Lukas Cramer *, Sinan Yavuz, Nana Schlage, Andreas Mühlen, Andreas Kitzig, Edwin Naroska and Gudrun Stockmanns

KneTex – Improvements to classification methods for a sensor system for rehabilitation after ACL surgery

https://doi.org/10.1515/cdbme-2022-1101

Abstract: Injuries and the associated surgery to the anterior cruciate ligament (ACL) can often trigger unpredictable effects, such as the so-called giving way effect, which is an uncontrolled buckling of the knee joint. For this purpose, the KneTex project has developed a smart textile-integrated sensor and actuator bandage system to record the movement of such patients and to monitor and support the rehabilitation process. Long-term monitoring and analysis of the movement data will identify patterns or gait types that can lead to a giving way effect. This paper describes the recent developments of the random forest model-based motion classification system developed within the project. Improvements have been achieved by reducing the number of features needed by 25% using feature importance analysis, speeding up the computation time by 14%, and increasing the classification efficiency. Feature elimination is a useful tool to improve classification systems in settings where feature count is high and feature importance analysis contributed by improving our understanding which sensor of our system are important for the motion classification task.

Keywords: KneTex, IMUs, motion classification, recognition system, feature importance, random forest

1 Introduction

In contrast to rehabilitation systems, which can be used in fixed locations due to the structure and architecture [1][2], the KneTex system, a knee bandage functionalized with sensors and actuators, can be used in everyday life due to the mobile and textile-integrated structure. After ACL surgery patients may experience unpredictable effects in their gait pattern, such as the so-called giving way phenomenon [3] which is difficult to measure due to its uncontrolled occurrence. These and similar events can be analysed using long-term monitoring assisted by the data from the sensors in the bandage. In this way, therapists in rehabilitation are provided with a measuring tool which augments the subjective perceptions of patients with objective observations and data, enabling planning and implementation of more individualized and adapted rehabilitation. The classification system contributes to the further understanding during which movements potential giving-way phenomena occur. Thereby it becomes possible to gain a deeper insight into potential triggers of giving way events, while keeping the ease of use of the bandage high, due to the automatic nature of the motion classification system.

2 System overview

The basic KneTex system, first published in [4][5], is made up of three units. The first unit is a bandage in combination with a mobile application specifically developed for this project. The bandage is equipped with various sensors to record movements and vibration actuators (details in [5]) to warn the user e.g., in case of incorrect movements. The used sensors include two 9-DoF IMU sensors and a barometer, which are needed for the classification system focused here. The KneTex system is used in pairs to analyse both halves of the body in comparison. A set contains one left and one right bandage and is directly connected to the KneTex smartphone application to control the bandages and to provide a real-time mode in which the classification system is integrated.

The application is also able to forward the encrypted data of the bandage via WLAN to unit two, the KneTex server. Once at the server, the data is processed and stored in a database. The last element, a "standard" IT system, is used to view and analyse the stored data by authorized users via the KneTex website [6].

***Corresponding author: Lukas Cramer:** Niederrhein University of Applied Sciences, Reinarzstr. 49, Krefeld, Germany, e-mail: lukas.cramer@hs-niederrhein.de
S. Yavuz, N. Schlage, A. Mühlen, A. Kitzig, E. Naroska, G. Stockmanns: Niederrhein University of Applied Sciences, Krefeld, Germany

3 Data acquisition

As described in section 2, the KneTex smartphone application is part of the first unit and is required for data acquisition. The recordings can be used to expand the KneTex movement database or for further analyses such as training the classification system (details [4][5]). In the latest version, the graphical design has been revised and new functionalities have been added. The user is now able to choose two recording

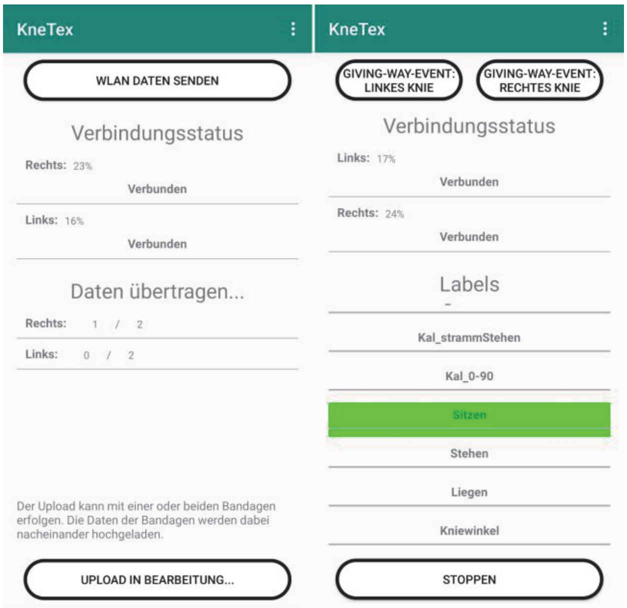

Figure 1: New UI design

modes. Mode one is the already presented offline mode [4]. It now runs completely offline without the need for prior login, allowing it to be used regardless of location. In addition, the recorded data can be uploaded to the server at any time later. For better usability the label list is constantly visible, and the selected activity is highlighted in green (see Figure 1 right). For the upload, the bandages are now connected via a local WLAN network, which allows a faster transfer (see Figure 1 left). Mode two is the newly designed real-time mode. In this mode, the bandages are connected via WLAN to the mobile application, where they forward their data live. The integrated classification system processes the incoming data directly on the smartphone, allowing the user to follow the recognized activity performed. In the background, the data is additionally sent to the KneTex server for further use.

For the data acquisition a standardized recording scheme is used, which defines the exact procedure of how to perform the recordings. This and the use of calibration poses before recordings, ensures that all the recordings meet the same standard.

4 Classification system

The classification system is an important part of the overall KneTex system. It can be used to understand the extent to which giving ways occur depending on the movement context. Using more data from real patients, we expect to be able to link specific types of movement to giving way events. Therefore, the classification is to be seen as a kind of preliminary stage, both for warning (warning speed, frequency of warning, type of warning, etc.) and retrospectively for the consideration by the health care provider. As described in the previous paper [5], the recognition and classification system is

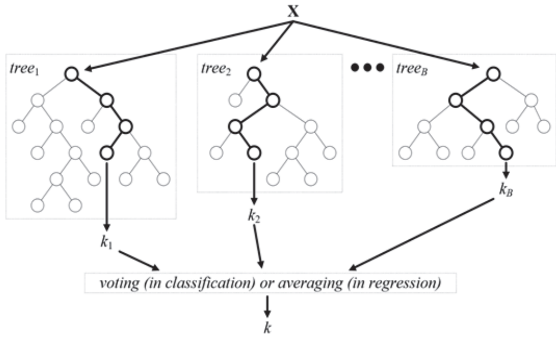

Figure 2: Voting in random forests [7]

based on a random forest model and is trained and evaluated with labeled data acquired by the KneTex bandage. The data includes sensor data (acceleration, pressure, temperature) as well as the label of the movement.

In the following sections, the improvements of the classification system are presented.

4.1 Training

For the classification it is necessary to train the random forest model. During training, many decision trees of the random forest model are built using the data acquired with the KneTex bandage. The system extracts 570 different features, which are used as input data x (see Figure 2). The output k of the classification system is the combination of results from the individual outputs k_B of the decision trees. After training the classifier on a PC, the model can be saved and used for processing on a smartphone. In the evaluation of the system (section 4.3), a total of 1000 trees were trained with a sliding window size of one second and a stride value of 0.5.

4.2 Evaluation as binary classifier

The first step before improving classification performance itself, was to improve our methods of performance evaluation, to get more accurate results and decide whether new changes made an impact. In our previous work, mainly accuracy was used as a metric to evaluate the performance of the classification system. The new evaluation method used, allows the consideration of our multi-class classification problem as a set of multiple binary classification problems, once for each class [8]. The new metrics include accuracy, specificity, precision, recall and F1score. The use of these metrics enables a comparability with other systems. Particularly useful is the F1Score, which combines multiple performance metrics into one. In section 4.4 we use it as an improved metric over accuracy to measure the classification performance.

4.3 Evaluation of feature importance

At present, 570 different features are used. In order to reduce the number of features, the important features must be determined. The high number of features is due to the fact that the sensors provide 19 different values for a singular measurement. One pressure value, and each of the two IMUs

Table 1: List of all used statistical features

mean	std - standard deviation	zqr - zero-crossing rate
maximum	mad – median absolute deviation	mcr - mean-crossing rate
minimum	sma - signal magnitude area	covariance
range	iqr – interquartile range	skewness
median	nrop – number of peaks	kurtosis

providing three axes of gyroscope, magnetometer, and acceleration data, totalling nine per IMU. Furthermore, for each of these 19 values, deltas are computed resulting in 36 values per measurement. With this technique, which is commonly used in speech recognition tasks, we aim to reduce constant offsets in the data introduced by e.g. gravity in the case of acceleration data. Using one-second-long sliding windows 15 statistical features (see Table 1) are extracted for each of the original 36 values resulting in 570 features in total.

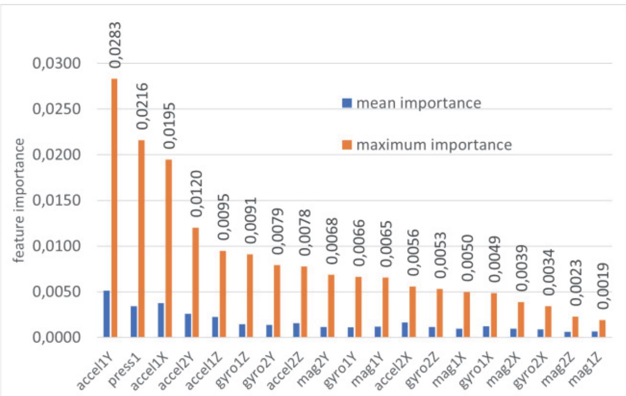

Figure 3: Maximum and mean importance per sensor-axis

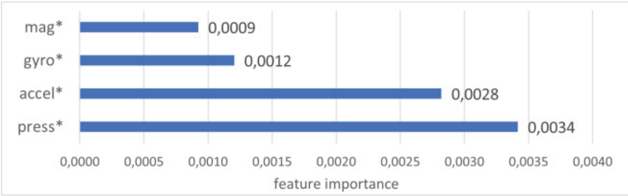

Figure 4: Mean importance per sensor when combining axes

To evaluate the feature importance, the average reduction in gini impurity for each feature is calculated, also called mean decrease in impurity (MDI)[9]. It was chosen for its simplicity resulting in almost no additional computational cost, because this metric is used by random forests to decide how to split and create the decision trees. Therefore, it is already being calculated in most random forest implementations and it only needs to be recorded. The feature, which has the largest average reduction is the most important feature according to this evaluation method. These values were collected during training of the random forest. For each split in each tree, the best improvement in gini impurity is recorded and accumulated separately for each feature.

The importance values for the 570 features were categorised by sensor and mean and maximum values were computed (see Figure 3). The three most important sensors by a large margin are y-axis acceleration of the upper thigh sensor (accel1Y), pressure sensor (press1) and x-axis acceleration of the upper thigh sensor (accel1X) (see Figure 3).

Combining all importance measures from the upper thigh sensor and comparing it to those of the lower shank sensor, we see a ~54% higher average importance of the former. This shows that, while the upper sensor is better suited for this movement classification task, the lower sensor still contributes valuable data. Overall, the magnetometer seems to be the least important sensor, with only one axis (mag2Y) being among the top ten (see Figure 3) and having an average importance of 0.0009 when combining all axes of upper and lower sensors, while the pressure and acceleration sensors have the highest importance by a large margin (see Figure 4). Looking at the

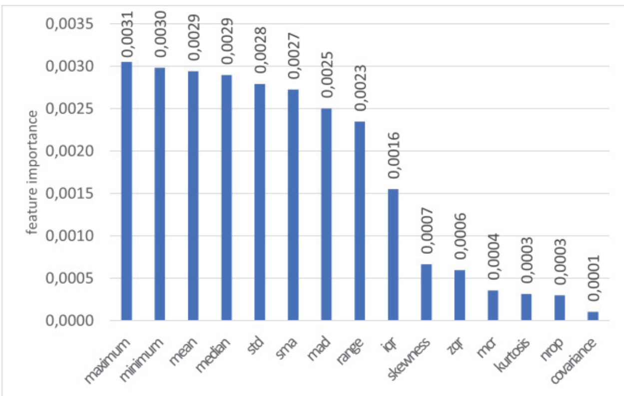

Figure 5: Mean importance per statistical feature

average importance of each used statistic from Table 1, a clear divide between the better and worse features is visible (see Figure 5). Even considering the maximum instead of the average for the four worst features (mcr, kurtosis, nrop, covariance), they don't surpass the average value of iqr (0.0016), which indicates their unsuitability for the task at hand. On average delta features performed worse than the original ones with an importance measure of 0.009 for the former and 0.0026 for the latter.

4.4 Elimination of features

To reduce the large number of features and the accompanying high computation cost, features were selected to not be used in the classification task. Based on the work in section 4.3, the least important features were eliminated. To compare the effect of the elimination, different thresholds were used

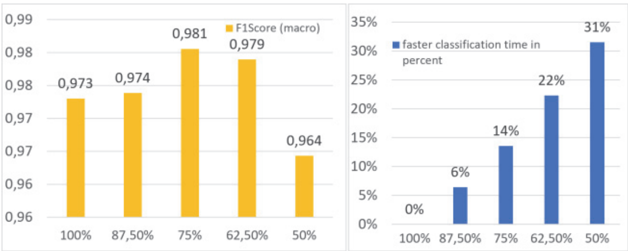

Figure 6: Left: improvement in classification (F1Score); right: improvement in computation time for classification

keeping respectively the best 87.5%, 75%, 62.5% and best 50% of features and discarding the rest. The four worst features shown in Figure 5 were not removed manually, and instead this dynamic approach was chosen to include potentially good sensor / statistical feature combinations.

Predictably, the more features were excluded, the faster the computation time for the classification task (see Figure 6 right). Additionally, it also improved the accuracy of the classification system (see Figure 6 left). Best results were achieved when removing only the bottom 25% of features and keeping the top 75%. Using this threshold an increase of 0.8 percentage points F1Score were obtained. Reduction of feature count by 50% led to worse scores than even the original classification results without removing features.

5 Conclusion and Outlook

In this paper an overview over current developments of the KneTex project was presented. These include the new UI design of the mobile app as well as improvements in the classification system. Using feature importance, less useful features were excluded, speeding up computation time and increasing classification performance. Best results were achieved, when excluding the worst 25% of features, resulting in a 14% speedup and an increase in F1Score of 0.8 percentage points. The speedup allowed implementing the system on mobile phones providing a real time feedback to the user.

Further steps might include replacing MDI with a better, less biased importance measure [9] as well as using the data gathered by the calibration poses to reduce errors introduced by misaligned sensors. In the future we plan to link specific types of movement to giving way events, using data collected through the bandage by automatically classifying movement types using the system described here, which may help the health care provider in further diagnosis.

Author Statement

This project was funded by the European Regional Development Fund (EFRE-0801285). Conflict of interest: Authors state no conflict of interest. Informed consent: All persons included in this project have consented to informed consent. Ethical approval: The actuator study was approved by the Ethics Council of the German Sport University Cologne (No. 116/2019).

References

[1] M. Ayoade, L. Baillie, „A Novel Knee Rehabilitation System for the Home“, CHI Conference on Human Factors in Computing Systems, Toronto Ontario Canada, 2014

[2] P.E. Taylor, G.J. Almeida, et al., „Classifying Human Motion Quality for Knee Osteoarthritis Using Accelerometers“, 32nd Annual International Conference of the IEEE, Buenos Aires Argentina, 2010

[3] A. Lynch, D. Logerstedt, et al., Consensus criteria for defining 'successful outcome' after acl injury and reconstruction: a delaware-oslo acl cohort investigation. British Journal of Sports Medicine, 49(5):335–342, 2015

[4] J. Demmer, A. Kitzig, et al., „Knetex - Development of a textile-integrated sensor system of feedback-supported rehabilitation after surgery of the anterior cruciate ligament“, in *Current Directions in Biomedical Engineering*, De Gruyter, 2020

[5] S. Yavuz, N. Schlage, et al., „Knetex - Improvements to a textile-integrated sensor system for feedback-supported rehabilitation after ACL surgery“, in *Current Directions in Biomedical Engineering*, De Gruyter, 2021

[6] Knetex [Online]. Available: https://knetex.hsnr.de

[7] A. Verikas, E. Vaiciukynas, et al., Electromyographic patterns during golf swing: Activation sequence profiling and prediction of shot effectiveness. Sensors, 16:592, 2016

[8] M. Sokolova and G. Lapalme. A systematic analysis of performance measures for classification tasks. Information Processing & Management, 45(4):427–437, 2009

[9] X. Li, W. Yu, et al., MDI Feature Importance Measure for Random Forests, 2019