



# Exploring raw data transformations on inertial sensor data to model user expertise when learning psychomotor skills

Miguel Portaz<sup>1,3</sup> · Alberto Corbi<sup>2,3</sup> · Alberto Casas-Ortiz<sup>1,3</sup> · Olga C. Santos<sup>1,3</sup>

Received: 28 June 2023 / Accepted in revised form: 6 January 2024  
© The Author(s) 2024

## Abstract

This paper introduces a novel approach for leveraging inertial data to discern expertise levels in motor skill execution, specifically distinguishing between experts and beginners. By implementing inertial data transformation and fusion techniques, we conduct a comprehensive analysis of motor behaviour. Our approach goes beyond conventional assessments, providing nuanced insights into the underlying patterns of movement. Additionally, we explore the potential for utilising this data-driven methodology to aid novice practitioners in enhancing their performance. The findings showcase the efficacy of this approach in accurately identifying proficiency levels and lay the groundwork for personalised interventions to support skill refinement and mastery. This research contributes to the field of motor skill assessment and intervention strategies, with broad implications for sports training, physical rehabilitation, and performance optimisation across various domains.

**Keywords** Psychomotor learning · Inertial data analysis · Multivariate time series forecasting · Martial arts · User modelling · Features extraction and selection · Coordinate systems · Quaternions · Data fusion

---

✉ Miguel Portaz  
mportaz@dia.uned.es

Alberto Corbi  
alberto.corbi@unir.net

Alberto Casas-Ortiz  
alberto.casasortiz@dia.uned.es

Olga C. Santos  
ocsantos@dia.uned.es

<sup>1</sup> Artificial Intelligence Department, Universidad Nacional de Educación a Distancia (UNED), Madrid, Spain

<sup>2</sup> Research Institute for Innovation and Technology in Education (UNIR iTED), Universidad Internacional de la Rioja (UNIR), Logroño, Spain

<sup>3</sup> PhyUM Research Center, Universidad Nacional de Educación a Distancia (UNED), Madrid, Spain

## 1 Introduction

According to some business intelligent market research databases, such as Grand View Research,<sup>1</sup> the number of connected wearable devices worldwide by 2025 is expected to be more than 1.5 billion, being the recognition of human activities one of their most predominant uses.

The recognition of human activities through wearable technology represents a groundbreaking approach with broad implications across several domains. This versatile technology finds invaluable applications in well-being and healthy ageing (Camp et al. 2021; Debes et al. 2016; Kaňtoch 2018), enabling individuals to maintain a high quality of life by tracking their movements and activities. Additionally, wearable-based activity recognition approaches plays a pivotal role in physical rehabilitation (Réby et al. 2023; Kim et al. 2021; Meng et al. 2020), offering real-time feedback for tailoring exercise plans to expedite and enhance recovery.

Furthermore, the integration of wearable activity recognition systems into human-machine interaction interfaces is revolutionising the way we engage with technology (Mannini and Sabatini 2010; Huang et al. 2022). Gesture-based controls and motion-sensitive interfaces offer intuitive and seamless interactions with devices, from gaming consoles to virtual reality environments (Dallel et al. 2023), enhancing user experiences and accessibility.

In the realm of sports, this technology serves as a game-changer, providing coaches and athletes with unprecedented insights into performance metrics (Camomilla et al. 2018). From analysing the bio-mechanics of a golf swing (Najafi et al. 2015) to dissecting the intricacies of a gymnast's routine (Krüger and Edelmann-Nusser 2009), wearable devices empower sports professionals to fine-tune training regimens and optimise athletic potential. In addition, the integration of wearable technologies for activity recognition offers a comprehensive perspective approach on how individuals interact with their environment.

However, the future of wearable technology should focus on give more comprehensive assessment on performance analysis (Mason et al. 2022), rather than focus on data analysis for identifying patterns. Although some efforts have been made to support personalised training (Smyth et al. 2021), there are still challenges to transform the captured data into user models that can offer personalised interaction to the users (Hopfgartner et al. 2020).

Embedded sensors within wearables capture intricate details of motion, allowing for comprehensive analysis by experts and delivering quantitative and qualitative insights. This combined approach, marrying precise measurements with expert judgement, ensures an integral understanding of motor actions, paving the way for tailored interventions and finely tuned training, injure prevention and rehabilitation programmes (Portaz et al. 2023). This approach not only considers the technical aspects of the system but also delves into the intricate taxonomies of psychomotor behaviour (Newell 2020). Understanding these diverse categories of motor behaviour provides a nuanced understanding of how our motor actions are planned, executed, and adapted, enriching the overall comprehension of human activity recognition systems.

---

<sup>1</sup> <https://www.grandviewresearch.com/industry-analysis/wearable-patch-market>.

Thus, our research aims to classify practitioners of psychomotor activities according to their expertise level in order to provide personalised multisensorial support (e.g. following an approach similar to Santos et al. (2016), where feedback is delivered through visual, auditory and/or tactile channels to provide a personalised response to the user aimed to support the corresponding learning process) when learning psychomotor behaviours.

Additionally, since Physics are intrinsic to the martial arts domain (Santos and Corbí 2019), and following the idea suggested in James et al. (2014), in our research we have explored if changes in the Cartesian coordinate system used by accelerometers can improve the user modelling. This includes the conversion from Cartesian coordinates to spherical or cylindrical coordinates for redefining points in terms of radial distance, inclination, azimuth angles and height, offering a more specialised perspective. We have also explored fusion approaches to represent spatial orientations, as well as rotations of elements in the three-dimensional space with the use of quaternions. In this context, we pose the following research question:

### RQ

Can inertial data transformations be reliably utilised to assess the proficiency level of a user in the process of acquiring motor skills?

To approach this research question, we have selected the martial arts domain, where it seems possible to identify beginners and experts in the practice of the art from the inertial signals collected (Corbí and Santos 2018). In fact, previous studies on martial arts show that expert's inertial signals are clearer, more regular and balanced (Heinz et al. 2006; Kunze et al. 2006). Furthermore, from a practical point of view, martial arts performance involves exercising physical movements in the sports domain, but they have established a detailed system to differentiate beginners and experts while they progress in the practice.

This paper is structured as follows. First, in Sect. 2 we present the background of our research, including taxonomies of psychomotor behaviour. Next, in Sect. 3 we present the materials and methods. Specifically, in Sect. 3.1 we described the datasets built from practitioners performing three different movements of two different martial arts. After that, in Sect. 3.2 we explain the geometrical representations used to define the spatial orientation of the movements carried out. Following, in Sect. 3.3, we present the modelling approaches applied. In Sect. 3.4, we present the experimental setup, and in Sect. 4, the results obtained are presented and followed by a discussion on the findings. Finally, in Sect. 5 some conclusions are outlined.

## 2 Background

According to Voelcker-Rehage (2008), motor development belongs to adaptive changes in movement behaviour over a person's lifespan, resulting in a long-term progress (*ontogenesis*). Motor learning involves lasting enhancements in motor skill proficiency resulting from training and targeted interventions, which are considered

short-term changes (*microgenesis*). Despite this distinction, learning and development are interdependent; the effectiveness of learning motor skills is influenced by developmental status, and learning outcomes also impact development.

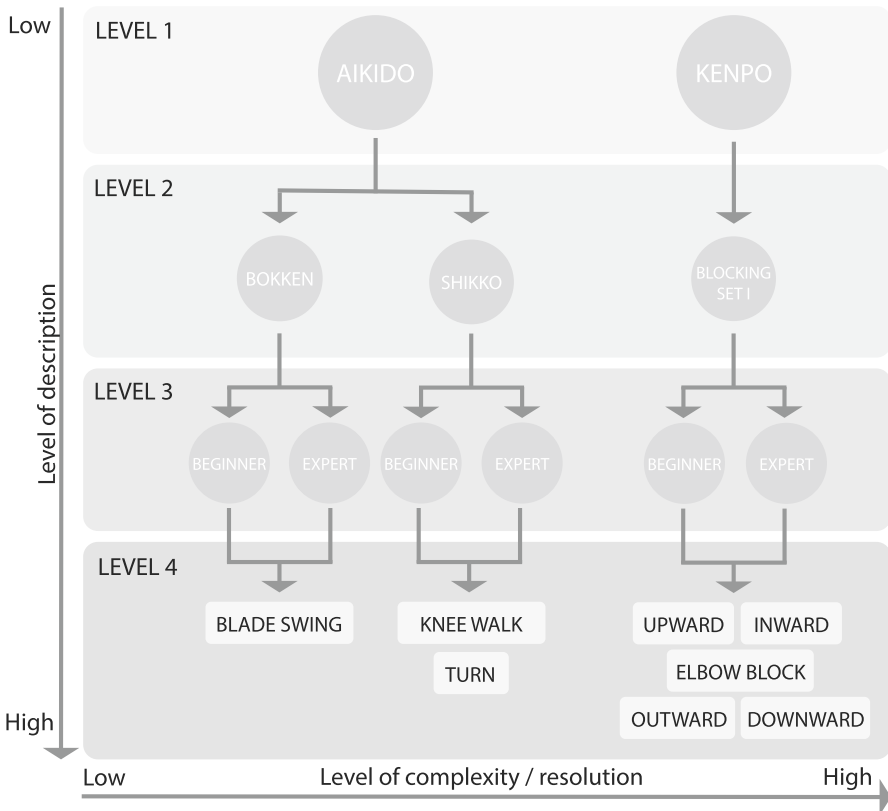
In this context, executing skilled motor actions entails achieving a specific objective with the highest level of assurance, like successfully performing a martial art movement (Schmidt and Lee 2014). A proficient motor behaviour involves a structured and harmonised series of deliberate movements involving body, head, and/or limb actions aimed at achieving a specific result. The coordination of various body parts is essential for executing this motion. Sensory and cognitive inputs play a vital role in shaping an individual's decisions regarding the action and in organising and refining the movements. In addition, motor ability is inherent to an individual, impacting the execution of diverse motor skills. For instance, when executing a punch or kick in martial arts, a practitioner engages in distinct components of the skill, including positioning, release, and follow-through. These actions draw upon underlying motor abilities such as hand-eye coordination, shooting accuracy, agility, and upper body strength (Magill and Anderson 2010).

In skilled motor performance, the goal is to carry out specific limb movements with precise timing and coordination to achieve a desired outcome. Establishing a well-defined ontology for psychomotor planning activities, structured in the form of a tasks hierarchical tree, holds immense significance (Paraschiakos et al. 2020). Such a structured method helps to analyse the relationships between different skills and assess them effectively. This not only enhances our comprehension of psychomotor planning but also serves as a valuable foundation for designing effective training programmes, rehabilitation strategies, and interventions in fields ranging from sports performance to physical therapy.

Our research follows the reverse and analogous process described in Ehatisham-Ul-Haq et al. (2020), evaluating the transition from a fine-grained activity (i.e. walk on your knees, turn and come back in the case of one of the martial arts analysed) to a coarse-grained assessment (expert or beginner) in a basic hierarchical ontology (Fig. 1), and involving a process of increasing generality and abstraction (Olugbade et al. 2022). This process allows a more generalised understanding of the subject matter (expertise level), providing a broader overview of related assessment, enabling to view information in a more encompassing manner, facilitating high-level categorisation and analysis.

## 2.1 Psychomotor learning systems

Psychomotor learning involves the integration of mental and muscular activity with the purpose of learning a motor competence by consolidating it into memory through repetitions (Gagne et al. 1992) and this is done in a gradual manner from a low-performance level (i.e. the learner can hardly recognise the movement) to a high-performance level (i.e. the learner has internalised the movement) (Santos 2016). Nonetheless, supporting personalised learning in this area is still in an early stage of research (Casas-Ortiz et al. 2023) due to the need of modelling more sophisticated or skilled behaviours. The importance of providing personalised support while executing these



**Fig. 1** Hierarchical ontology diagram for different martial art activities. Fine-grained activities are included in Level 4, as coarse-grained assessment is in Level 3

highly skilled movements relies on the relationship between psychomotor learning performance (Fitts and Posner 1967; Schmidt 1976), motor plasticity (Voelcker-Rehage and Willimczik 2006; Guglielmo 2012) and its reinforcement by the fact that psychomotor learning systems can be seen as a closed-loop system (Adams 1971) ready to provide feedback and error detection at latter stage.

Nevertheless, determining the user's personalised behaviours is mentioned as a current challenge (Qiu et al. 2022) and the way users perform the activities is only explored for authentication purposes (Lateef and Abbas 2022). Even so, some of these reviews acknowledge that users have different motion patterns so the way they perform the activities vary from one user to others (Zhang et al. 2022) (between-users variations) and even for a single user the performance of the activity can vary depending on their physical and mental state (Saha et al. 2022) or because the user motion pattern evolves (Miranda et al. 2022) (within-user variations).

Artificial intelligence algorithms on inertial data collected from wearables can be used to compare learner's executions with those from experts during the psychomotor learning process, and this can be used either (i) to recognise specific motion learning

units, or (ii) to assess the learning performance in a motion unit (Santos 2019). In these two cases, the modelling goal should differ, in particular, while in the latter the kind of movement performed is already known, but not how well the user is executing it. In our research, we focus on the recognition of complex human activities using inertial sensor data. Instead of classifying the activity performed, we focus on modelling the level of user motion performance through the knowledge obtained from the motion features, extracted of the movement performed.

## 2.2 Martial arts sensor integration

Our study centres on modelling users expertise level in martial arts movements through the information extracted from inertial sensors. The analysis of psychomotor movements in martial arts represents an opportunity for deploying personalised guided intervention during its practice (Santos 2017). Competence levels are clearly standardised in the form of belts, grading the degree of the practitioner's expertise level (Cox 1993). In particular, practitioners expertise improvement relays on minor details that are revealed when the movement is executed correctly following the Physics involved.

However, up to our knowledge, not many research works to date have focused on analysing martial arts to build personalised psychomotor learning systems. Nonetheless, we have found some works that aim to differentiate between novice and experts from inertial signals, such as Kunze et al. (2006) in Tai-Chi and Heinz et al. (2006) in Kung Fu. In turn, James et al. (2014) used four accelerometers on a practice wooden sword (known as *bokken* in the Aikidō jargon), which measured the inertial signals produced in a basic swing (*shōmen*) and found a correlation between experience level and sensor output. The analysis of the movement consisted in visualising the time series in a cylindrical coordinate system, which turned out to be a very useful way of movement tracking in a three-dimensional space. In other works, accelerometers and gyroscopes were placed on different parts of the body to assess the acceleration profile of a Judo student during the performance of a ukemi (soft fall) (Glowinski et al. 2016).

These works show the potential and feasibility of modelling the expertise level of martial arts practitioners with inertial data. Thus, in order to provide some insights about how to model the expertise level addressing the research question posed, we have built three specific datasets. These datasets contain inertial information for specific complex movements and from different practitioners. The sensors used to build these datasets (accelerometers, gyroscopes and magnetometers) are embedded into specific micro-electro-mechanical system (MEMS), often referred as inertial measurement unit (IMU).

## 2.3 Multivariate time series

The depiction of the inertial information gathered by any of the IMUs is in the form of multivariate time series (MTS) and follows a  $(x, y, z)$  three-dimensional Cartesian coordinate plane representation. Due to the circular nature of martial art movements and following (James et al. 2014), who opted to transform Cartesian values into a

$(\rho, \phi, z)$  three-dimensional cylindrical coordinate system, in our research we have explored different options to transform the raw Cartesian information, including transformations to a  $(r, \phi, \theta)$  three-dimensional spherical coordinate system.<sup>2</sup>

Moreover, besides the above coordinate systems transformations, we introduced another innovation for modelling users on psychomotor learning systems deployments, transforming the raw inertial data into a specific group of hyper-complex numbers, the quaternions. Therefore, we can estimate pitch and roll orientation motion fusing accelerometer and gyroscope information in form of quaternions. Although quaternions have been widely used in different disciplines not directly related to the recognition of human activity and mainly linked with multirotor unmanned aerial vehicles, like in Fresk and Nikolakopoulos (2013); Guerrero-Sánchez et al. (2017); Yang et al. (2017); Xing et al. (2019), Sabatini (2005) uses a quaternion-based integration method for gait analysis, applying the spherical Linear InterPolation (SLERP) procedure disclosed in Shoemake (1985). Furthermore, Sabatini (2006); Renaudin and Combettes (2014); Sung et al. (2018); Bergamini et al. (2014) use magnetic data to increase orientation accuracy and mitigate drift issues (Sabatini 2005). Thus, Sabatini (2006) introduces a quaternion-based extended Kalman filter (EKF) to figure out the orientation of rigid bodies, with applications in the analysis of human movements, while Renaudin and Combettes (2014) and Sung et al. (2018) use wearable sensor data, transformed in quaternions, for indoor pedestrian navigation, and for accurate motion estimation respectively. In Bergamini et al. (2014), similar approaches for manual and locomotion tasks are followed.

To recognise human activity or to assess expertise level, inertial MTS data gathered need to be processed prior to perform psychomotor behaviour classification and recognition. Several approaches can be followed to analyse the collected MTS.

At first glance, Bagnall et al. (2016) categorises classification techniques into different groups according to the type of discriminatory features used. The complete MTS can be analysed, as Zhou et al. (2006) did for estimating the upper limb motion registered, although a time series feature extraction workflow (i.e. preprocessing, segmentation, feature extraction, dimensionality reduction and classification) is a more efficient and effective process for the recognition of human activity (Avci et al. 2010). In this sense, Barandas et al. (2020) provides a complete library for simplifying the extraction of features in several domains, including temporal, statistical and spectral. In the sport domain, extracting features from inertial data was also the method used by Benson et al. (2018) for classifying running speed conditions.

Alternatively, the use of CNNs for MTS classification have also shown good results, as the outcomes obtained with the use of ResNet (Wang et al. 2016) or AlexNet (Fawaz et al. 2020). Since MTS data have essentially the same topology as images, it is possible to apply the same techniques used for image classification to MTS classification. Thus CNNs, which are effective for classifying images, should also be effective for classifying MTS data. The main drawback of these methods is their high computational complexity.

<sup>2</sup> The convention used in this article follows ISO 80000-2:2019, where  $\rho$  is the radial coordinate,  $\theta$  the azimuth,  $z$  the height,  $r$  the radial distance and  $\phi$  the polar angle.

These CNN methods use convolutional kernels to detect patterns in the input time series data. These kernels are convolved through a sliding dot product operation to generate a feature map. As disclosed in Goodfellow et al. (2016), kernels have some basic parameters: size (length), weights and bias, dilation and padding. The resultant kernel, although smaller, has the same basic structure as the MTS. In this case, each kernel is a vector of weights, with a bias term added to the result of the convolution operation between an input MTS and the weights of the given kernel.

In this sense, Choi et al. (2019) used feed-forward artificial neural network models to classify the inertial time series data estimating the centre of mass–centre of pressure (COM-COP) inclination angle during walking using a wearable magnetic IMU. In addition, the framework proposed in Dempster et al. (2020) achieves state-of-the-art accuracy in MTS classification but with a much lower computational expense than previous methods.

Thus, two different approaches seem of interest for our research problem: on the one hand, feature extraction on the time series, and on the other hand, a CNN-based approach.

### 3 Materials and methods

We now present the datasets used for the corresponding modelling process. We also set up the methods we have followed for transforming the original raw Cartesian data (baseline), included in the aforementioned datasets, into spherical and cylindrical coordinates systems. Prior to performing any form of classification, we also explored the benefit of reducing the datasets dimensionality by fusing the inertial data with quaternions. Finally, to infer performance level, we use and compare two modelling alternatives, one based on time series features extraction (Barandas et al. 2020) and the other on a convolutional neural network (Dempster et al. 2020).

#### 3.1 Datasets

To explore the research question, whether inertial data can be used to model the users' expertise level when learning psychomotor skills, and whether the physical characteristics of the movement can be considered in the modelling process, we have built three datasets that collect specific complex movements with different commercial IMUs and in real-world scenarios (i.e. we have collected free-living data on the wild).

The first two datasets gather movements from Aikidō. Aikidō is a non-aggressive Japanese martial art that consists of entering and turning movements that redirect the momentum of an opponent's attack, and ends with a throw or joint lock that terminates the technique (Seitz et al. 1991). To determine the expertise level in the Aikidō datasets, two groups were defined in terms of the corresponding grading belts: expert practitioners (1st kyū to 6th dan)<sup>3</sup> and beginners (7th to 2nd kyū). Note that

<sup>3</sup> Kyū and dan are two levels of rank in Japanese martial arts. The lower rank of kyū is for beginners, while the higher rank of dan is for experienced practitioners. In Japan, dan ranks are typically recognised with a black belt.



we considered 1st kyū Aikidō practitioners as experts because passing from this level to the 1st dan requires a formal external exam and some practitioners delay their examinations. This entails that they stay in the 1st kyū level formally (sometimes for years), but in practice, their mastery of this martial art already fits into the 1st dan.

The third dataset gathers some arm movements of American Kenpo Karate, a modern martial art focused on self-defence, that keeps a balance between tradition but applying modern ideas, like the principles of Physics. It uses quick, body-delivered strikes, enhanced by quick posture changes. Kenpo techniques are taught through scripted outlines that define a set of situations (e.g. an opponent attacks you frontally with a punch, while another opponent is grabbing your arms) (Parker 2009). In this case, the expertise level of the practitioners was defined by belt and years practicing the art. In American Kenpo Karate, a white or yellow belt practitioner is considered a beginner, an orange or purple belt practitioner is considered intermediate, a green or brown belt practitioner is considered advanced, and black belt practitioner is considered an expert. After that, there is a rank of ten black belt dans that ranges from instructor to professor, and finally master and grand master. Reaching this level can take several years of training and continuous learning of movements, concepts and effort.

### 3.1.1 Dataset 1: Aikido–Bokken shomenuchi

This first dataset (D1) gathers one of the Aikidō movements performed with a wooden sword (*bokken*). This movement it is used in James et al. (2014) and explored the utility of changing the coordinate system to improve the modelling of the movement taking into account the Physics behind its execution.

It includes 153 participants ( $N = 153$ ). Ages are between 18 and 69 years old and corresponds to 13 different performance levels, being the highest 6th dan (assigned to  $-6$  for computational purposes) and the lowest 7th kyū (assigned to 7). We considered experts from level  $-6$  to level 1 and beginners above. Other biographical and anthropometric data were also gathered, including the years of martial arts experience, weight, height, gender and the arm and forearm length.

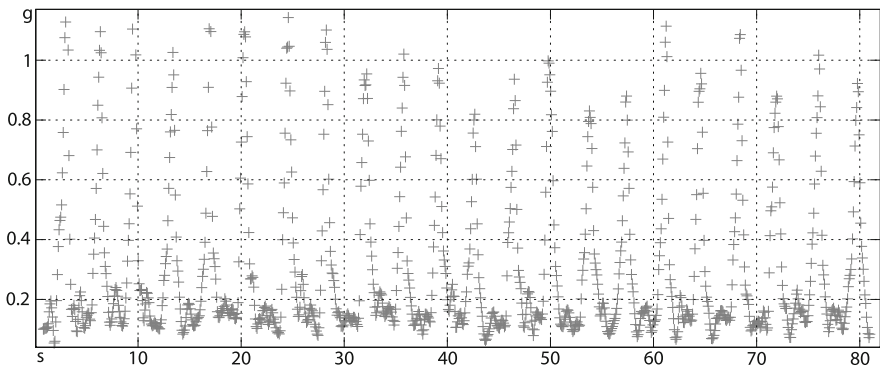
To collect the dataset, we used an Axivity AX3 device<sup>4</sup> attached to the tip of the *bokken* (see right picture on Fig. 2), following a similar approach to the one used in James et al. (2014). Moreover, as in that work, we also video-recorded the data collection to manually label and segment the signal obtained with the device. In particular, the goal of video recording all participants has been no other than to comfortably determine the starting and ending points of each lapse and the overall exercise time, besides easily keeping visual hints of any issues or events worth highlighting.

The movement included in this dataset consisted in performing repeated straight *bokken* swings (see Fig. 2), for a few seconds (80s in our setup). In this sense, the *Shomenuchi* movements itself is a strike to the top of someone else's head. Although this movement might seem simple at first sight, it requires a right body positioning and a correct sword grip. What is more, although speed, and thus, cadence, is important, there is no direct correlation between performer's levels and swing intensity, as shown

<sup>4</sup> AX3 Axivity, Ltd., Newcastle upon Tyne, UK..



**Fig. 2** Three left images represent one straight blade swing (*bokken shomenuchi*) sequence performed by an Aikidō practitioner (who authorised the recording and publication of the images). The rightmost image shows the experimental setup comprising an AX3 accelerometer tightly attached to the tip of one of these wooden Japanese swords



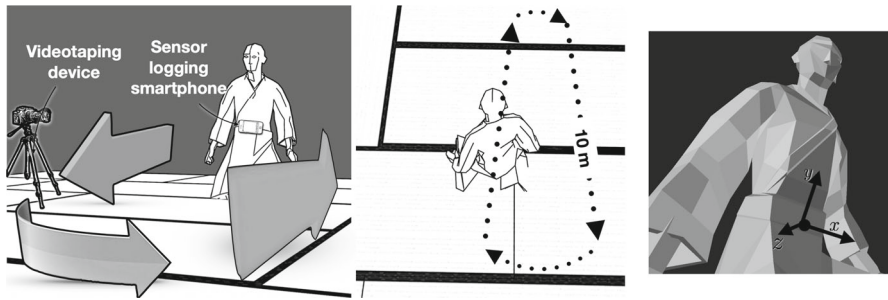
**Fig. 3** Intensity of the linear acceleration as measured in the Aikidō *bokken shomenuchi* setup depicted in Fig. 2

in Fig. 3, where the learner's performance represented shows roughly 22 blade swings during 80 s.

In Fig. 3, the linear acceleration intensity is in fractions of  $g$  (y-axis) and the duration of the exercise is in seconds (x-axis). In this case, the maximum intensity of almost all of these 22 blade swings registered is around  $1 g$ . During the first 30 s the movements registered show certain level of consistence with similar top intensity. After these 30 s top intensity levels are less regular, which may represent a symptom of possible fatigue in the learner.

### 3.1.2 Dataset 2: Aikido–Shikko (knee-walking)

To complement the *bokken shomenuchi* dataset, in Corbí and Santos (2018) we proposed modelling another characteristic movement in Aikidō called *shikko* (very similar to a rhythmic knee-walking and schematically depicted in Fig. 4). *Shikko* movement is difficult to master as it relies on keeping the body centre aligned. It also requires hours of practice and can cause long-term problems on the knees if it is not performed correctly and supervised by an instructor (Homma 2007). However, it is also very useful to develop awareness of one's own centre of mass, also known as *hara* in the Aikidō jargon. The *hara* is the name given in Aikidō for the gravity centre of the body



**Fig. 4** Left and centre images: schema of the movement and representation of the experimental setup. Right image: approximate location of the *hara* in a 3D human figure obtained with the origin to centre of mass operator in Wartmann (2001). The *hara* also represents the origin of the coordinate system of the inertial sensors that we have used to collect the data

and corresponds with just below or directly behind the umbilicus<sup>5</sup> (and is drawn in the right picture of Fig. 4). A correct control of the *hara* contributes to keep a stable position that is later needed for other stand-up techniques as it helps to achieve a correct and swift hip movement, which is essential to master the rest of the Aikidō techniques. Thus, it can (and effectively does) improve the learner's balance even for movements outside of the *shikko* practice itself. The rotational movement that is required to turn when knee-walking is particularly good for learning how to properly shift the hips during the practice of Aikidō, encouraging the development of a strong awareness of the body's centre of mass. Additionally, the *shikko* movement gathered in this dataset can also be analysed to help understand complex concepts of Physics when an embodied learning approach is used (Corbi et al. 2019).

This second dataset (D2) includes 185 participants ( $N = 185$ ). The ages of the participants are between 19 and 69 years old, and participants were classified into 13 different performance levels (from 6th dan to 7th kyū). As in D1, the highest was assigned to  $-6$  and the lowest 7, considering experts from level  $-6$  to level 1 and beginners above. Other biographical and anthropometric data (i.e. years of martial arts experience, weight, height, gender and arm and forearm length) were also gathered.

To collect the data for the D2 dataset, the Aikidō practitioners were asked to perform two 20 ms lapses (40 ms in total), as shown on the left two pictures in Fig. 4 (and also in Fig. 5), while their movements were recorded with the inertial sensors embedded inside a smartphone attached to the practitioners' waists using a fanny pack (also known as bum bag), carefully positioned between the skin (at the level of the navel) and the *aikidogi* (the traditional Japanese garment worn by Aikidō practitioners). This use of type of smartphones devices have proven to meet the expectations for movement recording, as evinced by Saponas et al. (2008); Kos et al. (2016); del Rosario et al. (2015). In this way, the physical variables acceleration and angular velocity were registered relative to the *hara*'s coordinate system using, respectively, the accelerometer and gyroscope packed in the smartphone. The information comprised both acceleration and gyroscopic data along the three ( $x$ ,  $y$ ,  $z$ ) spatial axes; hence, six axes are used.

<sup>5</sup> [https://en.wikipedia.org/wiki/Hara\\_\(tanden\)#In\\_the\\_martial\\_arts](https://en.wikipedia.org/wiki/Hara_(tanden)#In_the_martial_arts).



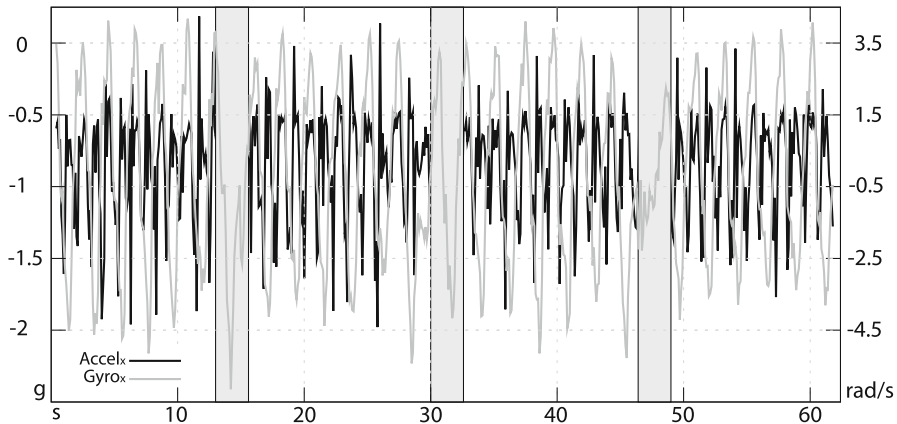
**Fig. 5** Some snapshots (from two of the authors of this research work) showing some straight (go and return) steps of the *shikko* exercise and the associated  $180^\circ$  turns or *ho tenkans* (two rightmost pictures in both rows). The practitioners are wearing an attached fanny pack (enclosing a smartphone) around their waist

The overall exercise was also video-recorded as in D1 with a second synchronised smartphone to facilitate manually segmenting and labelling of the movements. For this, the video metadata included precise timestamp information and both smartphones (including their corresponding data streams) were in sync thanks to the application of Network Time Protocol (Mills 2006, 1985) at the system level.

For this D2 dataset we split every *shikko* performance in any of the phases disclosed in Fig. 4. As a result, we segmented every *shikko* movement in up to 7 different phases (see Fig. 6). This segmentation was useful for either data augmentation and to identify if we can improve the modelling results for any specific phase using different dataset versions, e.g. in the case of the 3 different turn phases (*ho tenkans*) included in the *shikko* movement. Thus, D2 can be analysed in several ways, since it actually includes two different movements: knee-walking in a straight line (going and return) and turning, as represented in Figure 5.

In Fig. 6 we can observe the  $x$  spatial axis component corresponding to the accelerometer (right y-axis, dark line) and the gyroscope (left y-axis, light line) streams for a complete *shikko* exercise with four straight walks (two go's and two return's) and three turns (*ho tenkans*). Focusing on any of the four straight walks, every knee-walking step is represented by a succession of crests and troughs. Thus, counting each of these peaks, either in the acceleration or the gyro component, the number of steps made by this practitioner can be guesstimated. In any of the turn phases, we are only representing gyro information as an example of how well turns are performed following only  $x$  gyro component. In this sense, the last turn is less regular, which may represent a symptom of fatigue.

Regarding the technological setup, the IMU embedded inside the smartphone (model Apple iPhone 8 ®) was set to factory default for gathering inertial data for this dataset.



**Fig. 6** Component  $x$  of the accelerometer (left  $y$ -axis) and the gyroscope (right  $y$ -axis) streams for a complete *shikko* exercise with four straight walks (2 'go's and 2 'return's) and three turns or *ho tenkans* (highlighted in grey and referred to the two right most pictures in Fig. 5)

### 3.1.3 Dataset 3: American Kenpo Karate (Blocking Set I)

This third dataset (D3) gathers the Blocking set I of American Kenpo Karate,<sup>6</sup> which is the first defence set learned by an American Kenpo Karate practitioner. This set of movements has been designed to teach the learners to block hits aiming to different parts of the body. It is formed by six defensive arm movements (upward block, inward block, outward extended block, downward block, rear elbow block and push-down block). However, to gather the dataset, we added a start position and removed the last movement, which is complicated for new practitioners. Thus, we capture the six movements that are shown in Fig. 7. In this way, the D3 dataset complements the other two by including differentiated executions within the movement. In addition, it is easy to compare these different executions as well as to teach the set to facilitate the capturing process with different participants on the wild.

The dataset includes 16 participants ( $N = 16$ ) and 192 total movements (5 different movements performed twice by each practitioner and static noise to detect absence of motion, with right and left arms:  $(6 + 6)$  times  $16 = 192$ ). We evaluated 9 different performance levels, being the highest 18 and the lowest 0 for computational purposes. We considered experts from level 9 and beginners below.

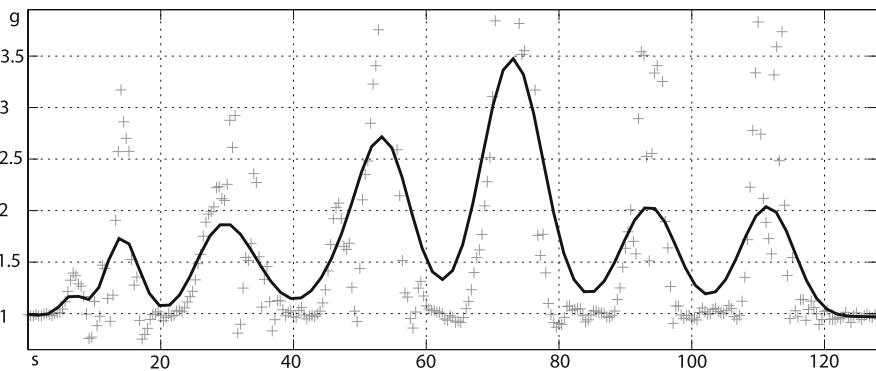
For gathering the inertial information, a smartphone (XiaoMi Mi A2) attached to the practitioners wrist through a running wristband was used, as depicted in the rightmost picture of Fig. 7. For data collection, an ad hoc motion capture software was also developed using Android Studio, whose details and setup are more extensively described in Casas-Ortiz and Santos (2021a, b).

The information obtained from the gyroscope was used to manually label and separate each movement, obtaining the total 192 movements. We only considered the movements performed using the martial artist's dominant hand for a total of 96 movements (192 divided by  $2 = 96$ ), as we centred this research on analysing the

<sup>6</sup> <https://github.com/Physical-User-Modeling-PhyUM/KSAS-Dataset>.



**Fig. 7** An American Kenpo Karate learner (one of the authors of this research work) showing the start position and the 5 blocks conforming the Blocking Set I included in this study (start position, upward block, inward block, outward extended block, downward block and rear elbow block). The right most picture shows the experimental setup consisting in an smartphone attached to the martial artist by means of a runner wristband



**Fig. 8** Intensity of the acceleration stream (grey points) from the set of the Kenpo dataset defence technique that appears in Fig. 7, executed by an expert. The black thick line is the Bézier curve approximation (from the experimental points) overlaid with the goal of highlighting each of the 6 sequential gestures

expertise level. This segmentation was also useful to analyse if we can improve the modelling results for any specific partial movement using different dataset versions. From the technical point of view, the main difference of this dataset (in comparison with the other Aikidō datasets), is that also comprised information from a magnetometer (for obtaining heading information).

A sample stream of the captured acceleration from a participant's execution of Blocking Set I is shown in Fig. 8. In the figure, each one of the oscillations represents a movement. The higher the oscillation, the higher the acceleration and the longer the movement. The reason of this is that each movement is usually executed approximately in the same amount of time, and longer movements must thus be quicker to be executed

**Table 1** Summary of the three datasets used in this research

Dataset	Discipline	Movement	#Mov	#Participants	#Sensor Axis
D1	Aikidō	<i>bokken shomenuchi</i>	1	153	3
D2	Aikidō	<i>shikko (go/return and turn)</i>	2	185	6
D3	American Kenpo Karate	Blocking Set I	6	16	9

**Table 2** Segmentations performed for analysing the datasets

Dataset	Reference (per practitioner)	#Participants (samples)	Segmentation	#Segmented samples
D1	Unique movement	156	1	156
D2	Unique movement	185	7 phases	1,193
D3	Differentiated moves	16	6 movements	96

in the same time as shorter movements. In this case, the participant is an expert that is executing the six movements (including the one removed from the dataset).

### 3.1.4 Summary of datasets

From now on, we will use the following tags to account any of these datasets:

- D1 (*bokken shomenuchi* dataset)
- D2 (*shikko* dataset)
- D3 (Blocking Set I dataset)

Data acquisition for all these datasets was made on real-world environments with free-living data and using commercial devices (i.e. a small accelerometer<sup>7</sup> in D1 and two smartphones in D2 and D3) that embed inertial, and occasionally magnetic and virtual, sensors and are attached to the martial art practitioner (or to an instrument intimately linked to the martial art, as in D1).

These movements and their associated datasets are summarised in Table 1. The column *sensor axis* represents the information collected by the sensors in any of the three ( $x, y, z$ ) dimensions as an independent variable or predictor. In particular, D1 uses only the three axes of the accelerometer, D2 uses six as it includes a gyroscope in addition to the accelerometer, and D3 uses nine, as it also includes a magnetometer. Graphical depictions of any of the movements are included in Fig. 3 (for D1), Fig. 6 (for D2) and Fig. 8 (for D3).

In Table 2, we depict the segmentations performed within each of the three datasets to explore different modelling approaches from the reference movements captured by the participants. For D1 no further segmentation was done as the movement captured is very short. For D2 dataset, seven different segmented phases were considered: 1st go, 1st turn, 1st return, 2nd turn, 2nd go, 3rd turn and 2nd return, thus resulting in two different movements: four straight walks (2 go's and 2 return's) and three turns. In this

<sup>7</sup> Note that attaching a smartphone on the tip of the *bokken* for the D1 dataset was not feasible.

**Table 3** Devices and sensors used for capturing the datasets

	D1	D2	D3
Device	Axivity AX3 ®	Apple iPhone 8 ®	Xiaomi Mi A2 ®
Accelerometer	ADXL345 <sup>1</sup>	SP18382 <sup>2</sup>	BMI120 <sup>2</sup>
Gyroscope	–	SP18382 <sup>2</sup>	BMI120 <sup>2</sup>
Magnetometer	–	–	AK09918 <sup>3</sup>

<sup>1</sup>ADXL345 is manufactured by Analog Devices, Inc

<sup>2</sup>SP18382 and BMI120 are manufactured by Bosch Sensortec GmbH. They are six-axis IMUs that embed accelerometer and gyroscope sensors under one unique encapsulation. The only difference is in their total thickness (BMI120 is 0.9 mm, representing roughly the industry standard, and SP18382 is 0.6 mm). As shown in Table 4, both output resolution and sensitivity are the same

<sup>3</sup>AK09918 is manufactured by Asahi Kasei Microdevices Corporation

case, although the ideal number of segmented samples should be 1295 (185 times 7), some of the phases were not completely registered because were not properly executed, e.g. 2 turns (*ho tenkans*) instead of 3, or because were not completely performed, e.g. due knee injuries, resulting in 1193 samples. For D3 dataset, the 16 participants performed the 6 different partial movements corresponding to the aforementioned arm blocks: start, upward block, inward block, outward extended block, downward block and rear elbow block, thus resulting in 96 segmented samples.

Table 3 contains a summary of all the sensors included in each of the devices. For each of these sensors, Table 4 shows details about sensor output resolution and sensitivity. Note that, in the case of sensitivities, comparing for instance BMI120 and ADXL345 accelerometers at minimum output resolution ( $\pm 2$  g), we have a sensitivity of .06 mg (or 1/16384 g) for the Bosch Sensortec GmbH model and 3.9 mg (or 1/256 g) for the Analog Devices, Inc. model. This means that every time the least significant bit (LSB) changes, we receive variations of .06 mg and 3.9 mg respectively. And therefore, the sensors used in either D2 and D3 datasets are roughly six times more sensitive than the one used in D1 dataset.

### 3.2 Coordinate systems and quaternions transformations

For recognising psychomotor behaviours, either related with martial arts, as in Santos (2019); Kunze et al. (2006); Heinz et al. (2006); Glowinski et al. (2016), or in general for any type of activity, as in Ariza-Colpas et al. (2022); Yuan et al. (2018), analyse the inertial data collected following a Cartesian coordinate system, which is the output information processed by the inertial sensors. In our research, besides using the original raw Cartesian system and the transformed cylindrical system, we have also explored the utility of a spherical coordinate system transformation. As disclosed in Sect. 1 and following the same principle stated in James et al. (2014), the reasons are grounded on the helicoidal nature of some martial arts, including Aikidō and Kenpo. A summary of the different transformations performed is included in Table 5. Note that Cartesian in column coordinates represents the raw original data as output information provided by the sensors (without any transformation), while spherical and cylindrical represent the



**Table 4** Sensor output resolution and sensitivity

model	sensor	axis	output resolution	sensitivity
ADXL345	accelerometer	3	$\pm 2 \text{ g} : 10 - \text{bit}$	256 LSB/g
			$\pm 4 \text{ g} : 11 - \text{bit}$	128 LSB/g
BMI120	accelerometer	3	$\pm 8 \text{ g} : 12 - \text{bit}$	64 LSB/g
			$\pm 16 \text{ g} : 13 - \text{bit}$	32 LSB/g
SP18382	accelerometer	3	$\pm 2 \text{ g} : 16 - \text{bit}$	16384 LSB/g
			$\pm 4 \text{ g} : 16 - \text{bit}$	8192 LSB/g
SP18382	gyroscope	3	$\pm 8 \text{ g} : 16 - \text{bit}$	4096 LSB/g
			$\pm 16 \text{ g} : 16 - \text{bit}$	2048 LSB/g
AK09918	magnetometer	3	$\pm 2000^\circ/\text{s} : 16\text{-bit}$	16.4 LSB/ $^\circ/\text{s}$
			$\pm 1000^\circ/\text{s} : 16\text{-bit}$	32.8 LSB/ $^\circ/\text{s}$
AK09918	magnetometer	3	$\pm 500^\circ/\text{s} : 16\text{-bit}$	65.6 LSB/ $^\circ/\text{s}$
			$\pm 250^\circ/\text{s} : 16\text{-bit}$	131.2 LSB/ $^\circ/\text{s}$
AK09918	magnetometer	3	$\pm 125^\circ/\text{s} : 16\text{-bit}$	262.4 LSB/ $^\circ/\text{s}$
			16-bit	.15 $\mu\text{T}/\text{LSB}$

**Table 5** Raw data and transformations made with the datasets used in this research. Original raw *Cartesian* data have been transformed into *spherical* and *cylindrical* coordinates systems. Subscript *a* for accelerometer data, subscript *g* for gyroscope data and subscript *m* for magnetometer data

dataset	coordinates	dimensions	raw + transformations
D1	Cartesian (raw)	3	$(x_a, y_a, z_a)$
	spherical	3	$(r_a, \phi_a, \theta_a)$
	cylindrical	3	$(\rho_a, \phi_a, z_a)$
D2	Cartesian (raw)	6	$(x_a, y_a, z_a) (x_g, y_g, z_g)$
	spherical	6	$(r_a, \phi_a, \theta_a) (r_g, \phi_g, \theta_g)$
	cylindrical	6	$(\rho_a, \phi_a, z_a) (\rho_g, \phi_g, z_g)$
D3	Cartesian (raw)	9	$(x_a, y_a, z_a) (x_g, y_g, z_g) (x_m, y_m, z_m)$
	spherical	9	$(r_a, \phi_a, \theta_a) (r_g, \phi_g, \theta_g) (r_m, \phi_m, \theta_m)$
	cylindrical	9	$(\rho_a, \phi_a, z_a) (\rho_g, \phi_g, z_g) (\rho_m, \phi_m, z_m)$

transformations of the raw Cartesian data into a spherical and cylindrical coordinates system, respectively.

Moreover, in this research, aside these coordinate transformations, quaternion representations were also derived from the collected  $(x, y, z)$  spatial data. As stated in Sect. 2, quaternions have proven to be specifically useful for describing spatial rotations, estimating pitch and roll orientations, encoding axis-angle rotation data, fusing raw data and consequently reducing the complexity of the datasets (Sabatini 2005). In fact, the classification of the skilled movements discussed in Sect. 2 demands to accurately define, either the position and the heading of the centre of mass of the body, as in the D2 dataset (Sect. 3.1.2), or the wrists, as in the D3 dataset (Sect. 3.1.3).

User modelling in the context of intelligent psychomotor learning systems also needs to orchestrate a way to provide certain level of feedback, as defined in the SMDD

framework (Santos 2016). Thus, instead of classifying the psychomotor behaviour performed, within a set build with different activities, we need to classify the level of expertise (within a set of different performances of the same activity). In this context, quaternion fusion, besides of providing a more accurate depiction of the movement, also reduces the complexity in the analysis for the personalisation support. Consequently, the process is computationally more affordable, as the classification is implemented with less dimensions, 4 in the case of quaternions versus 6 (in D2) or 9 (in D3) dimensions.

Before introducing the different methods used to derive quaternions, we define them as an extension of complex numbers with three square roots of  $-1$  ( $ijk$ ) instead of just  $i$ :

$$i^2 = j^2 = k^2 = i j k = -1 \quad (1)$$

Thus, the first component is a scalar real number  $s$  and the other 3 form a vector  $\vec{v}$  as follows:

$$q = q_0 + iq_1 + jq_2 + kq_3 = \langle s, v \rangle \quad (2)$$

where  $s = q_0$  and  $v = [q_1 \ q_2 \ q_3]$ . For convenience we will use only unit length *quaternions*:

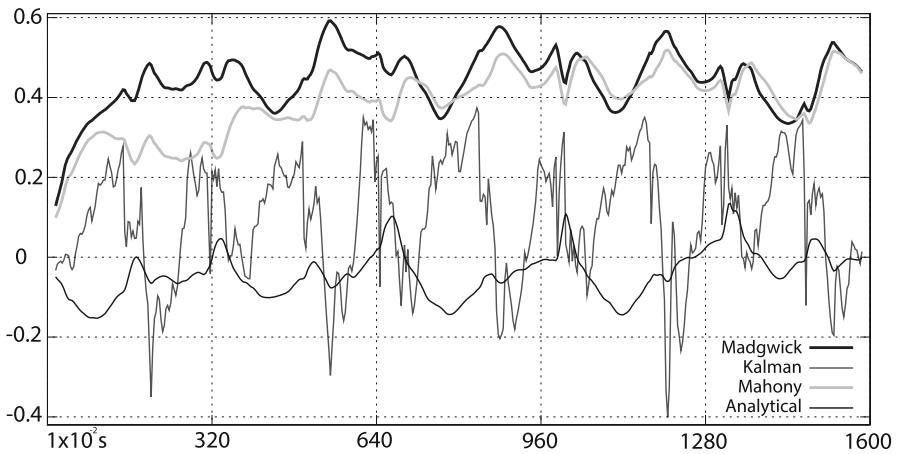
$$|q| = \sqrt{q_0^2 + q_1^2 + q_2^2 + q_3^2} = 1 \quad (3)$$

To fuse inertial data into quaternions, we need, at least, information from a six-axis device (which includes at least inertial information from accelerometer and gyroscope sensors, as in D2 and D3 datasets) being not possible to fuse the information gathered with a three-axis sensor (just an accelerometer sensor, as in D1 dataset). Consequently, we performed quaternion fusion in D2 and D3 datasets but not in D1 datasets, see Table 1 and Table 3.

Quaternion calculation from the original data was based on the framework described in Haslwanter (2020), which provides up to four different methods for fusing sensor data: analytical, Mahony (Mahony et al. 2008), Madgwick (Madgwick et al. 2011) and Kalman, as in Sabatini (2011); Guo and Hong (2019). Analytical method does not take into account magnetometer values, while the Mahony and Madgwick ones use this information for improving the accuracy. The Kalman method requires magnetometer information in order to properly operate. Without this sensor, quaternion fusion is often degraded (as evinced by Fig. 9).

Using the analytical method, simple quaternion integration is performed calculating orientation and position analytically from angular velocity and linear acceleration without drift compensation, so no magnetometer information is used. In the case of using Madgwick and Mahony methods, magnetometer information is optional. Madgwick method uses gradient descent filter and consequently is computationally more demanding. This method usually represents the most accurate transformation when magnetometer information is used (Ludwig et al. 2018) and includes gyroscope bias drift compensation.

Therefore, for each of the datasets included in this research (see Sects. 3.1.1, 3.1.2 and 3.1.3), we explore whether we can get better results in assessing the level of experience of martial arts practitioners by transforming and/or fusing the original raw inertial (and occasionally magnetic) data. The reasons for including the analysis



**Fig. 9** Value of the  $q_3$  component according to several methods for deriving quaternions and using sensor information encoded in Cartesian coordinates. The represented data correspond to D2 dataset, in particular to the first go lap of *shikko* exercise (performed by a beginner student). In this case, the magnetometer information was missing and was introduced as a constant in the computation for the Kalman method (this translates itself into a roughly degraded outcome)

of different coordinates system transformations and quaternion fusions are to have a better and more accurate representation of the psychomotor behaviour analysed. While coordinate systems transformations better characterise the nature of the motor skill, the quaternion fusion provides a more accuracy depiction of this. To perform a skilled movement classification, where all registries represents the same movement, we need to refine the inertial information to better represents the movement to classify, which we aim to achieve by performing transformations and fusions.

A summary of the different dataset versions used in this research is included in Table 6, where the baseline ( $v_0$ ) in any of the dataset portrays the original data without any transformation or fusion. In fact, we evaluated up to 33 different datasets, including the three original ones. In this sense, for D1 we only evaluated the raw Cartesian ( $v_0$ ) and both, spherical ( $v_1$ ) and cylindrical ( $v_2$ ) transformations, as quaternion fusion was not possible without gyroscope information. For either D2 and D3 datasets, we evaluated 15 different dataset versions for each of them, from ( $v_0$ ) to ( $v_{14}$ ). As an example, in the case of D3 dataset, version 7 ( $v_7$ ) corresponds to the original Cartesian D3 dataset transformed into a spherical coordinate system and then fused into quaternions following analytical method, that is, quaternion fusion without magnetometer data. Consequently, while any of the versions ( $v_0$ ), ( $v_1$ ) and ( $v_2$ ) of D3 represent a dataset with  $3 + 3 + 3 = 9$  dimensions (see Table 5), the version ( $v_7$ ) of dataset D3 represents the same, but with only 4 dimensions (see equation 2).

**Table 6** Summary of the 33 different versions built for this research, from the original raw Cartesian data gathered from the sensor without any modifications, referred as version 0 (v0) and considered as the baseline version for each of the datasets

dataset	version	coordinates	Fusion method (following [81])
D1	v0	Cartesian	-
	v1	spherical	-
	v2	cylindrical	-
D2	v0	Cartesian	-
	v1	spherical	-
	v2	cylindrical	-
	v3	Cartesian	Analytical
	v4	Cartesian	Madgwick
	v5	Cartesian	Mahony
	v6	Cartesian	kalman
	v7	spherical	Analytical
	v8	spherical	Madgwick
	v9	spherical	Mahony
	v10	spherical	kalman
	v11	cylindrical	Analytical
	v12	cylindrical	Madgwick
	v13	cylindrical	Mahony
v14	cylindrical	kalman	
D3	v0	Cartesian	-
	v1	spherical	-
	v2	cylindrical	-
	v3	Cartesian	Analytical
	v4	Cartesian	Madgwick
	v5	Cartesian	Mahony
	v6	Cartesian	kalman
	v7	spherical	Analytical
	v8	spherical	Madgwick
	v9	spherical	Mahony
	v10	spherical	kalman
	v11	cylindrical	Analytical
	v12	cylindrical	Madgwick
	v13	cylindrical	Mahony
v14	cylindrical	kalman	

### 3.3 Modelling approaches

We considered all the 33 dataset versions in Table 6, to explore ways to assess expertise level, modelling skilled human motion movements, and to explore ways to offer personalised support to students when learning psychomotor skills (Santos 2019).

As data is collected in form of MTS, the modelling processes need to identify trends and patterns in the inertial data gathered. This requires to analyse the relationship between dependent variables (i.e. the expertise level) and independent variables (i.e. the inertial data registered in each of the corresponding dataset versions obtained (15 for D2 and D3 and just 3 in the case of D1). To model these skilled movements performed by martial art practitioners, exploring the underlying relationships between outcome (expertise level) and predictors (inertial data), we followed two different approaches, as disclosed in Sect. 2.

To identify MTS trends and patterns, we need to extract the most relevant features which best represent the inertial data gathered. One of the two different approaches used in this research for this extraction is similar to the one used in Avci et al. (2010), which conducts the feature extraction in different domains, including time (extracting, for example, MTS mean, variance or standard deviation features) and frequency (extracting MTS spectral centroid or energy features among others). For this approach, we used (Barandas et al. 2020) library, which provides a reliable and rapid method to follow (Avci et al. 2010).

The second approach uses a neural networks based method instead for extracting features and was initially introduced in Jafari et al. (2007); Yang et al. (2008), although in neither case they focused on skilled analysis, performing movement classification but not expertise level assessment. Recent efforts have been made to classify human activity using neural networks, as in Dempster et al. (2020); Fawaz et al. (2020), where the convolutional kernels used in convolutional neural networks are applied to detect patterns in the input MTS (inertial data).

Although the first approach, without using neural networks, is computationally faster, deep learning approaches for classifying human activity have gained relevance (Babangida et al. 2022) and we can find different neural network approaches for human activity recognition, including (Tran et al. 2019; Fu et al. 2020; Wang and Miao 2018; Ferscha and Mattern 2004). However none of them analyse skilled movements and efficiency limitations are still challenging (Zhou et al. 2020), especially when one of the ultimate goals is to provide real-time feedback as required in intelligent psychomotor learning systems (Santos 2016). The reason of using random convolutional kernels in this research for classifying MTS that represent skilled movements is that the random kernels used have less computational requirements and consequently are faster than other neural network methods.

### 3.3.1 Time series feature extraction

The process included in Avci et al. (2010) for activity recognition follows the stages of pre-processing, segmentation, feature extraction, dimensionality reduction and classification. In the same way, this processing approach has already been discussed as appropriate to model the performance of psychomotor activities from inertial sensor data (Santos 2019). In any case, we will not delve into the first pre-processing stage because in the three datasets used here the raw signals obtained have been partially processed by the sensors themselves through their data processing units.

In order to evaluate how we can use extracted features from MTS, we followed (Barandas et al. 2020) and their Time Series Feature Extraction Library (TSFEL), which provides extraction methods for different features across 3 different domains (temporal, statistical and spectral).<sup>8</sup> The analysis performed in this research was configured using all these domains. In this sense and according to Barandas et al. (2020), most of the spectral domain features have higher computational complexity than the features included in the other domains (temporal and statistical). We applied this

<sup>8</sup> List of available features: [https://tsfel.readthedocs.io/en/latest/descriptions/feature\\_list.html](https://tsfel.readthedocs.io/en/latest/descriptions/feature_list.html).

extraction methods to each of the 33 different versions ( $3 + 15 + 15$ ) obtained from Sect. 3.2 and depicted in Table 6.

After performing the feature extraction stage on each of the versions included in Table 6 for each of the three different datasets, we carried out the dimensionality reduction stage. However, instead of Hartmann et al. (2021), which applied linear discriminant analysis (LDA) for dimensionality reduction, we applied least absolute shrinkage and selection operator (LASSO) linear model with iterative fitting along a regularisation path, as in Liu et al. (2020). We used this method as a general rule in all datasets for its speed and versatility and because it usually produces better results when the dimensionality is high, which is our case after applying TSFEL.

### 3.3.2 Random convolutional kernel transform

One step forward to the use of typical CNN methods for classifying MTS data is the use of Random Convolutional Kernel Transform (ROCKET) framework, introduced in Dempster et al. (2020). This method may achieve good accuracies, even with basic linear classifiers, while using a fraction of the computational complexity of other CNN methods, like Wang et al. (2016) or Fawaz et al. (2020). In contrast to the convolutional kernels typically used in CNNs, ROCKET generates a large number of random convolutional kernels which, in combination, capture features relevant for MTS classification.

ROCKET computes two features from each convolutional layer. The first feature is calculated using a standard approach called global/average max pooling,<sup>9</sup> which takes the maximum value from each part of the layer. The second feature is calculated using a unique approach called positive proportion value (ppv), which takes the proportion of positive values from each part of the layer. These two features are then combined to create a more robust feature map.

In our research, for convenience and simplicity we set ROCKET with the default number of random kernels ( $k = 10,000$ ), having  $2k$  features (20,000) per time series as output from the transform. Once the transformed feature map is generated we can use it as input data for any classification algorithm. Dempster et al. (2020) suggest to use some common linear algorithms like ridge regression classifier or logistic regression.

In addition for this research, we also tested some CNNs, including ResNet (Wang et al. 2016), InceptionTime (Fawaz et al. 2020), Time Le-Net (Fawaz et al. 2019; Guennec et al. 2016) or Time Warping Invariant Echo State Network (TWIESN) (Lukosevicius et al. 2006). Our outcomes using ROCKET surpassed those obtained with these CNNs, which were also adding high computational complexity to the classification. In this sense, our initial results comparing different CNNs methods versus ROCKET are similar to those already included in Dempster et al. (2020). Finally, while ROCKET execution takes minutes, CNNs may take hours with similar or lesser accuracy results.

<sup>9</sup> The difference between *global* and *average* is that the first method become the process of taking the maximum value from the entire result of the convolution whether the second method is the process of taking the average maximum value within a predetermined area of the convolution result.

**Table 7** List of classifiers used with TSFEL and ROCKET extracted features

Abbreviation	Classifier	parameters
RF	Random forest	n_estimators = 100, criterion = "gini"
LR	Logistic regression	solver = "lbfgs"
RCV	Ridge cross-validation	alphas=(0.1, 1.0, 10)
SGD	Stochastic gradient descent	max_iter=1000
kNN	k-Nearest neighbour	k = 4

### 3.4 Experimental setup

We have analysed 33 different dataset versions obtained from the 3 original datasets ( $v_0$ ,  $v_1$  and  $v_2$ ) as compiled in Table 6 and with the 2 methods disclosed in Sect. 3.3.1 (TSFEL) and Sect. 3.3.2 (ROCKET). In order to evaluate the classification behaviour of the proposed approaches for features extraction, we used some well-known classifiers included in the reputed scikit-learn library,<sup>10</sup> so we could compare the different proposals. A summary of the classifiers and the parameters used is included in Table 7. The parameters were established after several trials. The input for any of them is the features obtained by either method used: TSFEL and ROCKET.

In particular, the random forest (RF) classifier was chosen because it is more robust and easier to train than others. For this classifier, the number of trees was set to 100, and Gini impurity was established as the function to measure the quality of the splits. The logistic regression (LR) classifier was included as suggested while using ROCKET framework, and the algorithm (solver) for optimisation was the limited memory Broyden-Fletcher-Goldfarb-Shanno (LBFGS). The ridge cross-validation (RCV) classifier is also recommended by ROCKET framework for smaller datasets (due to fast cross-validation of the regularisation parameter). For the RCV classifier the values for the regularisation strength ( $\alpha$ ) were chosen accordingly for reducing the variance of the estimates. The stochastic gradient descent (SGD) classifier is particularly useful when the number of samples is very large. For SGD the maximum number of iterations taken for the solvers to converge was 1000. Finally, the k-nearest neighbour (kNN) classifier was chosen because is robust to noisy training data and effective when training data is large. The number of neighbours (k) was set to 4 after several trials and due showed good results. For the configuration of these classifiers, we decided to mainly use default values in order to better compare the different types of algorithms and representations, without worrying about differences in the parameters.

Before applying the classifiers, all MTS values were scaled using min-max normalisation with a feature range between 0 and 1. This was done basically because, as disclosed in Sect. 3.2 and in Eq. 3, we use only unit length quaternions, so most of the features obtained are already normalised in this scale.

Thus, applying the classifiers depicted in Table 7 we evaluated the accuracy on modelling users performing skilled movements. We also analysed if we can obtain better results using the baseline version ( $v_0$  in Table 6) or using different coordinates

<sup>10</sup> <https://scikit-learn.org/stable/>.

systems or quaternions (other versions than  $v0$  in Table 6). For all dataset versions included in Table 6, we performed MTS re-sampling method before applying any of the classifiers in Table 7. This operation was particularly useful while using ROCKET, where all kernel input data sizes were equal, but also when using TSFEL, for avoiding window overlaps.

As disclosed in Sects. 3.1.2 and 3.1.3, besides modelling the complete skilled movements performed by any of the martial artists in each dataset, we also extended our analysis for comparing specific segmented movements as disclosed in Table 2. Thus, in this research, for either D2 and D3 datasets, we analysed both the whole skilled movement and its corresponding segmentations into 7 phases or 6 partial movements, respectively.

We included a basic number of re-shuffling and splitting iterations (2) and the proportion of the dataset included in the test split was (30.%). In order to control the randomness of the training and testing indices produced, we used the same seed value for a reproducible output across multiple function calls. A visual overview of the experimental setup, specifically for D2 and D3 datasets, is presented in Fig. 10 and Fig. 11.

Initially, see Fig. 10, we start with basic transformations (2) from the original dataset collected (1). Then, sensor fusion is performed (3), resulting in a total of 15 dataset versions.<sup>11</sup> Then in Fig. 11, the next step (4) is the homogenisation performed with any of the dataset versions. Since martial artist performances differ in time, the execution of each practitioner is represented by a different number of rows, re sampling each performance we got homogeneous<sup>12</sup> input data for the next phase (5). As disclosed in Sects. 3.3.1 and 3.3.2, feature extraction (5) is performed with ROCKET and TSFEL. The resultant feature map and feature list (6) are finally used to feed the classifiers disclosed in Table 7.

## 4 Results and discussions

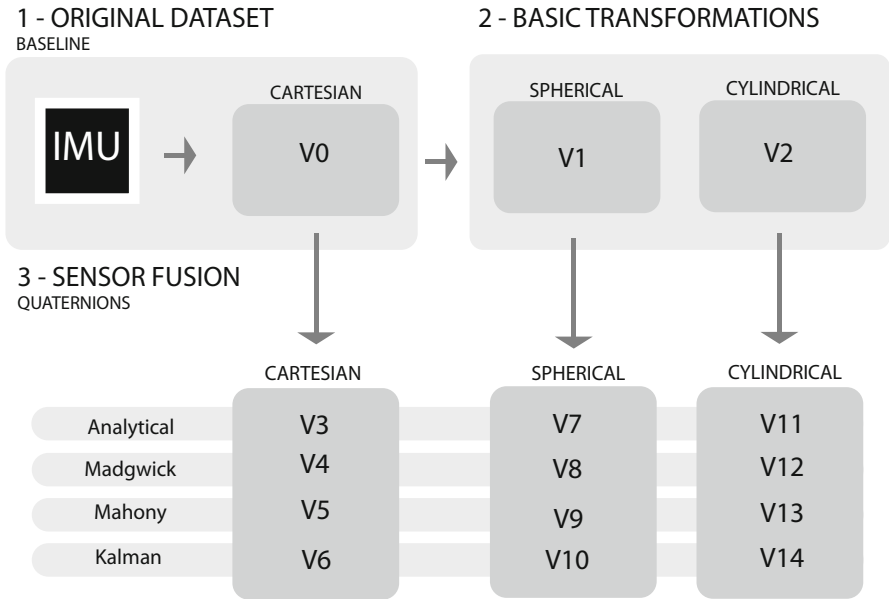
In this section, we provide a detailed study of the results obtained to assess the level of experience of martial artists when they execute skilled movements. As we have been anticipating, there are important differences when it comes to modelling users psychomotor behaviour depending on the type and form of the motion data collected. Therefore, as we might suppose, the accuracy in the results will differ accordingly depending on the dataset used.

Regarding the methods used for extracting the discriminatory features, in the case of TSFEL we obtained better results while using the statistical domain above the others (temporal and spectral). Besides, applying this domain we balanced the computational complexity, the total number of extracted features and the final accuracy obtained. In most of the cases, the resulting feature list, after removing redundancies and noise,

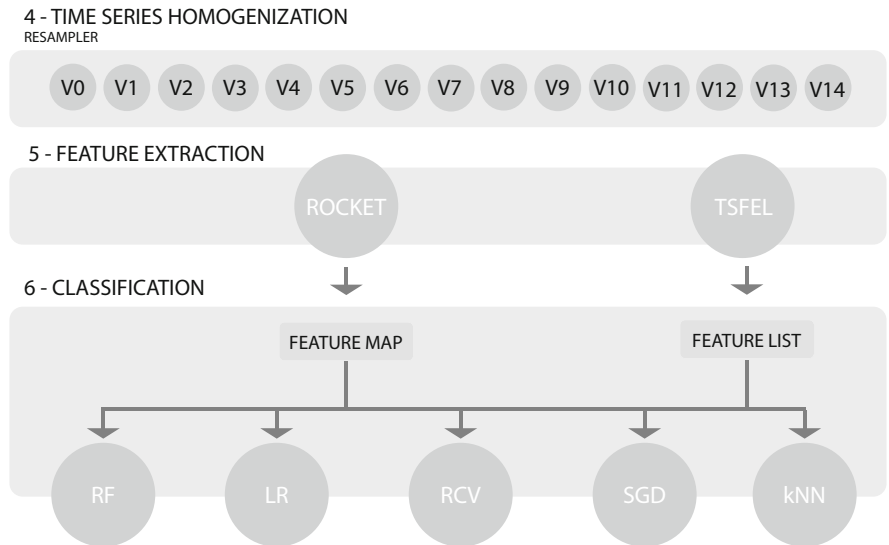
<sup>11</sup> As an example and in the case of D2, dataset versions V0, V1 and V2 have an structure of 596, 495 rows by 6 columns, while the remaining fusion versions (from V3 to V14) comprise a dataset of 596, 495 rows by 4 columns.

<sup>12</sup> As an example and in the case of D2, all dataset versions (any of them with 596, 495 rows) are re-sampled, resulting in datasets of 925, 000 rows (5000 rows for any of the 185 martial artists).





**Fig. 10** Transformations on a martial art dataset collected with inertial sensors (accelerometer, gyroscope and/or magnetometer) that result in alternative dataset versions for modelling the user expertise. V0 is considered the baseline version from which the initial v1 and v2 (basic transformations) are obtained. Then, for any of these 3 initial versions, a fusion into *quaternions* is performed



**Fig. 11** General overview of the experimental setup for D2 and D3 datasets. For any of the 15 different versions obtained during the process depicted in Fig. 11, a phase of time series homogenisation and feature extraction is made prior performing the final classification

**Table 8** D1—TSFEL statistical domain results. In bold the best result obtained

Dataset version	Highest accuracy (%)	Classifier	Samples
v0 (raw)-baseline	56.52	RF	153
v1	54.34	RF	153
v2	<b>60.89</b>	kNN	153

**Table 9** D1—ROCKET results. In bold the best result obtained

Dataset version	Highest accuracy (%)	Classifier	Samples
v0 (raw)—baseline	<b>67.39</b>	LR, RCV and SGD	153
v1	65.21	LR and RCV	153
v2	<b>67.39</b>	LR and SGD	153

usually include only few dimensions, just the most representatives. In the case of using ROCKET, the process is more straightforward, as the only parameter to choose is the number of random kernels. In any case, the features obtained with any of these two methods and for any of the dataset versions can be used directly as input data for any classification algorithm disclosed in Table 7.

In all the result tables disclosed in this section, in column *classifier* we are including the classifier or classifiers whose highest accuracy is obtained among those included in Table 7. Column *samples* represents the number of martial artist performances included in any of the different datasets (see Table 1).

#### 4.1 D1 dataset results

In the case of D1 dataset (see Sect. 3.1.1), the results obtained after extracting features with TSFEL (see Sect. 3.3.1) are summarised in Table 8. The results after classifying the feature map obtained using ROCKET (see Sect. 3.3.2) are disclosed in Table 9. Note that for D1 dataset we can only consider 3 different versions (see Table 6), as quaternion fusion was not possible to perform with this dataset because we only have information from one sensor (i.e. the accelerometer). This dataset also did not support either any kind of segmentation (see Table 2).

The best result for user (*martial artist*) modelling when performing the *bokken shomenuchi* movement is obtained while using the dataset version with *cylindrical* coordinate transformations (*v2*) and after applying *kNN* as classifier (see Table 7) in the case of using TSFEL (60.89%). In turn, ROCKET best results (67.39%) are obtained both while using baseline dataset version (*v0*) and dataset version with *cylindrical* coordinate transformations (*v2*) and after applying different classifiers (*LR*, *RCV* and *SGD*) (see Table 7). Furthermore, for the D1 dataset, even in the worst case version (*v1*), the lower accuracy result after applying ROCKET (65.21%) is better than any result obtained when using TSFEL (see Table 8).

## 4.2 D2 dataset results

As expected, with D2 dataset (see Sect. 3.1.2) we are improving the accuracy in modelling users (martial artists) from D1, because we have the information from two sensors (accelerometer and gyroscope). The results obtained after using TSFEL are summarised in Table 10, while the results applying ROCKET are available in Table 15. In either case, we are only showing the results obtained with the baseline ( $v0$ ) version and with basic transformation versions, ( $v1$ ) and ( $v2$ ), plus the best result obtained with any of the quaternion fusion versions (see Table 6).

Focusing first on the TSFEL modelling, upper half of the Table 10 (*segmented*) includes the summarised results obtained while analysing D2 dataset per phases. Version ( $v6$ ) corresponds to the quaternion fusion following the Kalman method (see Sect. 3.2) with Cartesian coordinates (see Table 6). In the lower half of the Table (*complete*), we have the results obtained while analysing D2 dataset per martial artist, considering the complete movement, given that in this case, contrary to the previous dataset, it is possible to split the movements into different phases (see Table 2). Version ( $v4$ ) corresponds to the quaternion fusion following the Madgwick method (see Sect. 3.2) with *Cartesian* coordinates (see Table 6).

After using TSFEL, the best results for user (*martial artist*) modelling, when performing the *shikko* movement, analysed per phases, are obtained while using baseline dataset version ( $v0$ ), and dataset version with *spherical* coordinate transformations ( $v1$ ) and after applying *RF* as classifier (see Table 7) in both cases. When analysing the complete movement, the best result is obtained while using baseline dataset version ( $v0$ ) and also after applying *RF* as classifier (see Table 7). The baseline dataset version ( $v0$ ) in the complete dataset is the one that obtains the highest accuracy for D2 dataset using *RF* classifier.

As an example of the use of TSFEL in this dataset, one of the features that is usually included in the resultant feature list, in almost all versions except in versions ( $v3$ ), ( $v5$ ) and ( $v6$ ), is the histogram of the signal, computed as follows:

$$n = \sum_{i=1}^k m_i \quad (4)$$

**Table 10** D2-TSFEL statistical domain results-summary. In bold the best result obtained

	dataset version	highest accuracy	classifier	samples
segmented	v0 (raw) - baseline	76.12%	RF	1184
	v1	76.12%	RF	1184
	v2	73.87%	RF and kNN	1184
	v6	65.16%	RF	1184
complete	v0 (raw) - baseline	<b>78.57%</b>	RF	185
	v1	75.0%	RF	185
	v2	71.42%	RF	185
	v4	71.42%	RF	185

**Table 11** List of the relevant features for the version  $v4$  of the D2 dataset

Input variable	Feature	Bin
$q_0$	ECDF	7
$q_1$	ECDF	6
$q_1$	Histogram	0
$q_1$	Histogram	1
$q_1$	Histogram	2
$q_1$	Root Mean Square	None <sup>1</sup>

<sup>1</sup>As this feature does not use bins, TSFEL calculates this feature only 4 times, 1 per input variable

where  $m_i$  represents the histogram in which  $n$  is the total number of observations and  $k$  the total number of bins (Barandas et al. 2020). For the versions ( $v4$ ), ( $v5$ ) and ( $v6$ ), the most representative feature that is included in all its feature list is the empirical cumulative distribution function (ECDF) along the time axis, which is also included in other versions, and which formula is as follows:

$$\frac{1}{n} \sum_{i=1}^n I[x_i \leq x] \quad (5)$$

where  $n$  is the number of data points,  $I$  is the indicator function, and  $x_i$  is the  $i$ th data point. In this case, the indicator function is used to count the number of data points that are less than or equal to a given value ( $x$ ). If the result of this comparison is true, then  $I$  is equal to 1, and 0 otherwise.

Other features, although present in some of the feature list versions, are less common, including variance, median, mean, root mean square, skewness and kurtosis among others. As a practical example of this, in Table 11 we are representing the most relevant extracted features for the  $v4$  version of the D2 dataset (see Table 6), that were used to feed the classification algorithms. In this case, after using LASSO for dimensionality reduction, we finally picked only 6 features from the total number of features obtained with TSFEL. Initially and for this dataset, for any of the 4 different input variables ( $q_0$ ,  $q_1$ ,  $q_2$  and  $q_3$ ) TSFEL calculates up to 16 different statistical domain features. TSFEL also calculates several times (10 by default and so called bins) some of this features (e.g. Histogram and ECDF), so each bin correspond to a different feature. Therefore, in Table 11, column bins represents the number of those that are more relevant. For example, in the case of  $q_1$  and Histogram there are up to 3 relevant bins: 0, 1 and 2, where the other 7 bins were not relevant in this case.

Note that, in addition to the above results, for D2 dataset and TSFEL method, we also evaluated movements considering only some stages separately (see Fig. 6), specifically the 2 different *going* phases (see Table 12), the 2 different *return* phases (see Table 13), as well as any of the 3 *turns* analysed jointly (see Table 14). In this case, for the *going* and *return* phase analysis, the best results are obtained with the *spherical* coordinate transformations ( $v1$ ) and the *turn* phase, due its circular nature, is better

**Table 12** D2—TSFEL only going movements results—summary. In bold the best result obtained

Dataset version	Highest accuracy (%)	Classifier	Samples
v0 (raw)—baseline	71.42	RF and RCV	185
v1	<b>75.</b>	RF	185
v2	73.21	RF	185
v4	67.85	LR	185

**Table 13** D2—TSFEL only return movements results—summary. In bold the best result obtained

Dataset version	Highest accuracy (%)	Classifier	Samples
v0 (raw)—baseline	66.07	RF and RCV	185
v1	<b>73.21</b>	RF	185
v2	71.42	RF	185
v10	71.42	RCV	185

**Table 14** D2—TSFEL only turn movements results—summary. In bold the best result obtained

Dataset version	Highest accuracy (%)	Classifier	Samples
v0 (raw)—baseline	63.04	RF and LR	185
v1	65.21	RF, LR and SGD	185
v2	65.21	RF, LR and RCV	185
v5	<b>76.08</b>	RF	185

**Table 15** D2—ROCKET results—summary. In bold the best result obtained

	dataset version	highest accuracy	classifier	samples
segmented	v0 (raw) - baseline	72.62%	kNN	1193
	v1	78.21%	RCV	1193
	v2	76.25%	LR	1193
	v6	68.43%	RF and LR	1193
complete	v0 (raw) - baseline	73.21%	kNN	185
	v1	<b>82.14%</b>	RCV and SGD	185
	v2	78.57%	LR and RCV	185
	v9	73.21%	LR	185

to model using *quaternions* (v5). In all these 3 analysis (see Table 12, Table 13 and Table 14) the best results are obtained after applying *RF* as classifier (see Table 7).

Finally, in the case of using ROCKET with D2 dataset, the best results for user (*martial artist*) modelling, when performing the *shikko* movement, analysed per phases, is obtained while using dataset version with *spherical* coordinate transformations (v1) and after applying *RCV* as classifier (see Table 7). The best result analysing the complete movement, is obtained while using dataset version with *spherical* coordinate transformations (v1) and after applying *RCV* and *SGD* as classifiers (see Table 7).

**Table 16** D3—TSFEL statistical domain results—summary. In bold the best result obtained

	dataset version	highest accuracy	classifier	samples
segmented	v0 (raw) - baseline	68.96%	kNN	96
	v1	65.51%	RF	96
	v2	65.51%	RF and kNN	96
	v4	75.86%	RF	96
complete	v0 (raw) - baseline	<b>80.0%</b>	RCV and SGD	16
	v1	<b>80.0%</b>	RF, SGD and kNN	16
	v2	<b>80.0%</b>	RF, RCV and SGD	16
	many	<b>80.0%</b>	kNN	16

**Table 17** D3—ROCKET results—summary. In bold the best result obtained

	dataset version	highest accuracy	classifier	samples
segmented	v0 (raw) - baseline	48.27%	kNN	96
	v1	58.62%	kNN	96
	v2	48.27%	RF and kNN	96
	v11	72.41%	RF and LR	96
complete	v0 (raw) - baseline	60.0%	RF, LR, RCV and kNN	16
	v1	60.0%	RF, LR, RCV and kNN	16
	v2	60.0%	LR all	16
	v14	<b>80.0%</b>	RF	16

In any case, considering both TSFEL and ROCKET methods for this D2 dataset, the best results in either case, regarding the whole movement or any of the segmented phases, are obtained with the baseline version (*v0*) in the case of using TSFEL (78.57%) or with one of the basic transformations (*v1*) in the case of using ROCKET (82.14%), representing the best efforts obtained for this D2 dataset after applying any of the methods disclosed in Sects. 3.3.1 and 3.3.2. On the other hand, quaternion versions are only useful to model turn phases separately, obtaining in this case a best result of 76.08% versus the 65.21% obtained with baseline and basic transformation versions.

Note that sample figures in the TSFEL analysis (1184 in Table 10) differs from those obtained with ROCKET (1193 in Table 15) because in the case of TSFEL, we removed some very short phases to avoid distortions in the extraction of features.

### 4.3 D3 dataset results

Reciprocally to Sect. 4.2, in the case of D3 dataset, we have a summary of TSFEL results in Table 16 and the summary of ROCKET results in Table 17. In either case, using TSFEL or ROCKET, in Tables 16 and 17, we are only showing the results obtained with the baseline (*v0*) version and with basic transformation versions, (*v1*) and (*v2*), plus the best result obtained with any of the quaternion fusion versions (see Table 6). In the same way than before, the upper half of the table (*segmented*) includes the summarised results obtained while analysing D3 dataset per partial movement. In

the lower half of the table (*complete*), we have the results obtained while analysing D3 dataset per martial artist, considering the complete movement.

For this D3 dataset and using TSFEL, the best results (80.%) for user (*martial artist*) modelling, when performing the *blocking set I* movement, analysed per partial movement (segmented), are obtained while using quaternion fusion (Madgwick) version with *Cartesian* coordinates (*v4*) and after applying *RF* as classifier (see Table 7). The best results for the complete movement are obtained while using many dataset versions, including baseline (*v0*), basic transformations, (*v1*) and (*v2*), and some quaternion fusions.

In the case of using ROCKET in this D3 dataset, the best results (80.%) for user (*martial artist*) modelling, when performing the *blocking set I* movement, analysed per segmented movements, are obtained while using quaternion fusion (Analytical) version with *cylindrical* coordinates (*v11*) and after applying *RF* and *LR* as classifiers (see Table 7). Analysing the complete movement, we obtained the best results while using quaternion fusion (Kalman) version with *cylindrical* coordinates (*v14*) and after applying *RF* as classifier (see Table 7). For this D3 dataset, the improvement of using *quaternions* is notorious, specifically when using ROCKET method and either when analysing the segmented and the complete movement, achieving results that are between 40. and 50.% better than the baseline or the basic transformation versions, see Table 17.

#### 4.4 Results findings and discussion

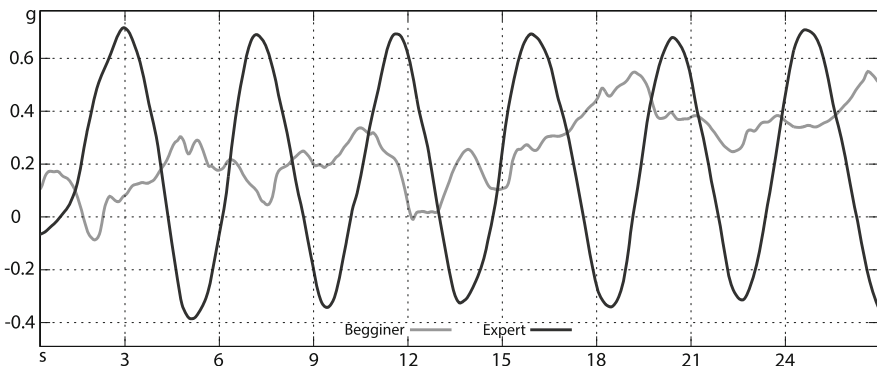
In Table 18 we include a summary the most representative findings obtained after analysing the results of Sects. 4.1, 4.2 and 4.3.

Despite the diverse nature of the different datasets used in this research, following the results introduced in Sects. 4.1, 4.2 and 4.3 we initially may conclude that, especially in the case of using complex datasets, such D2 and D3, we can model users by automatically classifying martial artists between experts and beginners through datasets built with inertial information. As expected, expert martial artists in D2 and D3, tend to fulfil the movements smoothly, showing autonomous behaviour in their performance. In Fig. 12, we have the signals gathered while an expert and a beginner are performing one of the *shikko* movement phases. As we can appreciate, beginner movements are erratic and less predictable.

In the same way, and following (Santos 2016), we may conclude that the learning experience in the psychomotor domain is not explained through the conscious knowledge of the discipline but associated with physical skills related to manual tasks and physical movements, acquired through the experience. In the case of D1, the basic swing (*shōmen*) movement performed with the *bokken*, the analysis is not so conclusive. We may explain this because the performance level in this movement is often difficult to appreciate, in addition to the fact that the sensor used in this dataset (see Table 3), only registered 3 dimensions and its sensitivity was also much lower compared with the one gathered in the other datasets (see Table 4). Thus, while in D2 and D3, over 80.% of the times we achieved to distinguish between beginner or expert practitioner, in D1, although we could not reach 70.% of accuracy, we could slightly distinguish between beginner or expert too.

**Table 18** Most relevant result findings—summary

1	Inertial datasets with 3 dimensions (accelerometer only) are limited for expertise level assessment
2	When using inertial datasets with 3 dimensions (accelerometer only), classification results considering the feature map extracted from random convolutional kernels (ROCKET) are better than when using the statistical features extracted from raw time series (TSFEL)
3	In inertial datasets with 6 dimensions (accelerometer and gyroscope information), the quaternion fusion does not provides any improvement in the classification accuracy due the degraded fusion
4	Quaternion fusion appears to be promising for the classification of eminently circular movements (turns). Spherical coordinates represent better the movement with constant and repeating patterns, even in datasets with only 6 dimensions (accelerometer and gyroscope)
5	When having 6 or more dimensions (accelerometer, gyroscope and/or magnetometer), the classification analysis of the complete movement versus the segmented ones, as depicted in Fig. 5, provides better results in any case
6	In inertial datasets with 9 dimensions (accelerometer, gyroscope and magnetometer), quaternion fusion provides better results in either case, when using TSFEL or ROCKET and especially when considering the segmented movements
7	The number of samples is crucial for improving the accuracy of the classification
8	Classification algorithm results may vary according to the nature of the movement, as well as to the features used

**Fig. 12** Fourth component of the Cartesian-based quaternion ( $q_3$ ) for two Aikidō practitioners (beginner and expert) while performing the first (of the four) straight lap of the *shikko-knee-walking* exercise



In this sense, all datasets used in this research include practitioners from 19 to 69 years old, for the datasets D1 and D2, and from 21 to 65 years old in the case of dataset D3, and consequently the proposed analysis demonstrates how we can infer performance level independently of other physical parameters, such age (Voelcker-Rehage and Willimczik 2006; Voelcker-Rehage 2008). The relevance of this is that our AI-driven analysis mainly considers mastering psychomotor skills as a gradual process where execution improvements comes from the experience, as discussed elsewhere (Santos 2019). Thus, improving belt rank means acquiring more experience, independently of other external factors.

With the proposed analysis, we fulfil the basis for personalising tangible psychomotor learning support, as defined in Santos (2016): sensing the learner's corporal movement and comparing this against the accurate movement. Thus, once the information collected has been analysed, we can compare beginner movement against the expert movement and decide whether it is appropriate or not to provide the tangible support.

Generally speaking, in the case of D2 dataset, we can assess expertise level with an accuracy of 78.57%, while using TSFEL, and with up to 82.14%, in the case of using ROCKET. In the case of D3 dataset, we reach an accuracy of 80.% using both methods, TSFEL and ROCKET.

In relation to D2 dataset, as the movement registered was basically rotational, transforms made from raw inertial data (*Cartesian*) to other coordinate systems (*spherical* and *cylindrical*) demonstrated, as expected, an improvement inferring performance level. In this sense, we may distinguish between the two approach methods used, comparing the results obtained with TSFEL and ROCKET. While using ROCKET, the highest accuracy (82.14%) is obtained using *spherical* coordinates during the whole *shikko* movement. On the contrary, with TSFEL we barely match the highest accuracy obtained with the raw inertial data collected: 75.% versus 78.57%. However, analysing TSFEL's results for this *shikko* movement separately per phases (Table 12, Table 13 and Table 14), we always reach better accuracies with *spherical* coordinates (goings and returns) or with quaternions (turns). This entails that both methods, TSFEL and ROCKET, work better with transformations (*spherical*) or fusions (quaternions) than with the baseline version. In addition, the extracting statistical domain features (TSFEL) are more adequate for analysing movement phases and convolutional kernels (ROCKET) are better analysing the whole movement. Using TSFEL in the most rotational *shikko* movement phase, the turns, we achieved the best accuracy result using quaternions (see Table 14). This is consistent with what was stated above, although we need to take into account that, unfortunately for this D2 dataset, magnetometer data was missing, so quaternion fusion was not as accurate as the one obtained in D3.

Regarding the D3 dataset, both methods, ROCKET or TSFEL, allowed better accuracy results using quaternions, as evinced in Tables 16 and 17. As stated in Sect. 3.1.3, the American Kenpo setup also comprised magnetometer information, which has resulted in a more accurate quaternion derivation. In this case, due we have a small number of practitioners (16), we estimated the analysis to be more realistic in the case of considering any of the six different movements registered separately, as phases for a total number of samples of 96. The results obtained show us the usefulness of using quaternions. In this case, as D3's movements were not as rotational as the movement

included in D2, the results using *spherical* or *cylindrical* were similar (ROCKET) or slightly worse (TSFEL) than those obtained in baseline (*Cartesian*).

Only in the case of using TSFEL, and related with the most representative features obtained after LASSO reduction, in the case of D2, together with some of the different histograms computed from the collected signals, we also found relevant the ECDF computation (Barandas et al. 2020). In the other hand and in the case of D3, the best quaternion accuracy was also with two of the different histograms computed. Thus, in both datasets, D2 and D3, the most relevant TSFEL feature used for distinguish between expert and beginner is the histogram, understood as the relative frequency which is equal to the frequency for an observed data value divided by the total number of data values in the sample.

We also confirmed that the main differences between TSFEL and ROCKET, is the time consumed in their executions: while ROCKET takes minutes, TSFEL extract features in seconds. As expected, although ROCKET is considerable faster than the mentioned CNN methods, its completion time it is not comparable with TSFEL, which is much faster. Thus, as the TSFEL approach is faster than ROCKET, this entails that it might be easy to implement into small devices due its low computer complexity (embedded devices).

Finally, the above is still valid for analysing specific behaviours that are not directly related with expertise level assessment, such fatigue monitoring or emotional level evaluation. As disclosed in Sects. 3.1.1 and 3.1.2, throughout the execution of the movement, we can observe how performance evolves, being able to appreciate certain symptoms of fatigue. Flexibility, as well as Joint Range of Motion (JROM) assessments (Thorpe et al. 2017), are elements that can be analysed, gathering for instance inertial information, to monitor fatigue in sports. Moreover, physical and mental fatigues are related, not only from the point of view of physical performance deterioration, but also via changes in technique execution (Russell et al. 2019). Thus, technical impairments can also be detected analysing inertial information in order to monitor mental fatigue. The foregoing can be directly extrapolated to the analysis of emotional factors and its relationship with psychomotor behaviours (Avalos et al. 2022).

#### 4.5 Limitations and future works

Although D2 is a very detailed and complete dataset to perform our research, the main limitation of this is the lack of information coming from a magnetometer sensor. Having these 3 more coordinates, we could have improved the quaternion fusion as confirmed with D3 dataset, the only dataset used with real magnetometer data. This highlights the need to include enriched inertial data to model in a more precise and accurate way the performances of the practitioners. Furthermore, we can add more devices and thus more sensors to improve the assessment.

In this case, D3 dataset, the number of martial artists included was only 16 and, although we performed some data augmentation, splitting every *blocking set I* movement into six other different movements, the reduced number of participants was also a limitation of this research. In the same way, it is essential to be able to orchestrate experiences to allow the collection of a greater amount of data. What is exposed in

this article, like most of the researches, would be benefited from the inclusion of a greater number of participants.

Despite the intrinsic disclosed limitations derived from the use of other capture devices for the purpose of solving the research question, such as video images (Tee et al. 2022; Ige and Mohd Noor 2022; Qiu et al. 2022; Gupta et al. 2022; Saha et al. 2022; Pereira and Gonçalves 2022; Babangida et al. 2022; Lateef and Abbas 2022), our research could have benefited from a multi modal and multi sensor approach, incorporating video images into the inertial analysis. To resolve this eventuality and to include video images to support inertial analysis, in future researches we will follow similar approaches to those considered with the iBAID (intelligent Basket AID) system (Portaz et al. 2023) where video images are used to support inertial analysis by providing a human-centred approach with the aim of recommending the physical activities and movements to perform when training in basketball, either to improve technique, to recover from an injury or even to stay active when ageing.

The current approach of analysing movement practice individually does have its limitations, particularly in scenarios where collaborative psychomotor efforts are essential. However, in these cases, and as outlined in Echeverria and Santos (2021), there exists a promising avenue for enhancement. By introducing an intermediate level between the existing Level 1 and Level 2 in the suggested hierarchical ontology (as shown in Fig. 1), we can effectively address whether movements are executed individually or within a collaborative scenario. This addition to the hierarchy not only expands the scope of applicability but also ensures a more comprehensive and adaptable approach to movement practice. It holds the potential to significantly enrich the learning experience and outcomes, particularly in contexts where collective coordination is paramount.

Another limitation is the current gender distribution, which shows a clear dominance of male martial artists, depending of the dataset, the percentage of female participants vary from 12.5% (D3) to 13.63% (D1) and 14.59% (D2). In order to fulfil one of the aims of this research, which is to develop psychomotor learning systems that favour gender equality and social inclusion, we should put more efforts into favouring more diverse data collection in the future.

## 5 Conclusions

This work demonstrates a promising correlation between the inertial sensor data gathered by inertial systems and the modelling of users' expertise levels during the acquisition of psychomotor skills. The findings suggest that such technology holds potential in providing valuable insights into skill acquisition and proficiency assessment. This novel approach not only enhances our understanding of learning processes but also opens new avenues for personalised and adaptive training methodologies. Further research and refinement in this area have the potential to revolutionise skill acquisition across various domains, benefiting both beginners and experts alike.

As an overview, we have tested whether it was possible to classify martial artists according to their expertise level and by measuring very specific movements or gestures, solving affirmatively the initial research question. These measurements have

been carried out just with simple inertial sensors (accelerometers, gyros and magnetometers) attached to a single spot on the practitioner's body or martial instrument (sword). Two martial arts have been researched as a proof-of-concept: Aikidō and American Kenpo Karate. The logged data streams have been subsequently transformed and prepared by several means prior to be fed into state-of-the-art classification algorithms. These transformations have comprised coordinate transformations, quaternion fusion, feature extraction and convolutional neural networks, applied to multivariate time series data.

Thus, the method proposed in this research to model users in the frame of psychomotor learning system, uses martial art datasets with transformed and fused inertial raw data, which was collected with magnetic and non-magnetic IMUs and applying different approaches for the classification of the discriminatory features extracted. One of the approaches has been chosen due it represents the current state-of-the-art for time series classification (i.e. TSFEL library), while the other approach used in this research has been specifically conceived for human activity recognition and MTS (i.e. ROCKET framework for CNN).

Results have shown that it is possible to achieve a  $> 82\%$  classification success rate. We also demonstrated how transforming raw data into spherical or cylindrical coordinates can improve the accuracy in the classification, overall if the studied movement includes rotational behaviour. The application of quaternions has also shown to be adequate in this setup (above all when magnetometer information is available, as it has been the case of American Kenpo Karate dataset).

Several other final conclusions may be taken into consideration regarding the method used. Thus, while the extraction of features (performed using the TSFEL library) may work better when analysing a unique movement or phase, ROCKET may produce better results when the analysis includes a complete movement or a joint sequence of movements.

**Acknowledgements** The research carried out is framed in the projects “INT2AFF: INTelligent INtra-subject development approach to improve actions in AFFectaware adaptive educational systems” funded by the Spanish Ministry of Science, Innovation and Universities, the Spanish Agency of Research and the European Regional Development Fund (ERDF) under Grant PGC2018-102279-B-I00, specifically within the scenario “SC4-motor skills” and “HUMANAIID- Sens: HUMAN-centered Assisted Intelligent Dynamic systems with SENSing technologies” funded by MCIN/AEI/10.13039/501100011033, and the European Union “NextGenerationEU”/PRTR under grant TED2021-129485B-C41. In addition, the authors would also like to thank all participants involved in the data collection, for both the Aikidō and American Kenpo Karate datasets. For the Aikidō dataset, they are the following: Judelda (Elda), Koen (El Campello), Aikidō Shogun (San Vicente), Vinalopó Budo (Elda), Aikidō Budo (Elche), Aikidō-iaido (Alicante), Aikidō Club Alicante (Alicante), Aikidō Daimyo (Alicante), Spa Columnata (Logroño), Aikidō Universidad de Alicante/Fight Team Crespo (San Vicente), Calasancio (Alicante), Kime (Madrid), Tomás Sánchez (Madrid), Zenrai (Murcia), Centro de Tecnificación (Alicante), Sintagma (Valencia), CEIP Alfonso X el Sabio (Novelda) and Universidad Autónoma de Madrid (Madrid). For the Kenpo Karate dataset, they are the following: Kenpo Karate Estudio (Madrid). The work performed by Alberto Corbi is partially funded by Universidad Internacional de la Rioja (UNIR, <http://transfer.unir.net>) through the Research Institute for Innovation & Technology in Education (UNIR iTED, <http://ited.unir.net>). He also would especially like to thank the Judelda and Sintagma dojos for giving permission to take some of the photos included in this paper.

**Author Contributions** Miguel Portaz contributed to conceptualisation, methodology, software, validation, formal analysis, investigation, writing—original draft, and visualisation; Alberto Corbi contributed to conceptualisation, investigation, methodology, supervision, resources, data curation, and writing—review editing; Alberto Casas-Ortiz contributed to resources and writing—review editing; Olga C. Santos con-

tributed to conceptualisation, investigation, methodology, supervision, resources, writing—review editing, project administration, and funding acquisition

**Funding** Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

## Declarations

**Conflict of interest** The authors declare no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Adams, J.A.: A closed-loop theory of motor learning. *J. Mot. Behav.* **3**(2), 111–150 (1971). <https://doi.org/10.1080/00222895.1971.10734898>
- Ariza-Colpas, P.P., Vicario, E., Oviedo-Carrascal, A.I., Aziz, S.B., Piñeres-Melo, M.A., Quintero-Linero, A., Patara, F.: Human activity recognition data analysis: history, evolutions, and new trends. *Sensors* **22**(9), 3401 (2022). <https://doi.org/10.3390/s22093401>
- Avalos, M., Zurita, F., Cachon, J.: Analysis of the psychomotor profile, learning difficulty and emotional level in preschoolers. *J. Hum. Sport Exercise* **17**(2), 389–399 (2022). <https://doi.org/10.14198/jhse.2022.172.14>
- Avci, A., Bosch, S., Marin-Perianu, M., Marin-Perianu, R., Havinga, P.: Activity recognition using inertial sensing for healthcare, wellbeing and sports applications: a survey, 11 (2010)
- Babangida, L., Perumal, T., Mustapha, N., Yaakob, R.: Internet of things (IoT) based activity recognition strategies in smart homes: a review. *IEEE Sens. J.* **22**(9), 8327–8336 (2022). <https://doi.org/10.1109/jsen.2022.3161797>
- Bagnall, A., Lines, J., Bostrom, A., Large, J., Keogh, E.: The great time series classification bake off: a review and experimental evaluation of recent algorithmic advances. *Data Min. Knowl. Disc.* **31**(3), 606–660 (2016). <https://doi.org/10.1007/s10618-016-0483-9>
- Barandas, M., Folgado, D., Fernandes, L., Santos, S., Abreu, M., Bota, P., Liu, H., Schultz, T., Gamboa, H.: TSFEL: time series feature extraction library. *SoftwareX* **11**, 100456 (2020). <https://doi.org/10.1016/j.softx.2020.100456>
- Benson, L.C., Clermont, C.A., Osis, S.T., Kobsar, D., Ferber, R.: Classifying running speed conditions using a single wearable sensor: optimal segmentation and feature extraction methods **71**, 94–99. <https://doi.org/10.1016/j.jbiomech.2018.01.034>
- Bergamini, E., Ligorio, G., Summa, A., Vannozzi, G., Cappozzo, A., Sabatini, A.: Estimating orientation using magnetic and inertial sensors and different sensor fusion approaches: accuracy assessment in manual and locomotion tasks. *Sensors* **14**(10), 18625–18649 (2014). <https://doi.org/10.3390/s141018625>
- Camomilla, V., Bergamini, E., Fantozzi, S., Vannozzi, G.: Trends supporting the in-field use of wearable inertial sensors for sport performance evaluation: a systematic review. *Sensors* **18**(3) (2018). <https://doi.org/10.3390/s18030873>
- Camp, N., Lewis, M., Hunter, K., Johnston, J., Zecca, M., Di Nuovo, A., Magistro, D.: Technology used to recognize activities of daily living in community-dwelling older adults. *Int. J. Environ. Res. Public Health* **18**(1) (2021). <https://doi.org/10.3390/ijerph18010163>
- Casas-Ortiz, A., Echeverria, J., Santos, O.C.: Chapter 18: Intelligent systems for psychomotor learning: a systematic review and two cases of study. *Handbook of Artificial Intelligence in Education*; du Boulay,

- B., Mitrovic, A., Yacef, K., Eds. Edward Edgar Publishing: Northampton, MA, USA, 390–421 (2023). <https://doi.org/10.4337/9781800375413.00030>
- Casas-Ortiz, A., Santos, O.C. (supervisor): Capturing, modelling, analyzing and providing feedback in martial arts with artificial intelligence to support psychomotor learning activities. UNED Master's thesis (2021). Retrieved October 1, 2023 from <http://www.ia.uned.es/docencia/posgrado/master-tfm-archivo.html>
- Casas-Ortiz, A., Santos, O.C.: AIED—ksas: An AI application to learn martial arts movements in on-line settings. (2021). Retrieved October 1, 2023 from <https://iaied.org/showcase/27NDIsN0KwloFfzXG7CD2>
- Choi, A., Jung, H., Mun, J.H.: Single inertial sensor-based neural networks to estimate COM-COP inclination angle during walking. *Sensors* **19**(13), 2974 (2019). <https://doi.org/10.3390/s19132974>
- Corbí, A., Santos, O.C.: MyShikko: modelling knee walking in aikido practice. In: 2018 ACM Adjunct Publication of the 26th Conference on User Modeling, Adaptation and Personalization, pp. 217–218 (2018). <https://doi.org/10.1145/3213586.3225225>
- Corbí, A., Santos, O.C., Burgos, D.: Intelligent framework for learning physics with aikido (martial art) and registered sensors. *Sensors* **19**(17), 3681 (2019). <https://doi.org/10.3390/s19173681>
- Cox, J.C.: Traditional asian martial arts training: a review. *Quest* **45**(3), 366–388 (1993). <https://doi.org/10.1080/00336297.1993.10484094>
- Dallel, M., Havard, V., Dupuis, Y., Baudry, D.: Digital twin of an industrial workstation: a novel method of an auto-labeled data generator using virtual reality for human action recognition in the context of human-robot collaboration. *Eng. Appl. Artif. Intell.* **118**, 105655 (2023). <https://doi.org/10.1016/j.engappai.2022.105655>
- Debes, C., Merentitis, A., Sukhanov, S., Niessen, M., Frangiadakis, N., Bauer, A.: Monitoring activities of daily living in smart homes: understanding human behavior. *IEEE Signal Proces. Mag.* **33**(2), 81–94 (2016). <https://doi.org/10.1109/MSP.2015.2503881>
- del Rosario, M., Redmond, S., Lovell, N.: Tracking the evolution of smartphone sensing for monitoring human movement. *Sensors* **15**(8), 18901–18933 (2015). <https://doi.org/10.3390/s150818901>
- Dempster, A., Petitjean, F., Webb, G.I.: ROCKET: exceptionally fast and accurate time series classification using random convolutional kernels. *Data Min. Knowl. Disc.* **34**(5), 1454–1495 (2020). <https://doi.org/10.1007/s10618-020-00701-z>
- Echeverria, J., Santos, O.C.: Toward modeling psychomotor performance in karate combats using computer vision pose estimation. *Sensors* **21**(24) (2021). <https://doi.org/10.3390/s21248378>
- Ehatisham-Ul-Haq, M., Azam, M.A., Amin, Y., Naeem, U.: C2fhar: coarse-to-fine human activity recognition with behavioral context modeling using smart inertial sensors. *IEEE Access* **8**, 7731–7747 (2020). <https://doi.org/10.1109/ACCESS.2020.2964237>
- Fawaz, H.I., Forestier, G., Weber, J., Idoumghar, L., Muller, P.-A.: Springer science and business media—deep learning for time series classification: a review. *Data Min. Knowl. Disc.* **33**(4), 917–963 (2019). <https://doi.org/10.1007/s10618-019-00619-1>
- Fawaz, H.I., Lucas, B., Forestier, G., Pelletier, C., Schmidt, D.F., Weber, J., Webb, G.I., Idoumghar, L., Muller, P.-A., Petitjean, F.: InceptionTime: finding AlexNet for time series classification. *Data Min. Knowl. Disc.* **34**(6), 1936–1962 (2020). <https://doi.org/10.1007/s10618-020-00710-y>
- Ferscha, A., Mattern, F. (eds.): *Pervasive Computing: Second International Conference, Pervasive 2004, Linz/Vienna, Austria, April 18–23, 2004: Proceedings. Lecture notes in computer science, vol. 3001.* Springer. Meeting Name: Pervasive (2004)
- Fitts, P.M., Posner, M.I.: *Human Performance* (1967). Human performance. Brooks/Cole
- Fresk, E., Nikolakopoulos, G.: Full quaternion based attitude control for a quadrotor. In: 2013 European Control Conference (ECC), pp. 3864–3869 (2013). <https://doi.org/10.23919/ECC.2013.6669617>
- Fu, B., Damer, N., Kirchbuchner, F., Kuijper, A.: Sensing technology for human activity recognition: a comprehensive survey. *IEEE Access* **8**, 83791–83820 (2020). <https://doi.org/10.1109/ACCESS.2020.2991891>
- Gagne, R.M., Briggs, L.J., Wager, W.W.: *Principles of Instructional Design* (1992), 4th edn. Harcourt Brace Jovanovich College Publishers
- Glowinski, S., Blazejewski, A., Krzyzynski, T.: Inertial sensors and wavelets analysis as a tool for pathological gait identification. In: Springer International Publishing, *Innovations in Biomedical Engineering*, pp. 106–114 (2016). [https://doi.org/10.1007/978-3-319-47154-9\\_13](https://doi.org/10.1007/978-3-319-47154-9_13)
- Goodfellow, I., Bengio, Y., Courville, A.: *Deep Learning*. MIT Press (2016). Retrieved October 1, 2023 from <http://www.deeplearningbook.org>

- Guennec, A.L., Malinowski, S., Tavenard, R.: Data augmentation for time series classification using convolutional neural networks (2016)
- Guerrero-Sánchez, M.E., Abaunza, H., Castillo, P., Lozano, R., García-Beltrán, C.D.: Quadrotor energy-based control laws: a unit-quaternion approach. *J. Intell. Robot. Syst.* **88**(2–4), 347–377 (2017). <https://doi.org/10.1007/s10846-017-0528-3>
- Guglielmo, E.: The ageing brain neuroplasticity and lifelong learning. *eLearning Papers* **29**, 1–7 (2012)
- Guo, H., Hong, H.: Research on filtering algorithm of MEMS gyroscope based on information fusion. *Sensors* **19**(16), 3552 (2019). <https://doi.org/10.3390/s19163552>
- Gupta, N., Gupta, S.K., Pathak, R.K., Jain, V., Rashidi, P., Suri, J.S.: Human activity recognition in artificial intelligence framework: a narrative review. *Artif. Intell. Rev.* **55**(6), 4755–4808 (2022). <https://doi.org/10.1007/s10462-021-10116-x>
- Hartmann, Y., Liu, H., Schultz, T.: Feature space reduction for human activity recognition based on multi-channel biosignals. In: SCITEPRESS Proceedings of the 14th International Joint Conference on Biomedical Engineering Systems and Technologies (2021). <https://doi.org/10.5220/0010260802150222>
- Haslwanter, T.: skinematics (2020). Retrieved October 1, 2023, from <http://work.thaslwanter.at/skinematics/html/>
- Heinz, E.A., Kunze, K.S., Gruber, M., Bannach, D., Lukowicz, P.: Using wearable sensors for real-time recognition tasks in games of martial arts—an initial experiment. In: 2006 IEEE Symposium on Computational Intelligence and Games, pp. 98–102. IEEE. <https://doi.org/10.1109/CIG.2006.311687>
- Homma, G.: No suwariwaza (kneeling techniques) on nippon kan, 245–254 (2007)
- Hopfgartner, F., Kay, J., Rapp, A.: Preface to the special issue on harnessing personal tracking data for personalization and sense-making. *User Model. User-Adap. Inter.* **30**(3), 323–329 (2020). <https://doi.org/10.1007/s11257-020-09271-7>
- Huang, J., Cao, Y., Wang, Y.-W.: Adaptive proxy-based sliding mode control for a class of second-order nonlinear systems and its application to pneumatic muscle actuators. *ISA Trans.* **124**, 395–402 (2022). <https://doi.org/10.1016/j.isatra.2020.09.009>
- Ige, A.O., Mohd Noor, M.H.: A survey on unsupervised learning for wearable sensor-based activity recognition. *Appl. Soft Comput.* **127**, 109363 (2022). <https://doi.org/10.1016/j.asoc.2022.109363>
- Jafari, R., Li, W., Bajcsy, R., Glaser, S., Sastry, S.: Physical activity monitoring for assisted living at home. In: 4th International Workshop on Wearable and Implantable Body Sensor Networks (BSN 2007), Springer Berlin Heidelberg, pp. 213–219 (2007). [https://doi.org/10.1007/978-3-540-70994-7\\_37](https://doi.org/10.1007/978-3-540-70994-7_37)
- James, D.A., Uroda, W., Gibson, T.: Dynamics of swing: a study of classical Japanese Swordsmanship Using Accelerometers, 7 (2014)
- Kaňtoch, E.: Recognition of sedentary behavior by machine learning analysis of wearable sensors during activities of daily living for telemedical assessment of cardiovascular risk. *Sensors* **18**(10) (2018). <https://doi.org/10.3390/s18103219>
- Kim, J.-K., Lee, K.B., Kim, J.-C., Hong, S.G.: Patient identification based on physical rehabilitation movements using skeleton data. In: 2021 International Conference on Information and Communication Technology Convergence (ICTC), pp. 1572–1574 (2021). <https://doi.org/10.1109/ICTC52510.2021.9621049>
- Kos, A., Tomažič, S., Umek, A.: Suitability of smartphone inertial sensors for real-time biofeedback applications. *Sensors* **16**(3), 301 (2016). <https://doi.org/10.3390/s16030301>
- Krüger, A., Edelmann-Nusser, J.: Biomechanical analysis in freestyle snowboarding: application of a full-body inertial measurement system and a bilateral insole measurement system. *Sports Technol.* **2**(1–2), 17–23 (2009)
- Kunze, K., Barry, M., Heinz, E.A., Lukowicz, P., Majoe, D., Gutknecht, J.: Towards recognizing tai chi: an initial experiment using wearable sensors. In: 3rd International Forum on Applied Wearable Computing 2006, pp. 1–6 (2006)
- Lateef, R.A., Abbas, A.R.: Human activity recognition using smartwatch and smartphone: a review on methods, applications, and challenges. *Iraqi J. Sci.*, 363–379 (2022). <https://doi.org/10.24996/ij.s.2022.63.1.34>
- Liu, P., Wang, H., Zheng, S., Zhang, F., Zhang, X.: Parkinson’s disease diagnosis using neostriatum radiomic features based on t2-weighted magnetic resonance imaging. *Front. Neurol.* **11** (2020). <https://doi.org/10.3389/fneur.2020.00248>
- Ludwig, S., Burnham, K., Jimenez, A., Touma, P.: Comparison of attitude and heading reference systems using foot mounted MIMU sensor data: basic, madgwick, and mahony. In: Sohn, H. (ed.) *SPIE Sensors*

- and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2018 (2018). <https://doi.org/10.1117/12.2296568>
- Lukosevicius, M., Popovici, D., Jaeger, H., Siewert, U.: Time Warping Invariant Echo State Networks. Retrieved June 15, 2021, from [https://opus.jacobs-university.de/files/660/02\\_twiesn\\_iubtechreport.pdf](https://opus.jacobs-university.de/files/660/02_twiesn_iubtechreport.pdf) (2006)
- Madgwick, S.O.H., Harrison, A.J.L., Vaidyanathan, R.: IEEE—estimation of IMU and MARG orientation using a gradient descent algorithm. In: 2011 IEEE International Conference on Rehabilitation Robotics (2011). <https://doi.org/10.1109/icorr.2011.5975346>
- Magill, R., Anderson, D.I.: *Motor Learning and Control*. McGraw-Hill Publishing New York (2010)
- Mahony, R., Hamel, T., Pfimlin, J.-M.: Nonlinear complementary filters on the special orthogonal group. *IEEE Trans. Autom. Control* **53**(5), 1203–1218 (2008). <https://doi.org/10.1109/tac.2008.923738>
- Mannini, A., Sabatini, A.M.: Machine learning methods for classifying human physical activity from on-body accelerometers. *Sensors* **10**(2), 1154–1175 (2010). <https://doi.org/10.3390/s100201154>
- Mason, R., Pearson, L.T., Barry, G., Young, F., Lennon, O., Godfrey, A., Stuart, S.: Wearables for running gait analysis: a systematic review. *Sports Med.* **53**(1), 241–268 (2022). <https://doi.org/10.1007/s40279-022-01760-6>
- Meng, Z., Zhang, M., Guo, C., Fan, Q., Zhang, H., Gao, N., Zhang, Z.: Recent progress in sensing and computing techniques for human activity recognition and motion analysis. *Electronics* **9**(9) (2020). <https://doi.org/10.3390/electronics9091357>
- Mills, D.L.: *Computer Network Time Synchronization: the Network Time Protocol* (2006). CRC/Taylor & Francis
- Mills, D.L.: Network time protocol (NTP). Technical report, RFC Editor (1985). <https://doi.org/10.17487/rfc0958>
- Miranda, L., Viterbo, J., Bernardini, F.: A survey on the use of machine learning methods in context-aware middlewares for human activity recognition. *Artif. Intell. Rev.* **55**(4), 3369–3400 (2022). <https://doi.org/10.1007/s10462-021-10094-0>
- Najafi, B., Lee-Eng, J., Wrobel, J.S., Goebel, R.: Estimation of center of mass trajectory using wearable sensors during golf swing. *J. Sports Sci. Med.* **14**(2), 354 (2015)
- Newell, K.M.: What are fundamental motor skills and what is fundamental about them? *J. Motor Learn. Dev.* **8**(2), 280–314 (2020). <https://doi.org/10.1123/jmld.2020-0013>
- Olugbade, T., Bienkiewicz, M., Barbareschi, G., D'amato, V., Oneto, L., Camurri, A., Holloway, C., Björkman, M., Keller, P., Clayton, M., Williams, A.C.D.C., Gold, N., Becchio, C., Bardy, B., Bianchi-Berthouze, N.: Human movement datasets: an interdisciplinary scoping review. *ACM Comput. Surv.* **55**(6) (2022). <https://doi.org/10.1145/3534970>
- Paraschiakos, S., Cachucho, R., Moed, M., van Heemst, D., Mooijaart, S., Slagboom, E.P., Knobbe, A., Beekman, M.: Activity recognition using wearable sensors for tracking the elderly. *User Model. User-Adap. Inter.* **30**(3), 567–605 (2020). <https://doi.org/10.1007/s11257-020-09268-2>
- Parker, E.: *Infinite Insights Into Kenpo*. BookSurge Publishing (2009)
- Pereira, J.G., Gonçalves, J.: Human activity recognition: A review. In: 2022 IEEE 10th International Symposium on Digital Forensics and Security (ISDFS) (2022). <https://doi.org/10.1109/isdfs55398.2022.9800781>
- Portaz, M., Manjarrés, A., Santos, O.C.: Towards human-centric psychomotor recommender systems. In: UMAP '23 Adjunct: Adjunct Proceedings of the 31st ACM Conference on User Modeling, Adaptation and Personalization (UMAP '23 Adjunct), 6 (2023). <https://doi.org/10.1145/3563359.3596993>
- Qiu, S., Zhao, H., Jiang, N., Wang, Z., Liu, L., An, Y., Zhao, H., Miao, X., Liu, R., Fortino, G.: Multi-sensor information fusion based on machine learning for real applications in human activity recognition: state-of-the-art and research challenges. *Inf. Fusion* **80**, 241–265 (2022). <https://doi.org/10.1016/j.inffus.2021.11.006>
- Réby, K., Dulau, I., Dubrasquet, G., Aïmar, M.B.: Graph transformer for physical rehabilitation evaluation. In: 2023 IEEE 17th International Conference on Automatic Face and Gesture Recognition (FG), pp. 1–8 (2023). <https://doi.org/10.1109/FG57933.2023.10042778>
- Renaudin, V., Combettes, C.: Magnetic, acceleration fields and gyroscope quaternion (MAGYQ)-based attitude estimation with smartphone sensors for indoor pedestrian navigation. *Sensors* **14**(12), 22864–22890 (2014). <https://doi.org/10.3390/s141222864>
- Russell, S., Jenkins, D., Smith, M., Halson, S., Kelly, V.: The application of mental fatigue research to elite team sport performance: new perspectives. *J. Sci. Med. Sport* **22**(6), 723–728 (2019). <https://doi.org/10.1016/j.jsams.2018.12.008>



- Sabatini, A.M.: Quaternion-based strap-down integration method for applications of inertial sensing to gait analysis. *Med. Biol. Eng. Comput.* **43**(1), 94–101 (2005). <https://doi.org/10.1007/bf02345128>
- Sabatini, A.M.: Quaternion-based extended kalman filter for determining orientation by inertial and magnetic sensing. *IEEE Trans. Biomed. Eng.* **53**(7), 1346–1356 (2006). <https://doi.org/10.1109/TBME.2006.875664>
- Sabatini, A.M.: Estimating three-dimensional orientation of human body parts by inertial/magnetic sensing. *Sensors* **11**(2), 1489–1525 (2011). <https://doi.org/10.3390/s110201489>
- Saha, A., Rajak, S., Saha, J., Chowdhury, C.: A survey of machine learning and meta-heuristics approaches for sensor-based human activity recognition systems. *J. Ambient. Intell. Human. Comput.* (2022). <https://doi.org/10.1007/s12652-022-03870-5>
- Santos, O.C.: Beyond cognitive and affective issues: designing smart learning environments for psychomotor personalized learning (2016). In: Spector, M.J., Lockee, B.B., Childress, M.D. (eds.) *Learning, Design, and Technology: An International Compendium of Theory, Research, Practice, and Policy*, pp. 1–24. Springer. [https://doi.org/10.1007/978-3-319-17727-4\\_8-1](https://doi.org/10.1007/978-3-319-17727-4_8-1)
- Santos, O.C.: Psychomotor learning in martial arts. In: *ACM Adjunct Publication of the 25th Conference on UMAP User Modeling, Adaptation and Personalization* (2017). <https://doi.org/10.1145/3099023.3099107>
- Santos, O.C.: Training the body: the potential of AIED to support personalized motor skills learning. *Int. J. Artif. Intell. Educ.* **26**(2), 730–755 (2016). <https://doi.org/10.1007/s40593-016-0103-2>
- Santos, O.C.: Artificial intelligence in psychomotor learning: modeling human motion from inertial sensor data. *Int. J. Artif. Intell. Tools* **28**(4), 1940006 (2019). <https://doi.org/10.1142/S0218213019400062>
- Santos, O.C., Corbí, A.: Can aikido help with the comprehension of physics? a first step towards the design of intelligent psychomotor systems for steam kinesthetic learning scenarios. *IEEE Access* **7**, 176458–176469 (2019). <https://doi.org/10.1109/ACCESS.2019.2957947>
- Santos, O.C., Uria-Rivas, R., Rodriguez-Sanchez, M.C., Boticario, J.G.: An open sensing and acting platform for context-aware affective support in ambient intelligent educational settings. *IEEE Sens. J.* **16**(10), 3865–3874 (2016). <https://doi.org/10.1109/JSEN.2016.2533266>
- Saponas, S., Lester, J., Froehlich, J.E., Fogarty, J.A., Landay, J.A.: iLearn on the iPhone: Real-time human activity classification on commodity mobile phones. Technical Report UW-CSE-08-04-02, Microsoft. Retrieved October 1, 2023 from <https://www.microsoft.com/en-us/research/publication/ilearn-iphone-real-time-human-activity-classification-commodity-mobile-phones/>
- Schmidt, R.A., Lee, T.D.: *Motor Learning and Performance: from Principles to Application* (2014), Fifth edition edn. Human Kinetics
- Schmidt, R.A.: Control processes in motor skills. *Exercise and Sport Sciences Reviews* **4**(1), 229–262 (1976)
- Seitz, F.C., Olson, G.D., Stenzel, T.E.: A martial arts exploration of elbow anatomy: iikkyo/i (Aikido’s first teaching). *Perceptual and Motor Skills* **73**(3\_suppl), 1227–1234 (1991). <https://doi.org/10.2466/pms.1991.73.3f.1227>
- Shoemake, K.: Animating rotations with quaternion curves, pp. 245–254. *Proc. SIGGRAPH 85*, ACM Press (1985)
- Smyth, B., Lawlor, A., Berndsen, J., Feely, C.: Recommendations for marathon runners: on the application of recommender systems and machine learning to support recreational marathon runners. *User Model. User-Adap. Inter.* **32**(5), 787–838 (2021). <https://doi.org/10.1007/s11257-021-09299-3>
- Sung, Y., Guo, H., Lee, S.-G.: Motion quaternion-based motion estimation method of MYO using k-means algorithm and Bayesian probability. *Soft. Comput.* **22**(20), 6773–6783 (2018). <https://doi.org/10.1007/s00500-018-3379-3>
- Tee, W.Z., Dave, R., Seliya, J., Vanamala, M.: A close look into human activity recognition models using deep learning. In: *2022 IEEE 3rd International Conference on Computing, Networks and Internet of Things (CNIOT)* (2022). <https://doi.org/10.1109/cniot55862.2022.00043>
- Thorpe, R.T., Atkinson, G., Drust, B., Gregson, W.: Monitoring fatigue status in elite team-sport athletes: implications for practice. *Int. J. Sports Physiol. Perform.* **12**(s2), 2–27234 (2017). <https://doi.org/10.1123/ijsp.2016-0434>
- Tran, S.N., Nguyen, D., Ngo, T.-S., Vu, X.-S., Hoang, L., Zhang, Q., Karunanithi, M.: On multi-resident activity recognition in ambient smart-homes. *Artif. Intell. Rev.* **53**(6), 3929–3945 (2019). <https://doi.org/10.1007/s10462-019-09783-8>
- Voelcker-Rehage, C.: Motor-skill learning in older adults: a review of studies on age-related differences. *Eur. Rev. Aging Phys. Activity* **5**(1), 5–16 (2008). <https://doi.org/10.1007/s11556-008-0030-9>

- Voelcker-Rehage, C., Willimczik, K.: Motor plasticity in a juggling task in older adults: a developmental study. *Age Ageing* **35**(4), 422–427 (2006). <https://doi.org/10.1093/ageing/af025>
- Wang, W., Miao, C.: Multi-resident activity recognition with unseen classes in smart homes. In: 2018 IEEE SmartWorld, Ubiquitous Intelligence and Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCCom/IOP/SCI) (2018). <https://doi.org/10.1109/smartworld.2018.00147>
- Wang, Z., Yan, W., Oates, T.: Time series classification from scratch with deep neural networks: a strong baseline (2016) [arxiv: 1611.06455](https://arxiv.org/abs/1611.06455)
- Wartmann, C.: *The Blender Book* (2001). Linux Journal Press
- Xing, H., Chen, Z., Wang, C., Guo, M., Zhang, R.: Quaternion-based complementary filter for aiding in the self-alignment of the mems imu. In: 2019 IEEE International Symposium on Inertial Sensors and Systems (INERTIAL), pp. 1–4 (2019). <https://doi.org/10.1109/ISISS.2019.8739728>
- Yang, J.-Y., Wang, J.-S., Chen, Y.-P.: Using acceleration measurements for activity recognition: an effective learning algorithm for constructing neural classifiers. *Pattern Recognit. Lett.* **29**(16), 2213–2220 (2008). <https://doi.org/10.1016/j.patrec.2008.08.002>
- Yang, H., Lee, Y., Jeon, S.-Y., Lee, D.: Multi-rotor drone tutorial: systems, mechanics, control and state estimation. *Intell. Serv. Robot.* **10**(2), 79–93 (2017). <https://doi.org/10.1007/s11370-017-0224-y>
- Yuan, Y., Xun, G., Ma, F., Wang, Y., Du, N., Jia, K., Su, L., Zhang, A.: IEEE—MuVAN: a multi-view attention network for multivariate temporal data. In: 2018 IEEE International Conference on Data Mining (ICDM) (2018). <https://doi.org/10.1109/icdm.2018.00087>
- Zhang, S., Li, Y., Zhang, S., Shahabi, F., Xia, S., Deng, Y., Alshurafa, N.: Deep learning in human activity recognition with wearable sensors: a review on advances. *Sensors* **22**(4), 1476 (2022). <https://doi.org/10.3390/s22041476>
- Zhou, H., Hu, H., Tao, Y.: Inertial measurements of upper limb motion. *Med. Biol. Eng. Comput.* **44**(6), 479–487 (2006). <https://doi.org/10.1007/s11517-006-0063-z>
- Zhou, X., Liang, W., Wang, K.I.-K., Wang, H., Yang, L.T., Jin, Q.: Deep-learning-enhanced human activity recognition for internet of healthcare things. *IEEE Internet Things J.* **7**(7), 6429–6438 (2020). <https://doi.org/10.1109/jiot.2020.2985082>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Miguel Portaz** has held key roles in management, business development, and international sales for prominent technology multinational corporations. Holding a Master's Degree in Research in Artificial Intelligence, he is currently pursuing his PhD in Intelligent System at UNED and he is researcher in the HUMANAID project. Additionally, he holds positions as a Board Member and Chief of Staff of the PhyUM Research Center. His research is dedicated to the development of autonomous or semi-autonomous, adaptive, human-centered decision systems, with a specific focus on personalized psychomotor systems that integrate AI ethical principles.

**Dr. Alberto Corbi** is a senior researcher at the UNIR iTED (Research Institute for Innovation and Technology in Education) and as an assistant professor at the Engineering School, which are both part of the Universidad Internacional de La Rioja (UNIR). With a background in ocean-atmosphere interaction, he is currently involved in a variety of research fields: e-learning standards, systems interoperability, medical physics, radiological protection, STEM education, the monitoring of physical activities, the social implications of technology (with emphasis on social networks), e-health advancement (with emphasis on Alzheimer's disease characterization and clinical information standards) and environmentalism. He has published over 20 research papers on all these subjects, and he is a frequent speaker and knowledge disseminator at radios, podcast shows, scientific workshops, the general press, academic settings, and outreach events.

**Alberto Casas-Ortiz** is a software engineer specialized in scientific software and research for health-related fields such as biomechanics, neurotoxicity, and genetics. He holds a Master's degree in Artificial Intelligence and is pursuing a PhD in Intelligent Systems at UNED. Currently, Alberto works as a software

engineer for Athlete Performance Research at Stanford University. His research interests include applied artificial intelligence for health sciences and motor skill learning.

**Dr. Olga C. Santos** is an Associate Professor with the Artificial Intelligence Department of the Computer Science School at UNED, the Spanish Open University. She is main researcher in the HUMANOID project, the president of the IAIED Society and the Head of the PhyUM Research Center, among other responsibilities. She has received several recognitions, such as the Young Researcher Award in 2014 by the IEEE Technical Committee on Learning Technology. She has participated in 16 research projects (UE, National) and published more than 150 articles. She has been practicing Aikido since 2011, and is currently first dan (black belt). She has also taken some courses taught by AETAIKI, the Spanish Association of Aikido Technicians, aimed to achieve the certification of fukushido (assistant instructor) in Aikido. Her current research interests focus on combining artificial intelligence with sensing technologies to support personalized psychomotor learning following a human-centric approach.