types of sensors used, targeted activities, and the steps of the activity recognition process, the aim was to present a view of the activity recognition field and their specific challenges when applying it to motocross. According to this review, most of the proposals use wearable sensors and machine learning algorithms for the recognition of safety related activities like falling or jumping. However, due to new developments in the area of neural networks and their potential to be applied to activity recognition, their use has to be explored in future research.

References

[1] Carl J Caspersen, Kenneth E Powell, and Gregory M Christenson. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public health reports*, 100(2):126, 1985.

[2] Jake K Aggarwal and Michael S Ryoo. Human activity analysis: A review. *ACM Computing Surveys (CSUR)*, 43(3):16, 2011.

[3] Cassim Ladha, Nils Y Hammerla, Patrick Olivier, and Thomas Plötz. Climbax: skill assessment for climbing enthusiasts. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*, pages 235–244. ACM, 2013.

[4] Eduardo Velloso, Andreas Bulling, Hans Gellersen, Wallace Ugulino, and Hugo Fuks. Qualitative activity recognition of weight lifting exercises. In *Proceedings of the 4th Augmented Human International Conference*, pages 116–123. ACM, 2013.

[5] Marcus J Colby, Brian Dawson, Jarryd Heasman, Brent Rogalski, and Tim J Gabbett. Accelerometer and gps-derived running loads and injury risk in elite australian footballers. *The Journal of Strength & Conditioning Research*, 28(8):2244–2252, 2014.

[6] Tsutomu Terada, Masakazu Miyamae, Yasue Kishino, Takahito Fukuda, and Masahiko Tsukamoto. An event-driven wearable system for supporting pit-crew and audiences on motorbike races. *J. Mobile Multimedia*, 5(2):140–157, 2009.

[7] Cloe Cummins, Rhonda Orr, Helen OConnor, and Cameron West. Global positioning systems (gps) and microtechnology sensors in team sports: a systematic review. *Sports Medicine*, 43(10):1025–1042, 2013.

[8] Alberto Gobbi. Motocross injuries: Incidence, management, and prevention. In *Sports Injuries*, pages 2841–2852. Springer, 2015.

[9] A Noelle Larson and Amy L McIntosh. The epidemiology of injury in atv and motocross sports. In *Epidemiology of Injury in Adventure and Extreme Sports*, volume 58, pages 158–172. Karger Publishers, 2012.

[10] Le Nguyen Ngu Nguyen, Daniel Rodríguez-Martín, Andreu Català, Carlos Pérez-López, Albert Samà, and Andrea Cavallaro. Basketball activity recognition using wearable inertial measurement units. In *Proceedings of the XVI International Conference on Human Computer Interaction*, page 60. ACM, 2015.

[11] Che-Chang Yang and Yeh-Liang Hsu. A review of accelerometry-based wearable motion detectors for physical activity monitoring. *Sensors*, 10(8):7772–7788, 2010.

[12] Ye Liu, Liqiang Nie, Li Liu, and David S Rosenblum. From action to activity: Sensor-based activity recognition. *Neurocomputing*, 181:108–115, 2016.

[13] Liming Chen, Chris D Nugent, and Hui Wang. A knowledge-driven approach to activity recognition in smart homes. *IEEE Transactions on Knowledge and Data Engineering*, 24(6):961–974, 2012.

[14] Saisakul Chernbumroong, Shuang Cang, Anthony Atkins, and Hongnian Yu. Elderly activities recognition and classification for applications in assisted living. *Expert Systems with Applications*, 40(5):1662–1674, 2013.

[15] Mi Zhang and Alexander A Sawchuk. Human daily activity recognition with sparse representation using wearable sensors. *IEEE journal of Biomedical and Health Informatics*, 17(3):553–560, 2013.

[16] Mian Yao, Qi Zhang, Menghua Li, Huiqi Li, Yunkun Ning, Gaosheng Xie, Guoru Zhao, Yingnan Ma, Xing Gao, and Zongzhen Jin. A wearable pre-impact fall early warning and protection system based on mems inertial sensor and gprs communication. In *2015 IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, pages 1–6. IEEE, 2015.

[17] Robert F Buckman, Jay A Lenker, and Donald J Kolehmainen. Air bag inflation device, March 28 2006. US Patent 7,017,195.

[18] Knmv vrijwilligers.

[19] Paul Anthony Yuen. Swim monitor using a three-axis accelerometer, March 25 2014. US Patent 8,681,006.

[20] Intel insports.

[21] Abderrahmane Boubezoul, Stéphane Espié, Bruno Larnaudie, and Samir Bouaziz. A simple fall detection algorithm for powered two wheelers. *Control Engineering Practice*, 21(3):286–297, 2013.

[22] Jussi Parviainen, Jussi Colliri, Timo Pihlström, Jarmo Takala, Kari Hanski, and Aki Lumiaho. Automatic crash detection for motor cycles. In *IECON 2014-40th Annual Conference of the IEEE Industrial Electronics Society*, pages 3409–3413. IEEE, 2014.

[23] Donald Selmanaj, Matteo Corno, and Sergio M Savaresi. Hazard detection for motorcycles via accelerometers: A self-organizing map approach. 2016.

[24] Ferhat Attal, Abderrahmane Boubezoul, Allou Samé, and Latifa Oukhellou. Powered-two-wheeler safety critical events recognition using a mixture model with quadratic logistic proportions. In *ESANN 2015-23rd European Symposium on Artificial Neural Networks*, pages pp–421, 2015.

[25] Stefan Dernbach, Barnan Das, Narayanan C Krishnan, Brian L Thomas, and Diane J Cook. Simple and complex activity recognition through smart phones. In *Intelligent Environments (IE), 2012 8th International Conference on*, pages 214–221. IEEE, 2012.

[26] Jeffrey R Giles and Charles D Ross. The physics of motocross. *The Physics Teacher*, 34:220, 1996.

[27] Rim Helaoui, Mathias Niepert, and Heiner Stuckenschmidt. Recognizing interleaved and concurrent activities: A statistical-relational approach. In *Pervasive Computing and Communications (PerCom), 2011 IEEE International Conference on*, pages 1–9. IEEE, 2011.

[28] J Mark Davis and Stephen P Bailey. Possible mechanisms of central nervous system fatigue during exercise. *Medicine and science in sports and exercise*, 29(1):45–57, 1997.

[29] Magni Mohr, Peter Krustrup, and Jens Bangsbo. Match performance of high-standard soccer players with special reference to development of fatigue. *Journal of sports sciences*, 21(7):519–528, 2003.

[30] Giovanni Diraco, Alessandro Leone, and Pietro Siciliano. An active vision system for fall detection and posture recognition in elderly healthcare. In *2010 Design, Automation & Test in Europe Conference & Exhibition (DATE 2010)*, pages 1536–1541. IEEE, 2010.

[31] Marc Schlipsing, Jan Salmen, Marc Tschentscher, and Christian Igel. Adaptive pattern recognition in real-time video-based soccer analysis. *Journal of Real-Time Image Processing*, pages 1–17, 2014.

[32] Hamed Pirsiavash and Deva Ramanan. Detecting activities of daily living in first-person camera views. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pages 2847–2854. IEEE, 2012.

[33] Michael S Ryoo and Larry Matthies. First-person activity recognition: What are they doing to me? In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2730–2737, 2013.

[34] Lei Gao, AK Bourke, and John Nelson. Evaluation of accelerometer based multi-sensor versus single-sensor activity recognition systems. *Medical engineering & physics*, 36(6):779–785, 2014.

[35] Andrea Mannini, Stephen S Intille, Mary Rosenberger, Angelo M Sabatini, and William Haskell. Activity recognition using a single accelerometer placed at the wrist or ankle. *Medicine and science in sports and exercise*, 45(11):2193, 2013.

[36] Edmond Mitchell, David Monaghan, and Noel E O'Connor. Classification of sporting activities using smartphone accelerometers. *Sensors*, 13(4):5317–5337, 2013.

[37] Daniel Rodriguez-Martin, Albert Samà, Carlos Perez-Lopez, Andreu Català, Joan Cabestany, and Alejandro Rodriguez-Molinero. Svm-based posture identification with a single waist-located triaxial accelerometer. *Expert systems with applications*, 40(18):7203–7211, 2013.

[38] SE Slawson, LM Justham, AA West, PP Conway, MP Caine, and R Harrison. Accelerometer profile recognition of swimming strokes (p17). In *The Engineering of Sport 7*, pages 81–87. Springer, 2009.

[39] Muhammad Shoaib, Hans Scholten, and Paul JM Havinga. Towards physical activity recognition using smartphone sensors. In *Ubiquitous Intelligence and Computing, 2013 IEEE 10th International Conference on and 10th International Conference on Autonomic and Trusted Computing (UIC/ATC)*, pages 80–87. IEEE, 2013.

[40] Wanmin Wu, Sanjoy Dasgupta, Ernesto E Ramirez, Carlyn Peterson, and Gregory J Norman. Classification accuracies of physical activities using smartphone motion sensors. *Journal of medical Internet research*, 14(5):e130, 2012.

[41] Huajun Liu, Xiaolin Wei, Jinxiang Chai, Inwoo Ha, and Taehyun Rhee. Realtime human motion control with a small number of inertial sensors. In *Symposium on Interactive 3D Graphics and Games*, pages 133–140. ACM, 2011.

[42] Miikka Ermes, Juha Pärkkä, Jani Mäntyjärvi, and Ilkka Korhonen. Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions. *IEEE Transactions on Information Technology in Biomedicine*, 12(1):20–26, 2008.

[43] Chris Thompson, Jules White, Brian Dougherty, Adam Albright, and Douglas C Schmidt. Using smartphones to detect car accidents and provide situational awareness to emergency responders. In *International Conference on Mobile Wireless Middleware, Operating Systems, and Applications*, pages 29–42. Springer, 2010.

[44] Kai Kunze and Paul Lukowicz. Dealing with sensor displacement in motion-based onbody activity recognition systems. In *Proceedings of the 10th international conference on Ubiquitous computing*, pages 20–29. ACM, 2008.

[45] Philip Kelly, C Ó Conaire, Noel E O'Connor, and Jessica Hodgins. Motion synthesis for sports using unobtrusive lightweight body-worn and environment sensing. In *Computer Graphics Forum*, volume 32, pages 48–60. Wiley Online Library, 2013.

[46] Cliodhna Ní Scanaill, Sheila Carew, Pierre Barralon, Norbert Noury, Declan Lyons, and Gerard M Lyons. A review of approaches to mobility telemonitoring of the elderly in their living environment. *Annals of biomedical engineering*, 34(4):547–563, 2006.

[47] Sasank Reddy, Min Mun, Jeff Burke, Deborah Estrin, Mark Hansen, and Mani Srivastava. Using mobile phones to determine transportation modes. *ACM Transactions on Sensor Networks (TOSN)*, 6(2):13, 2010.

[48] Yi Wang, Jialiu Lin, Murali Annavaram, Quinn A Jacobson, Jason Hong, Bhaskar Krishnamachari, and Norman Sadeh. A framework of energy efficient mobile sensing for automatic user state recognition. In *Proceedings of the 7th international conference on Mobile systems, applications, and services*, pages 179–192. ACM, 2009.

[49] Samuli Hemminki, Petteri Nurmi, and Sasu Tarkoma. Accelerometer-based transportation mode detection on smartphones. In *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems*, page 13. ACM, 2013.

[50] Piero Zappi, Clemens Lombriser, Thomas Stiefmeier, Elisabetta Farella, Daniel Roggen, Luca Benini, and Gerhard Tröster. Activity recognition from on-body sensors: accuracy-power trade-off by dynamic sensor selection. In *Wireless sensor networks*, pages 17–33. Springer, 2008.

[51] Narayanan C Krishnan, Colin Juillard, Dirk Colbry, and Sethuraman Panchanathan. Recognition of hand movements using wearable accelerometers. *Journal of Ambient Intelligence and Smart Environments*, 1(2):143–155, 2009.

[52] Ivan Laptev, Marcin Marszalek, Cordelia Schmid, and Benjamin Rozenfeld. Learning realistic human actions from movies. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. IEEE, 2008.

[53] Dean M Karantonis, Michael R Narayanan, Merryn Mathie, Nigel H Lovell, and Branko G Celler. Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring. *IEEE transactions on information technology in biomedicine*, 10(1):156–167, 2006.

[54] Angelo Maria Sabatini, Gabriele Ligorio, Andrea Mannini, Vincenzo Genovese, and Laura Pinna. Prior-to-and post-impact fall detection using inertial and barometric altimeter measurements. *IEEE transactions on neural systems and rehabilitation engineering*, 24(7):774–783, 2016.

[55] Nishkam Ravi, Nikhil Dandekar, Preetham Mysore, and Michael L Littman. Activity recognition from accelerometer data. In *Aaai*, volume 5, pages 1541–1546, 2005.

[56] Ghazaleh Panahandeh, Nasser Mohammadiha, Arne Leijon, and Peter Händel. Continuous hidden markov model for pedestrian activity classification and gait analysis. *IEEE Transactions on Instrumentation and Measurement*, 62(5):1073–1083, 2013.

[57] Stephen J Preece, John Yannis Goulermas, Laurence PJ Kenney, and David Howard. A comparison of feature extraction methods for the classification of dynamic activities from accelerometer data. *IEEE Transactions on Biomedical Engineering*, 56(3):871–879, 2009.

[58] Allah Bux, Plamen Angelov, and Zulfiqar Habib. Vision based human activity recognition: a review. In *Advances in Computational Intelligence Systems*, pages 341–371. Springer, 2017.

[59] Donghai Guan, Tinghuai Ma, Weiwei Yuan, Young-Koo Lee, and AM Jehad Sarkar. Review of sensor-based activity recognition systems. *IETE Technical Review*, 28(5):418–433, 2011.

[60] Piyush Gupta and Tim Dallas. Feature selection and activity recognition system using a single triaxial accelerometer. *IEEE Transactions on Biomedical Engineering*, 61(6):1780–1786, 2014.

[61] Zhenyu He and Lianwen Jin. Activity recognition from acceleration data based on discrete consine transform and svm. In *Systems, Man and Cybernetics, 2009. SMC 2009. IEEE International Conference on*, pages 5041–5044. IEEE, 2009.

[62] Thomas Plötz, Nils Y Hammerla, and Patrick Olivier. Feature learning for activity recognition in ubiquitous computing. In *IJCAI Proceedings-International Joint Conference on Artificial Intelligence*, volume 22, page 1729, 2011.

[63] Ali Sharif Razavian, Hossein Azizpour, Josephine Sullivan, and Stefan Carlsson. Cnn features off-the-shelf: an astounding baseline for recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 806–813, 2014.

[64] Michael C Stevens, Wei Lu, Changhong Wang, Stephen J Redmond, and Nigel H Lovell. A kalman filter to estimate altitude change during a fall. In *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the*, pages 5889–5892. IEEE, 2016.

[65] Zhen-Peng Bian, Junhui Hou, Lap-Pui Chau, and Nadia Magnenat-Thalmann. Fall detection based on body part tracking using a depth camera. *IEEE journal of biomedical and health informatics*, 19(2):430–439, 2015.

[66] Georgios Mastorakis and Dimitrios Makris. Fall detection system using kinects infrared sensor. *Journal of Real-Time Image Processing*, 9(4):635–646, 2014.

[67] Lina Tong, Quanjun Song, Yunjian Ge, and Ming Liu. Hmm-based human fall detection and prediction method using tri-axial accelerometer. *IEEE Sensors Journal*, 13(5):1849–1856, 2013.

[68] Uwe Maurer, Asim Smailagic, Daniel P Siewiorek, and Michael Deisher. Activity recognition and monitoring using multiple sensors on different body positions. In *Wearable and Implantable Body Sensor Networks, 2006. BSN 2006. International Workshop on*, pages 4–pp. IEEE, 2006.

[69] Thomas Cover and Peter Hart. Nearest neighbor pattern classification. *IEEE transactions on information theory*, 13(1):21–27, 1967.

[70] Chien-Liang Liu, Chia-Hoang Lee, and Ping-Min Lin. A fall detection system using k-nearest neighbor classifier. *Expert systems with applications*, 37(10):7174–7181, 2010.

[71] Sulaimon A Bashir, Daniel C Doolan, and Andrei Petrovski. Clustering and nearest neighbour based classification approach for mobile activity recognition. 2016.

[72] Xing Su, Hanghang Tong, and Ping Ji. Activity recognition with smartphone sensors. *Tsinghua Science and Technology*, 19(3):235–249, 2014.

[73] Young-Seol Lee and Sung-Bae Cho. Activity recognition using hierarchical hidden markov models on a smartphone with 3d accelerometer. In *International Conference on Hybrid Artificial Intelligence Systems*, pages 460–467. Springer, 2011.

[74] Anil K Jain, Jianchang Mao, and K Moidin Mohiuddin. Artificial neural networks: A tutorial. *Computer*, 29(3):31–44, 1996.

[75] Ming Zeng, Le T Nguyen, Bo Yu, Ole J Mengshoel, Jiang Zhu, Pang Wu, and Joy Zhang. Convolutional neural networks for human activity recognition using mobile sensors. In *Mobile Computing, Applications and Services (MobiCASE), 2014 6th International Conference on*, pages 197–205. IEEE, 2014.

[76] Zhiwei Deng, Arash Vahdat, Hexiang Hu, and Greg Mori. Structure inference machines: Recurrent neural networks for analyzing relations in group activity recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4772–4781, 2016.

[77] Dorra Trabelsi, Samer Mohammed, Faicel Chamroukhi, Latifa Oukhellou, and Yacine Amirat. An unsupervised approach for automatic activity recognition based on hidden markov model regression. *IEEE Transactions on Automation Science and Engineering*, 10(3):829–835, 2013.

[78] Miao Yu, Yuanzhang Yu, Adel Rhuma, Syed Mohsen Raza Naqvi, Liang Wang, and Jonathon A Chambers. An online one class support vector machine-based person-specific fall detection system for monitoring an elderly individual in a room environment. *IEEE journal of biomedical and health informatics*, 17(6):1002–1014, 2013.

[79] Maja Stikic, Tâm Huynh, Kristof Van Laerhoven, and Bernt Schiele. Adl recognition based on the combination of rfid and accelerometer sensing. In *Pervasive Computing Technologies for Healthcare, 2008. PervasiveHealth 2008. Second International Conference on*, pages 258–263. IEEE, 2008.

[80] Jamie A Ward, Paul Lukowicz, and Gerhard Tröster. Evaluating performance in continuous context recognition using event-driven error characterisation. In *International Symposium on Location-and Context-Awareness*, pages 239–255. Springer, 2006.

[81] Xiaofei Ji, Lu Zhou, and Qianqian Wu. A novel action recognition method based on improved spatio-temporal features and adaboost-svm classifiers. *International Journal of Hybrid Information Technology*, 8(5):165–176, 2015.

[82] Eric Bauer and Ron Kohavi. An empirical comparison of voting classification algorithms: Bagging, boosting, and variants. *Machine learning*, 36(1):2, 1998.

[83] Ji Zhu, Hui Zou, Saharon Rosset, and Trevor Hastie. Multi-class adaboost. *Statistics and its Interface*, 2(3):349–360, 2009.

[84] Yago Saez, Alejandro Baldominos, and Pedro Isasi. A comparison study of classifier algorithms for cross-person physical activity recognition. *Sensors*, 17(1):66, 2016.

[85] Jamie A Ward, Paul Lukowicz, and Hans W Gellersen. Performance metrics for activity recognition. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2(1):6, 2011.

[86] H Wallach. Evaluation metrics for hard classifiers. *Cavendish Lab., Univ. Cambridge, Cambridge, online available at http://www. inference. phy. cam. ac. uk/hmw26/, Tech. Rep. Technical Report*, 2006.

[87] James A Hanley and Barbara J McNeil. The meaning and use of the area under a receiver operating characteristic (roc) curve. *Radiology*, 143(1):29–36, 1982.

[88] Bashar Altakouri, Gerd Kortuem, Agnes Grünerbl, Kai Kunze, and Paul Lukowicz. The benefit of activity recognition for mobile phone based nursing documentation: A wizard-of-oz study. In *Wearable Computers (ISWC), 2010 International Symposium on*, pages 1–4. IEEE, 2010.

[89] Qiang Li, John A Stankovic, Mark A Hanson, Adam T Barth, John Lach, and Gang Zhou. Accurate, fast fall detection using gyroscopes and accelerometer-derived posture information. In *Wearable and Implantable Body Sensor Networks, 2009. BSN 2009. Sixth International Workshop on*, pages 138–143. IEEE, 2009.