Article

A Super-Bagging Method for Volleyball Action Recognition Using Wearable Sensors

Fasih Haider [1,*] **, Fahim A. Salim** [2] **, Dees B.W. Postma** [3] **, Robby van Delden** [3] **,
Dennis Reidsma** [3] **, Bert-Jan van Beijnum** [2] **and Saturnino Luz** [1]

[1] Usher Institute, Edinburgh Medical School, The University of Edinburgh, Edinburgh EH16 4UX, UK;
   S.Luz@ed.ac.uk
[2] Biomedical Signals and Systems, University of Twente, 7500 AE Enschede, The Netherlands;
   f.a.salim@utwente.nl (F.A.S.); b.j.f.vanbeijnum@utwente.nl (B.-J.v.B.)
[3] Human Media Interaction, University of Twente, 7500 AE Enschede, The Netherlands;
   d.b.w.postma@utwente.nl (D.B.W.P.); r.w.vandelden@utwente.nl (R.v.D.); d.reidsma@utwente.nl (D.R.)
* Correspondence: Fasih.Haider@ed.ac.uk

check for
updates

Abstract: Access to performance data during matches and training sessions is important for coaches
and players. Although there are many video tagging systems available which can provide such access,
these systems require manual effort. Data from Inertial Measurement Units (IMU) could be used
for automatically tagging video recordings in terms of players' actions. However, the data gathered
during volleyball sessions are generally very imbalanced, since for an individual player most time
intervals can be classified as "non-actions" rather than "actions". This makes automatic annotation
of video recordings of volleyball matches a challenging machine-learning problem. To address
this problem, we evaluated balanced and imbalanced learning methods with our newly proposed
'super-bagging' method for volleyball action modelling. All methods are evaluated using six classifiers
and four sensors (i.e., accelerometer, magnetometer, gyroscope and barometer). We demonstrate that
imbalanced learning provides better unweighted average recall, (UAR = 83.99%) for the non-dominant
hand using a naive Bayes classifier than balanced learning, while balanced learning provides better
performance (UAR = 84.18%) for the dominant hand using a tree bagger classifier than imbalanced
learning. Our super-bagging method provides the best UAR (84.19%). It is also noted that the
super-bagging method provides better averaged UAR than balanced and imbalanced methods in 8
out of 10 cases, hence demonstrating the potential of the super-bagging method for IMU's sensor
data. One of the potential applications of these novel models is fatigue and stamina estimation e.g.,
by keeping track of how many actions a player is performing and when these are being performed.

Keywords: sensor fusion; behavior analysis; social signal processing; machine learning; bagging;
boosting; action recognition; wearable technologies; multimodal systems

1. Introduction

   Top performance in sports depends on training programs designed by team staff, with a regime
of physical, technical, tactical and perceptual–cognitive exercises. Depending on how athletes perform,
exercises are adapted, or the program could be redesigned. State of the art data science methods have
led to groundbreaking changes. Data come from sources such as position and motion of athletes in
basketball [1], and baseball and football match statistics [2].

   Furthermore, new hardware platforms have appeared, such as LED displays integrated into
a sports court [3] and custom tangible sports interfaces [4], which offer possibilities for hybrid training
with a mix of technological and non-technological elements [3]. This has led to novel kinds of

exercises [4,5] including real-time feedback that can be well-fitted to the specifics of athletes in a highly controlled way. Data science tools can then be used to put well-fitted modifications to the parameters of such training. These developments are not limited to elite sports. Interaction technologies are also used for youth sports (e.g., the widely used player development system (www.dotcomsport.nl—last accessed May 2020)), and school sports and Physical Education [6].

Identifying actions automatically in sport activities is important and numerous studies have been conducted for that purpose [7–10]. Wearable devices such as Inertial Measurement Units (IMUs) [11,12] are becoming increasingly popular for sports related action analysis because of their reasonable price as well as portability [10]. While researchers have proposed different configurations in terms of number and placement of sensors [13], it is ideal to keep the number of sensors to a minimum due to cost, setup effort and player's comfort [13–16]. However, the data gathered during a volleyball session generally suffer from the class imbalance or 'curse of imbalanced data sets' problem, which is among the top 10 data mining problems today. The performance can degrade significantly for classifiers that assume a well-balanced class distribution and equal cost for misclassification [17]. To tackle that problem, different techniques are available such as oversampling [18–21], undersampling, decision trees [22] or ensemble methods [23]. One disadvantage of oversampling is that it could lead to model over-fitting, while the undersampling method will lose information when balancing the classes. Ensemble methods are also proposed to tackle the class imbalanced problem [23,24]. There are different types of ensemble methods including bagging [25] and boosting [26]. A bagging algorithm randomly select subsets of a data set to train an ensemble of classifiers while a boosting algorithm selects a subset of features to train an ensemble. The method proposed in this study (i.e., super-bagging method) is based on balanced (undersampling) and imbalanced (full sampling) learning instead of randomly selecting subsets of data sets for training.

This study extends our previous work [27–29] in which we used IMU data from both dominant and non-dominant wrists for classification of action and non-actions events (i.e., a two class problem). The previous study [29] provided us with interesting results regarding the role of the non-dominant hand in volleyball action and non-action classification. The current study demonstrates both balanced and imbalanced learning, and proposes a novel super-bagging method for classification of action and non-action events (i.e., two class problem). A potential application of the proposed models is in fatigue and stamina estimation [8]. This paper's contributions to the field are:

1. Proposal of a novel ensemble method (i.e., the super-bagging method) and its demonstration for volleyball action modelling,
2. Evaluation of the super-bagging method against undersampling (i.e, balanced learning) , full sampling (i.e., imbalanced learning) and ensemble (i.e., tree bagger) methods for volleyball action modelling,
3. Demonstration of the role of dominant and non-dominant hand for volleyball action modelling using super-bagging method,
4. Evaluation of all four IMU sensors separately and in combination for volleyball action modelling using different learning methods (i.e., balanced learning, imbalanced learning and super-bagging methods).

2. Related Work

Many approaches have been proposed for human activity recognition. They can be categorized mainly into two main categories: wearable sensor-based and vision-based. Vision-based methods employ cameras to detect and recognize activities using several computer vision techniques e.g., Zivkovic et al. propose a robust player segmentation algorithm. Novel features are extracted from video frames, and finally, classification results for different classes of tennis strokes using Hidden Markov Model are reported [30].

Wearable sensor-based methods collect input signals from wearable sensors mounted on human bodies such as accelerometer and gyroscope. For example, Liu et al. [31] identified temporal patterns

among actions and used those patterns to represent activities for the purpose of automatic recognition. Kautz et al. [32] presented an automatic monitoring system for beach volleyball based on wearable sensor devices which are placed at the wrist of the dominant hand of players. Beach volleyball *serve* recognition from a wrist-worn gyroscope is proposed in Cuspinera et al. [33]. Jarit et al. [34] showed that the grip strength of non-dominant and dominant hands is almost the same for college baseball players.

Inertial Measurement Units (IMUs) [11,12] have been used to detect sport activities in different sports e.g., soccer; Schuldhaus et al. use a custom-made system comprise of sensors and memory to collect data regarding the lower extremities of soccer players to classify shot pass in soccer [35]. The usage of wearable devices is not limited to sports, e.g., Wang et al. [36] use wearable sensors to form a wireless body area network to sense various physiological parameters of the human body, while others [37,38] have crafted ways to make the process energy efficient and secure. In tennis, Pei et al. use the JY-61 sensor to acquire motion information, such as accelerometer, that is used to detect tennis stroke type such as forehand, backhand and serve by using acceleration data as well as angular velocity [10]. Similarly, Kos et al. placed a miniature wearable IMU device on the player's forearm to classify the common type of tennis strokes [39].

Particularly for volleyball, Vales-Alons et al. developed a Smart Coaching Assistant for professional volleyball players to analyze exercise quality control by analyzing repetitions of the same action using dynamic programming [8]. Bagautdinov et al. use a neural network approach to detect individual activity to infer joint team activities in the context of volleyball games [9]. In their work, Wang et al. assessed the skill of volleyball spikers. The level of the players were classified into three levels of group such as elite, sub-elite and amateur by a Support Vector Machine (SVM) classifier [13].

It can be concluded that there are multiple studies that take into account the use of IMUs sensors and computer vision for sports related events. However, one of the limitations of a computer vision approach is that it cannot work well in the volleyball setting when the player positions change and when the sight of a player is occluded by some other player. Hence, the IMUs sensors is a good fit for volleyball settings. It is also noted that while there are quite a few studies focused on volleyball action modelling, most of the studies take into account the role of the dominant hand particularly for volleyball action modelling and the role of the non-dominant hand is less explored in sports related activities.

3. Our Approach

The presented paper extends upon the ideas presented in our previous work [27–29]. Figure 1 shows the overall system architecture. The presented paper focuses on step 2 of the proposed system. However, to give the reader the full idea, this section provides a summary of all the steps of the proposed approach.
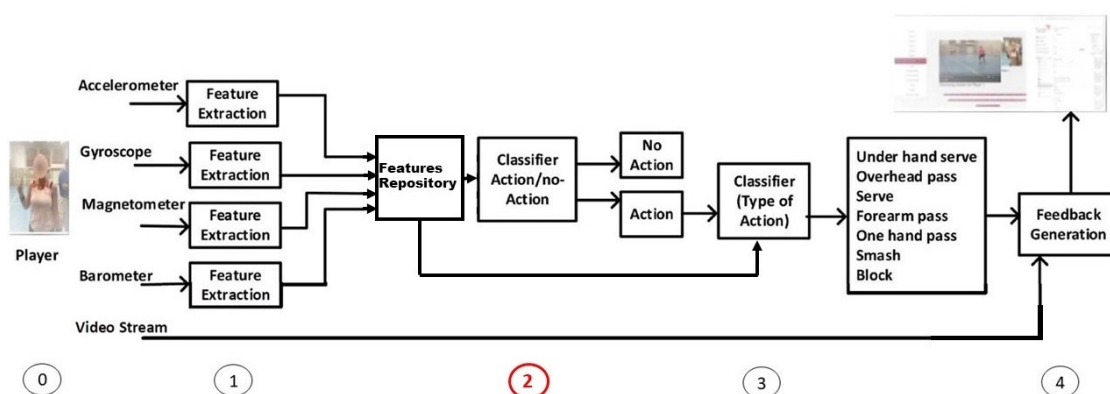

Figure 1. Block diagram of prototype system architecture adopted from [28].

Data were collected for 9 female volleyball players who wore IMUs on both wrists and were encouraged to play naturally during their routine training session i.e., step (0) in Figure 1. The hardware used in this study are the Xsens MTw Awinda (https://www.xsens.com/products/mtw-awinda—last accessed May 2020) IMU sensors [11] and two video cameras. The video streams are synchronized with the IMU's sensor data streams for further processing.

3.1. Data Annotation

To obtain the ground truth for machine-learning model training, the video recording was annotated using the Elan software (see Figure 2) [40]. Three annotators annotated the video. Since volleyball actions performed by players are quite distinct there is no ambiguity in terms of inter-annotator agreement. The quality of the annotation is evaluated by a majority vote i.e., if all annotators have annotated the same action or if an annotator might have missed or mislabeled an action.
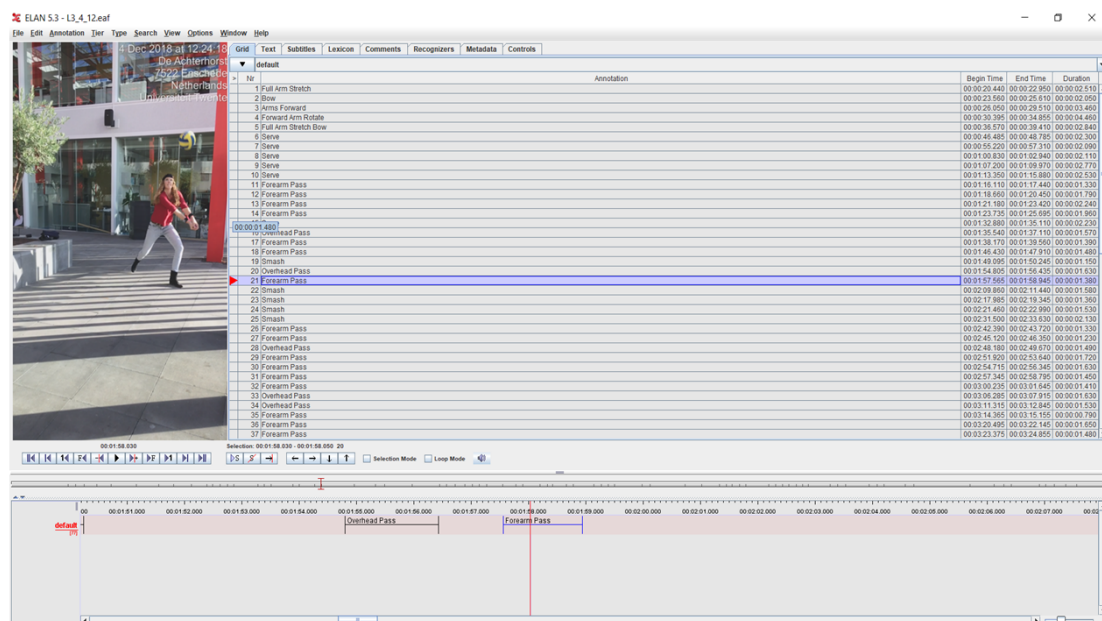

Figure 2. Annotation example with Elan annotation tool [28].

As a result, for the action case and the non-action case there were 1453 and 24,412 s of data, respectively. Table 1 shows the amount of data (in seconds) for each player. This data set is made available to the research community upon request. The annotators annotated the type of volleyball actions such as under hand serve, overhead pass, serve, forearm pass, one hand pass, smash, underhand pass and any other activity such as walking is considered to be non-action. Table 1 also details the number of volleyball actions performed by each player.

Table 1. Data Set Description: Time taken by each player (ID) to perform actions and non-actions, number and type of actions performed by each player (ID) and Dominant Hand (DH; Right(R) or Left(L)) information.

ID	DH	Action(sec)	Non-Action(sec)	# Actions	Forearm Pass	One hand Pass	Overhead Pass	Serve	Smash	Underhand Serve	Block
S_1	R	198	3055.25	120	40	3	16	0	29	28	4
S_2	L	193.75	3061	125	36	2	14	32	15	0	6
S_3	R	191	3030	116	50	3	3	34	25	0	1
S_5	R	176.75	3054.5	124	46	2	19	21	28	4	4
S_6	R	228.5	3009	150	30	1	70	0	12	30	7
S_7	R	135.5	3080.25	106	39	4	13	0	14	34	2
S_8	R	146.25	3077.5	105	34	4	16	34	17	0	0
S_9	R	183.25	3044.5	144	42	1	58	33	4	1	5
total	–	1453	24,412	990	317	20	209	154	144	97	49

3.2. Auto-tagging System Prototype

The proposed system performs classification in two stages i.e., step (2) and step (3). In step (2) binary classification (detection of start and end times of an action) is performed to identify if a players is performing an action or not using supervised machine learning at frame level [29]. After detecting the start and end times of an action, in step (3) (Figure 1), the type of volleyball action performed by the players is classified using supervised machine-learning algorithms. Once the action type is identified, its information along with the timestamp is stored in a repository for indexing purposes. Information related to the video, players and actions performed by the players are indexed and stored as documents in tables or cores in the Solr search platform [41]. An example of a 'Smash' indexed by Solr is shown as below:

```
"id":"25_06_Player_1_action_2"
"player_id":["25_06_Player_1"],
"action_name":["Smash"],
"timestamp":["00:02:15"],
"_version_":1638860511128846336
```

An interactive system is developed to allow player and coaches, access to performance data by automatically supplementing video recordings of training sessions and matches. The interactive system is developed as a web application. The server-side is written using the asp.net MVC framework. The front-end is developed using HTML5/Javascript. Figure 3 shows a screen shot of the front-end of the developed system. The player list and actions list are dynamically populated by querying the repository. The viewer can filter the actions by player and action-type (e.g., overhead pass by player 3). Once a particular action item is clicked or taped, the video is automatically jumped to the time interval where the action is being performed.

Currently the developed system lets a user filter types of action performed by each player, the details of the interactive system is described in [27,28].
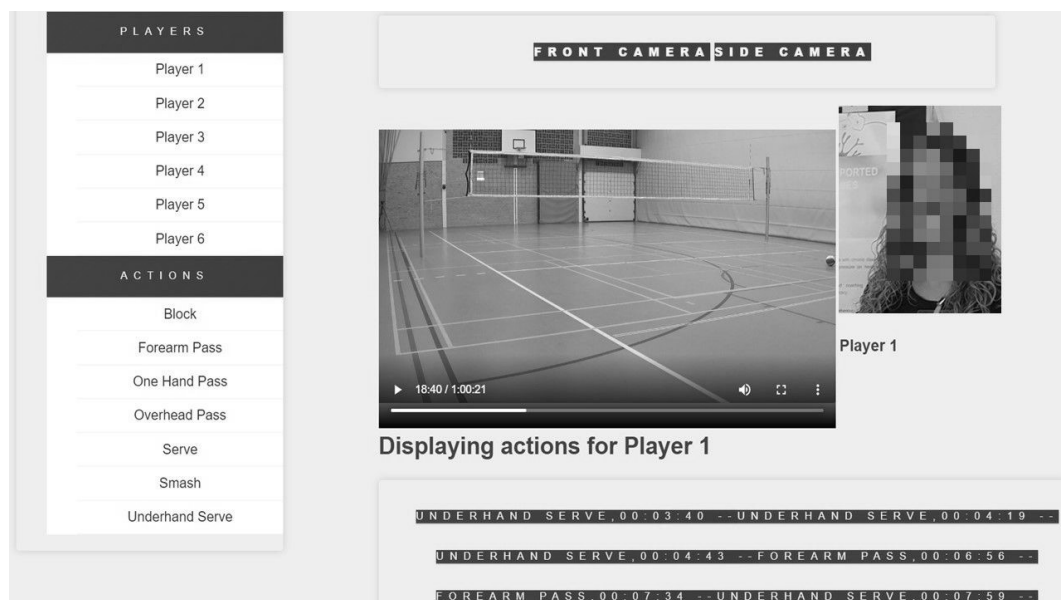

Figure 3. Interactive front-end system [28].

4. Super-Bagging Method

This section describes the super-bagging method for training a classifier for imbalanced data. We call the method super-bagging because, like bagging methods, an ensemble is trained on multiple

subsets of the data, but in contrast to regular bagging methods, rather than taking random subsets of the data, our method builds on top of balanced undersampling and unbalanced full sampling data sets.
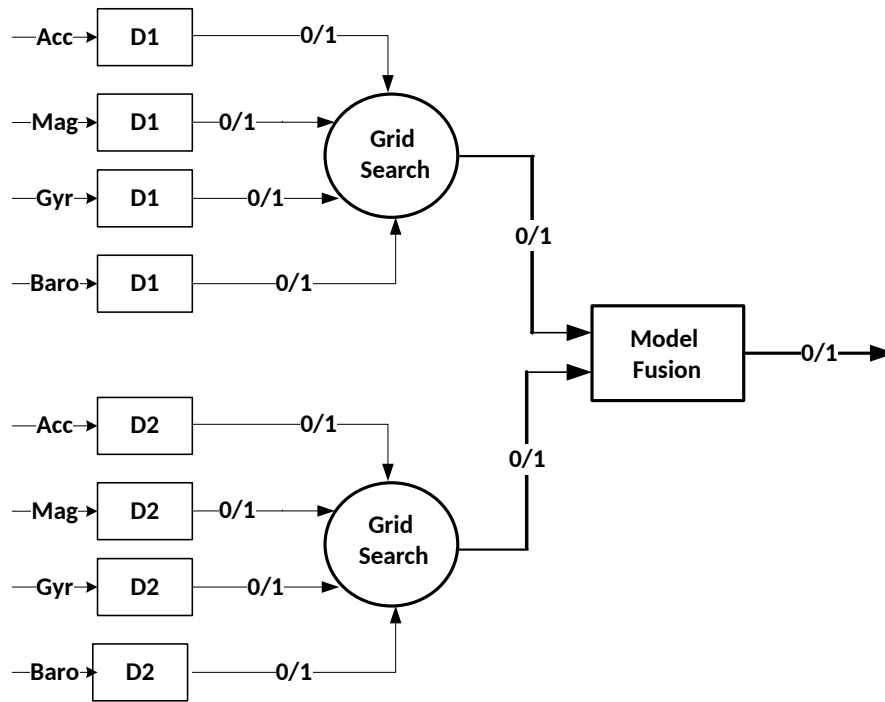

Figure 4. Super-bagging method for fusion of all sensors (the grid search is only for sensor fusion, in case of one sensor such as an accelerometer there is no grid search but late fusion), M_{D_1} and M_{D_2} represent the machine-learning models trained on D_1 and D_2 respectively.

Given a standard training set D of size n (i.e., observations), super-bagging generates two new training sets D_1 with size n and D_2 with size n'. All observations are repeated in D_1, so $D = D_1$. However, D_2 contains a subset of D_1. Let n observations be distributed into two classes m_1 and m_2, with t_1 and t_2 observations respectively, so that $n = t_1 + t_2$ and $t_2 > t_1$. Then, let $n' = t_1 + t_2'$ where $t_2' = t_1$ (i.e., $n' = 2t_1$). This results in two training sets, one of which is a full imbalanced training set (D_1) and the other is a balanced training set (D_2). Two machine learning models have been trained using training sets D_1 and D_2. Each of the model results are fused using decision (late) fusion method i.e., labelling an instance as a non-action event in case of unanimity only. In case of fusion of sensors, the number of votes to label an instance as non-action is searched through a grid-search algorithm. The architecture of the algorithm is shown in Figure 4.

5. Experimentation

This section describes the machine-learning models training using balanced, imbalanced and super-bagging methods for action and non-action events recognition.

5.1. The Data Set

We evaluated the super-bagging method using the data set [28] which we collected with the aim of developing volleyball action recognition components to be used in interactive digital-physical volleyball exercise applications [42]. This data set was collected during a volleyball training session as described in Section 3. The data set is highly imbalanced: around 94% of the data belong to the non-action class. Hence, different machine-learning approaches need to be explored in this setting.

5.2. Feature Extraction

We have extracted time domain features by applying basic six statistical-functionals such as mean, standard deviation, median, mode, skewness and kurtosis which are extracted over a frame length (i.e., time window) of 0.5 seconds 50% overlapping frames step (1) of Figure 1. The gyroscope, magnetometer and accelerometer data are three-dimensional, which is why we get 6×3 features over a frame for the sensors and barometer is single-dimensional data which results in 6×1 features for each frame in total.

5.3. Classification Methods

The classification experiments were performed using six different methods, namely decision trees (DT, with leaf size of 10, where the leaf size is optimized through a grid search within a range of 1 to 20), nearest neighbor (KNN with K = 5, where K parameter is optimized through a grid search within a range of 1 to 10), linear discriminant analysis (LDA), Tree Bagger (TB, with 50 trees and a leaf size of 10 where leaf size is optimized through a grid search within a range of 1 to 20), Naive Bayes (NB, with kernel distribution assumption optimized through a grid search for kernel smoothing density estimate, Multinomial distribution, Multivariate multinomial distribution and Normal distribution) and support vector machines (SVM, with a linear kernel (optimized by trying different kernel function i.e., linear, Gaussian, RBF and polynomial) with box constraint of 0.5 (optimized by trying a grid search between 0.1 to 1.0), and sequential minimal optimization solver (optimized by trying different solvers i.e., iterative single data algorithm, L1 soft-margin minimization by quadratic programming and sequential minimal optimization)).

The classification methods are implemented in MATLAB (http://uk.mathworks.com/products/matlab/ (December 2018)) using the statistics and machine-learning toolbox. The classifier hyper-parameters maximum ranges (such as K = 10) are set using hit and trial method. A leave-one-subject-out (LOSO) cross-validation setting was adopted, where the training data do not contain any information of the validation subjects. To assess the classification results, we used the Unweighted Average Recall (UAR) instead of overall accuracy as the data set is highly imbalanced. The unweighted average recall is the arithmetic average of recall of both classes.

5.4. Experiments

The overall action frames for eight players were 5812 frames while the in non-action case there were 97,648 frames. One can understand from the samples that the data set is imbalanced. To evaluate the performance of the IMU sensor, we trained machine-learning models using balanced as well as imbalanced data sets for the recognition of action and non-action frames. It is done using different classifiers and we evaluated their effectiveness for handling balanced and imbalanced data sets (i.e., IMU's sensors) for volleyball action recognition as some classifier are less affected by the class imbalance nature such as NB and KNN. We have conducted mainly three experiments as follow:

- **Experiment 1 (M_{D_1}):** training is performed on the imbalanced data sets (i.e., D_1) in terms of action and non-actions and validation is performed on the imbalanced data set (i.e., D_1) in leave-one-subject out settings. The prior-probabilities of classifiers are set according to the classes distribution.
- **Experiment 2 (M_{D_2}):** training is performed on the balanced data sets (i.e., D_2) in terms of actions and non-actions, where same number of non-actions events (selected randomly) and action events for each player are used. The validation is performed on the imbalanced data set (i.e., D_1) in leave-one-subject out settings. The prior-probabilities of classifiers are set to be equal for both classes as in this setting the distribution of classes is same.
- **Experiment 3 ($M_{D_1} + M_{D_2}$):** training is performed using the super-bagging method and validation is performed on the imbalanced data set in leave-one-subject out settings.

6. Results and Discussions

This section describes the results of machine-learning models for action and non-action events and demonstrates the discriminative power of different IMU sensors placed on the dominant and non-dominant hand.

6.1. Experiment 1 (M_{D_1}): Imbalanced Learning Method

The UAR of the dominant hand and non-dominant hand for all sensors are shown in Tables 2 and 3 respectively. These results indicate that the non-dominant hand (83.99%) provides better UAR than the dominant hand (79.83%), with NB being the best classifier for action detection. The results indicated that the accelerometer provides the best averaged UAR of 69.83% for the dominant hand and 73.24% for the non-dominant hand. The averaged UAR also indicates that the accelerometer (74.14%) and magnetometer (73.52%) provide better UAR on the non-dominant hand than on the dominant hand. The average UAR of fusion results indicate that the non-dominant hand provides better results (74.42%) than the dominant hand (70.81%).

Table 2. Imbalanced Learning method results for the dominant hand: Unweighted Average Recall (UAR) in % along with the standard deviation (Std) of UAR for each fold (i.e., subject) of classification. The bold figures indicate the best results.

Sensor	TB		DT		KNN		NB		SVM		LDA		Avg.
	UAR	Std	UAR	Std	UAR	Std	UAR	Std	UAR	Std	UAR	Std	UAR
Acc.	70.17	0.02	70.83	0.02	68.83	0.02	**79.83**	0.03	59.77	0.02	69.56	0.03	69.83
Mag.	60.67	0.03	63.10	0.02	57.12	0.02	74.16	0.03	50.00	0	67.71	0.03	62.13
Gyr.	61.55	0.03	64.07	0.03	60.78	0.02	74.58	0.03	53.35	0.02	64.86	0.03	63.20
Baro.	58.43	0.03	59.22	0.05	56.53	0.04	57.24	0.06	53.01	0.01	56.78	0.03	56.87
Fusion	70.37	0.07	70.75	0.10	68.77	0.08	80.30	0.02	60.14	0.02	74.53	0.03	**70.81**

Table 3. Imbalanced Learning method results for non-dominant hand: Unweighted Average Recall (UAR) in % along with the standard deviation (Std) of UAR for each fold (i.e., subject) of classification. The bold figures indicate the best results.

Sensor	TB		DT		KNN		NB		SVM		LDA		Avg.
	UAR	Std	UAR	Std	UAR	Std	UAR	Std	UAR	Std	UAR	Std	UAR
Acc.	68.59	0.05	71.53	0.13	72.98	0.12	**83.99**	0.06	66.47	0.08	75.90	0.09	73.24
Mag.	58.41	0.02	76.61	0.11	67.67	0.09	80.83	0.11	66.75	0.09	75.74	0.10	71.00
Gyr.	60.37	0.03	61.42	0.05	58.85	0.03	75.71	0.07	50.00	0	64.70	0.04	61.84
Baro.	52.16	0.02	40.86	0.21	38.56	0.22	31.53	0.21	50.00	0	50.53	0.00	43.94
Fusion	71.64	0.06	71.85	0.24	66.93	0.25	79.58	0.08	73.59	0.10	82.93	0.09	**74.42**

6.2. Experiment 2 (M_{D_2}): Balanced Learning Method

The UAR of dominant hand and non-dominant hand for all sensors are shown in Tables 4 and 5 respectively. These results indicate that the dominant hand (84.18%) provides better UAR than the non-dominant hand (82.16%), with TB being the best classifier for action detection. The results indicated that the accelerometer provides the best averaged UAR of 82.29% for the dominant hand and 78.26% for the non-dominant hand. The averaged UAR also indicates that all sensors provide better UAR on the dominant hand than on the non-dominant hand. The average UAR of fusion results indicate that the dominant hand provides better results (81.00%) than the non-dominant hand (78.26%).

Table 4. Balanced Learning method results for the dominant hand: Unweighted Average Recall (UAR) in % along with the standard deviation (Std) of UAR for each fold (i.e., subject) of classification. The bold figures indicate the best results.

Sensor	TB		DT		KNN		NB		SVM		LDA		Avg.
	UAR	Std	UAR	Std	UAR	Std	UAR	Std	UAR	Std	UAR	Std	UAR
Acc.	**84.18**	0.03	81.99	0.02	82.50	0.02	82.19	0.03	82.35	0.02	80.52	0.02	**82.29**
Mag.	81.71	0.02	77.47	0.02	74.86	0.02	79.25	0.04	79.50	0.03	79.08	0.03	78.65
Gyr.	77.91	0.05	73.72	0.03	75.48	0.04	75.94	0.04	74.17	0.04	72.78	0.03	75.00
Baro.	58.51	0.09	57.19	0.06	56.80	0.08	59.30	0.08	61.45	0.03	61.01	0.03	59.04
Fusion	83.10	0.03	79.46	0.03	80.69	0.03	80.83	0.04	81.57	0.03	80.32	0.02	81.00

Table 5. Balanced Learning method results for non-dominant hand: Unweighted Average Recall (UAR) in % along with the standard deviation (Std) of UAR for each fold (i.e., subject) of classification. The bold figures indicate the bestresults.

Sensor	TB		DT		KNN		NB		SVM		LDA		Avg.
	UAR	Std	UAR	Std	UAR	Std	UAR	Std	UAR	Std	UAR	Std	UAR
Acc.	**82.16**	0.03	78.90	0.03	80.33	0.03	81.71	0.02	81.28	0.03	79.84	0.04	**80.70**
Mag.	77.59	0.04	74.80	0.03	69.59	0.04	75.31	0.04	76.69	0.04	75.90	0.05	74.98
Gyr.	76.79	0.03	72.84	0.02	73.42	0.03	74.74	0.04	75.35	0.03	75.10	0.04	74.71
Baro.	53.07	0.04	51.57	0.03	50.22	0.04	49.46	0.06	55.88	0.02	56.07	0.02	52.72
Fusion	79.59	0.03	76.70	0.03	76.18	0.03	78.25	0.03	79.60	0.04	79.24	0.04	78.26

6.3. Experiment 3 ($M_{D_1} + M_{D_2}$): Super-Bagging Method

The UAR of dominant hand and non-dominant hand for all sensors are shown in Tables 6 and 7 respectively. These results indicate that the dominant hand (84.19%) provides better UAR than the non-dominant hand (82.93%), with TB being the best classifier for action detection. The results indicated that the accelerometer provides the best averaged UAR of 82.43% for the dominant hand and 80.91% for the non-dominant hand. The averaged UAR also indicates that the all sensor provide better UAR on the dominant hand than on the non-dominant hand. The average UAR of fusion results indicates that the dominant hand provides better results (81.91%) than the non-dominant hand (80.08%).

Table 6. Super-bagging results for the dominant hand: Unweighted Average Recall (UAR) in % along with the standard deviation (Std) of UAR for each fold (i.e., subject) of classification. The bold figures indicate the best results.

Sensor	TB		DT		KNN		NB		SVM		LDA		Avg.
	UAR	Std	UAR	Std	UAR	Std	UAR	Std	UAR	Std	UAR	Std	UAR
Acc.	**84.19**	0.03	82.70	0.02	82.50	0.02	82.19	0.03	82.35	0.02	80.67	0.02	**82.43**
Mag.	81.67	0.02	77.97	0.02	74.86	0.02	79.18	0.04	79.50	0.03	79.08	0.05	78.71
Gyr.	77.91	0.05	74.34	0.03	75.48	0.04	75.95	0.04	74.17	0.04	72.80	0.03	75.11
Baro.	58.51	0.09	57.25	0.06	56.80	0.08	59.32	0.08	61.45	0.03	61.04	0.03	59.06
Fusion	82.87	0.03	80.22	0.03	80.59	0.03	81.24	0.03	81.58	0.03	80.66	0.02	81.19

Table 7. Super-bagging results for non-dominant hand: Unweighted Average Recall (UAR) in % along with the standard deviation (Std) of UAR for each fold (i.e., subject) of classification. The bold figures indicate the best results.

Sensor	TB		DT		KNN		NB		SVM		LDA		Avg.
	UAR	Std	UAR	Std	UAR	Std	UAR	Std	UAR	Std	UAR	Std	UAR
Acc.	82.40	0.03	79.38	0.06	80.30	0.05	**82.93**	0.04	80.50	0.04	79.93	0.05	**80.91**
Mag.	77.59	0.04	78.95	0.04	71.36	0.06	77.89	0.06	77.21	0.04	77.82	0.05	76.80
Gyr.	76.79	0.03	72.05	0.03	73.15	0.03	72.37	0.04	75.35	0.03	74.63	0.04	74.06
Baro.	53.07	0.04	47.35	0.10	45.66	0.10	39.80	0.11	55.88	0.02	56.07	0.02	49.64
Fusion	80.11	0.03	80.12	0.08	77.65	0.08	80.87	0.06	80.13	0.04	81.58	0.05	80.08

6.4. Discussion

The results reported above indicate that Exp 3 (i.e., super-bagging) improved the UAR and provided the best average UAR of 81.19% and 80.08% for dominant and non-dominant hand, respectively. We have also noted that the best UAR was obtained using the TB classifier. TB with super-bagging improve the UAR of fusion for non-dominant hand from 79.59% to 80.11% but results in slight decrease in UAR for the dominant hand from 83.10% to 82.87%.

It is demonstrated that the imbalanced learning provides better UAR (83.99%) for the non-dominant hand using a Naive Bayes classifier than balanced learning, as Naive Bayes does not work with an assumption of balanced distribution. The balanced learning provides better UAR of 84.18% for the dominant hand using the tree bagger classifier than imbalanced learning. It could be due to the reason that the dominant hand requires less information (i.e., the movements of the non-dominant hand do not vary a lot while performing a volleyball action) for action modelling than the non-dominant hand. The super-bagging method provides the best UAR of 84.19%. To get further insight of the results we reported the confusion matrix of the best results as shown in Figure 5. To compare the result of super-bagging (84.19%) and balanced learning (84.18%), we set a null hypothesis that both methods provide the same results for a mid-p value McNemar test. The test rejects the null hypothesis with $p = 2.0432 \times 10^{-36}$.

Figure 5. Dominant Hand: Confusion Matrices for tree bagger classifier for all learning methods i.e., balanced, imbalanced and super-bagging methods.

However it is also noted that imbalanced learning (83.99% with NB) is more accurate in capturing the non-dominant hand information than balanced learning (82.1% with tree bagger) and super-bagging method (82.93% with NB). To get further insights of the of the results, we reported the average UAR in Table 8. From Table 8, it is noted that the super-bagging method provides better averaged UAR in 8 out of 10 cases than balanced and imbalanced methods.

Table 8. Results summary: The bold figures indicate the best averaged UAR (%) of each sensor for Dominant Hand (DH) and Non-dominant Hand (NDH).

Sensor	Imbalanced		Balanced		Super-Bagging	
	DH	**NDH**	**DH**	**NDH**	**DH**	**NDH**
Acc.	69.83	73.24	82.29	80.70	**82.43**	**80.91**
Mag.	62.13	71.00	78.65	74.98	**78.71**	**76.80**
Gyr.	63.20	61.84	75.00	**74.71**	**75.11**	74.06
Baro.	56.87	43.94	59.04	**52.72**	**59.06**	49.64
Fusion	70.81	74.42	81.00	78.26	**81.19**	**80.08**

The previous study [43] provided us with interesting results regarding the role of the non-dominant hand in volleyball action and non-action classification. However, in that study, we used an imbalanced learning method which suggests that the non-dominant hand provides more accurate results than the dominant hand. The current study uses both balanced and imbalanced learning and our newly proposed super-bagging method, and suggests that the dominant hand provides more accurate results

than the non-dominant hand for balanced learning and super-bagging method. It is also demonstrated that the balanced learning provides higher average UAR (81.00%) than imbalanced learning (70.81%), which are even more marked with 'super-bagging method' with a UAR of 81.19% for the dominant hand. It is indicating that super-bagging can capture more information than balanced and imbalanced learning methods. However, these results need further research to investigate/analyze different data sets for multiple applied machine-learning problems such as emotion recognition and type of volleyball action recognition.

The previous work detailed in Section 2 is focusing on a small number of sensors instead of the evaluation of four sensors which are used in this study. It is also noted that while there are quite a few studies focused on volleyball action modelling, most of the studies take into account the role of the dominant hand particularly for volleyball action modelling and the role of the non-dominant hand is less explored in sports related activities. This study demonstrates the role of both dominant and non-dominant hand movements. The proposed novel method (i.e., super-bagging method) is a fusion of imbalanced and balanced learning method which results in using the full data set (no missing information) for training and avoids the 'curse of imbalanced data set' using only two classifiers in an ensemble. The potential application of the proposed models can be interesting for fatigue and stamina estimation [8], where players/trainers are only interested in determining the amount of actions performed regardless of their type.

7. Conclusions

This article demonstrated the relevance of a balanced (undersampling method), imbalanced (full sampling) and super-bagging method for volleyball action modelling. Machine-learning models operating on IMU's sensors provided UAR of up to 84.19%, which is well above the chance level of 50%. The undersampling method provided more accurate results than the full sampling method which is more marked with our super-bagging method. It is also noted that the undersampling method provided better results for the dominant hand than full sampling method. However, the full sampling method provided better results for the non-dominant hand compared to the undersampling method. It is also noted that the super-bagging method provides a better averaged UAR in 8 out of 10 cases for sensors than balanced and imbalanced methods. Hence, demonstrating the potential of a super-bagging method for IMU's sensor data. The difference is small but it is the first testing of a super-bagging method which encourages further exploration of this method on different machine-learning problems and by also adjusting the weights of both classifiers in super-bagging ensemble and exploring score fusion methods. In the future, we aim to extend this research by incorporating different frequency domain features such as spectrogram, and to employ the super-bagging method to evaluate its generalizability particularly for multi-class problems.

References

1. Thomas, G.; Gade, R.; Moeslund, T.B.; Carr, P.; Hilton, A. Computer vision for sports: Current applications and research topics. *Comput. Vis. Image Underst.* **2017**, *159*, 3–18. [CrossRef]
2. Stensland, H.K.; Landsverk, Ø.; Griwodz, C.; Halvorsen, P.; Johansen, D.; Gaddam, V.R.; Tennøe, M.; Helgedagsrud, E.; Næss, M.; Stenhaug, M.; et al. Bagadus: An integrated real time system for soccer analytics. *ACM Trans. Multimed. Comput. Commun. Appl.* **2014**, *10*, 1–21. [CrossRef]
3. Kajastila, R. Motion Games in Real Sports Environments. *Interactions* **2015**, *22*, 44–47. [CrossRef]
4. Ludvigsen, M.; Fogtmann, M.H.; Grønbæk, K. TacTowers: An interactive training equipment for elite athletes. In Proceedings of the 8th ACM Conference on Designing Interactive Systems, Aarhus, Denmark, 16–20 August 2010; pp. 412–415. [CrossRef]
5. Jensen, M.M.; Rasmussen, M.K.; Mueller, F.F.; Grønbæk, K. Keepin' it Real. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15), Seoul, Korea, 18–23 April 2015; pp. 2003–2012. [CrossRef]
6. Koekoek, J.; van der Mars, H.; van der Kamp, J.; Walinga, W.; van Hilvoorde, I. Aligning Digital Video Technology with Game Pedagogy in Physical Education. *J. Phys. Educ. Recreat. Dance* **2018**, *89*, 12–22. [CrossRef]
7. Matejka, J.; Grossman, T.; Fitzmaurice, G. Video Lens: Rapid Playback and Exploration of Large Video Collections and Associated Metadata. In Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology, Honolulu, HI, USA, 5–8 October 2014; pp. 541–550. [CrossRef]
8. Vales-Alonso, J.; Chaves-Dieguez, D.; Lopez-Matencio, P.; Alcaraz, J.J.; Parrado-Garcia, F.J.; Gonzalez-Castano, F.J. SAETA: A Smart Coaching Assistant for Professional Volleyball Training. *IEEE Trans. Syst. Man Cybern. Syst.* **2015**, *45*, 1138–1150. [CrossRef]
9. Bagautdinov, T.; Alahi, A.; Fleuret, F.; Fua, P.; Savarese, S. Social scene understanding: End-to-end multi-person action localization and collective activity recognition. In Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2017), Honolulu, HI, USA, 21–26 July 2017; pp. 3425–3434. [CrossRef]
10. Pei, W.; Wang, J.; Xu, X.; Wu, Z.; Du, X. An embedded 6-axis sensor based recognition for tennis stroke. In Proceedings of the 2017 IEEE International Conference on Consumer Electronics (ICCE 2017), Bengaluru, India, 8–11 January 2017; pp. 55–58. [CrossRef]
11. Bellusci, G.; Dijkstra, F.; Slycke, P. *Xsens MTw: Miniature Wireless Inertial Motion Tracker for Highly Accurate 3D Kinematic Applications*; Xsens Technologies B.V.: Enschede, The Netherlands, 2018; pp. 1–9. [CrossRef]
12. X-IO Technologies. NG-IMU. 2019. Available online: http://x-io.co.uk/ngimu/ (accessed on 24 June 2019).
13. Wang, Y.; Zhao, Y.; Chan, R.H.; Li, W.J. Volleyball Skill Assessment Using a Single Wearable Micro Inertial Measurement Unit at Wrist. *IEEE Access* **2018**, *6*, 13758–13765. [CrossRef]
14. Cancela, J.; Pastorino, M.; Tzallas, A.T.; Tsipouras, M.G.; Rigas, G.; Arredondo, M.T.; Fotiadis, D.I. Wearability assessment of a wearable system for Parkinson's disease remote monitoring based on a body area network of sensors. *Sensors* **2014**, *14*, 17235–17255. [CrossRef] [PubMed]
15. Ismail, S.I.; Osman, E.; Sulaiman, N.; Adnan, R. Comparison between Marker-less Kinect-based and Conventional 2D Motion Analysis System on Vertical Jump Kinematic Properties Measured from Sagittal View. In *Proceedings of the 10th International Symposium on Computer Science in Sports (ISCSS)*; Springer: Cham, Switzerland, 2016; Volume 392, pp. 11–17. [CrossRef]
16. von Marcard, T.; Rosenhahn, B.; Black, M.J.; Pons-Moll, G. Sparse Inertial Poser: Automatic 3D Human Pose Estimation from Sparse IMUs. *Comput. Graph. Forum* **2017**, *36*, 349–360. [CrossRef]
17. Japkowicz, N.; Stephen, S. The class imbalance problem: A systematic study. *Intell. Data Anal.* **2002**, *6*, 429–449. [CrossRef]
18. Chawla, N.V.; Bowyer, K.W.; Hall, L.O.; Kegelmeyer, W.P. SMOTE: synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **2002**, *16*, 321–357. [CrossRef]
19. García, V.; Sánchez, J.S.; Martín-Félez, R.; Mollineda, R.A. Surrounding neighborhood-based SMOTE for learning from imbalanced data sets. *Prog. Artif. Intell.* **2012**, *1*, 347–362. [CrossRef]
20. He, H.; Bai, Y.; Garcia, E.A.; Li, S. ADASYN: Adaptive synthetic sampling approach for imbalanced learning. In Proceedings of the 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence), Hong Kong, China, 1–6 June 2008; pp. 1322–1328.

21. Zhou, L. Performance of corporate bankruptcy prediction models on imbalanced data set: The effect of sampling methods. *Knowl. Based Syst.* **2013**, *41*, 16–25. [CrossRef]

22. Liu, W.; Chawla, S.; Cieslak, D.A.; Chawla, N.V. *A Robust Decision Tree Algorithm for Imbalanced Data Sets*; SIAM: Philadelphia, PA, USA, 2010; pp. 766–777.

23. Lemaître, G.; Nogueira, F.; Aridas, C.K. Imbalanced-learn: A python toolbox to tackle the curse of imbalanced data sets in machine learning. *J. Mach. Learn. Res.* **2017**, *18*, 559–563.

24. Wang, S.; Yao, X. Using class imbalance learning for software defect prediction. *IEEE Trans. Reliab.* **2013**, *62*, 434–443. [CrossRef]

25. Breiman, L. Bagging predictors. *Mach. Learn.* **1996**, *24*, 123–140. [CrossRef]

26. Freund, Y.; Schapire, R.E. Experiments with a new boosting algorithm. In Proceedings of the Thirteenth International Conference Machine Learning, Bari, Italy, 3–6 July 1996; Volume 96, pp. 148–156.

27. Salim, F.; Haider, F.; Tasdemir, S.B.Y.; Naghashi, V.; Tengiz, I.; Cengiz, K.; Postma, D.; Delden, R.V.; Reidsma, D.; Luz, S.; et al. A Searching and Automatic Video Tagging Tool for Events of Interest During Volleyball Training Sessions. In *2019 International Conference on Multimodal Interaction*; ACM: New York, NY, USA, 2019; ICMI '19, pp. 501–503. [CrossRef]

28. Salim, F.A.; Haider, F.; Tasdemir, S.; Naghashi, V.; Tengiz, I.; Cengiz, K.; Postma, D.B.W.; Delden, R.V.; Reidsma, D.; Luz, S.; et al. Volleyball Action Modelling for Behavior Analysis and Interactive Multi-modal Feedback. In Proceedings of the 15th International Summer Workshop on Multimodal Interfaces (eNTERFACE'19), Ankara, Turkey, 8 July–2 August 2019.

29. Haider, F.; Salim, F.; Naghashi, V.; Tasdemir, S.B.Y.; Tengiz, I.; Cengiz, K.; Postma, D.; Delden, R.V.; Reidsma, D.; van Beijnum, B.J.; et al. Evaluation of Dominant and Non-Dominant Hand Movements For Volleyball Action Modelling. In Proceedings of the Adjunct of the 2019 International Conference on Multimodal Interaction, Suzhou, China, 14–18 October 2019; pp. 8:1–8:6. [CrossRef]

30. Zivkovic, Z.; van der Heijden, F.; Petkovic, M.; Jonker, W. Image Segmentation and Feature Extraction for Recognizing Strokes in Tennis Game Videos. In Proceedings of the 7th Annual Conference of the Advanced School for Computing and Imaging, Heijen, The Netherlands, 30 May–1 June 2001; pp. 262–266.

31. Liu, Y.; Nie, L.; Liu, L.; Rosenblum, D.S. From Action to Activity. *Neurocomputing* **2016**, *181*, 108–115. [CrossRef]

32. Kautz, T.; Groh, B.H.; Hannink, J.; Jensen, U.; Strubberg, H.; Eskofier, B.M. Activity recognition in beach volleyball using a Deep Convolutional Neural Network. *Data Min. Knowl. Discov.* **2017**, *31*, 1678–1705. [CrossRef]

33. Cuspinera, L.P.; Uetsuji, S.; Morales, F.; Roggen, D. Beach volleyball serve type recognition. In Proceedings of the 2016 ACM International Symposium on Wearable Computers, Heidelberg, Germany, 12–16 September 2016; pp. 44–45.

34. Jarit, P. Dominant-hand to nondominant-hand grip-strength ratios of college baseball players. *J. Hand Ther.* **1991**, *4*, 123–126. [CrossRef]

35. Schuldhaus, D.; Zwick, C.; Körger, H.; Dorschky, E.; Kirk, R.; Eskofier, B.M. Inertial Sensor-Based Approach for Shot/Pass Classification During a Soccer Match. In Proceedings of the KDD Workshop on Large-Scale Sports Analytics, Sydney, Australia, 10–13 August 2015; Volume 27, pp. 1–4.

36. Wang, D.; Huang, Q.; Chen, X.; Ji, L. Location of three-dimensional movement for a human using a wearable multi-node instrument implemented by wireless body area networks. *Comput. Commun.* **2020**, *153*, 34–41. [CrossRef]

37. Pirbhulal, S.; Wu, W.; Li, G.; Sangaiah, A.K. Medical Information Security for Wearable Body Sensor Networks in Smart Healthcare. *IEEE Consum. Electron. Mag.* **2019**, *8*, 37–41. [CrossRef]

38. Sodhro, A.H.; Sangaiah, A.K.; Sodhro, G.H.; Lohano, S.; Pirbhulal, S. An energy-efficient algorithm for wearable electrocardiogram signal processing in ubiquitous healthcare applications. *Sensors* **2018**, *18*, 923. [CrossRef] [PubMed]

39. Kos, M.; Ženko, J.; Vlaj, D.; Kramberger, I. Tennis Stroke Detection and Classification Using Miniature Wearable IMU Device. In Proceedings of the 2016 International Conference on Systems, Signals and Image Processing (IWSSIP), Bratislava, Slovakia, 23–25 May 2016; pp. 1–4. [CrossRef]

40. Lausberg, H.; Sloetjes, H. Coding gestural behavior with the NEUROGES-ELAN system. *Behav. Res. Methods* **2009**, *41*, 841–849. [CrossRef] [PubMed]

41. Velasco, R. *Apache Solr: For Starters*; CreateSpace Independent Publishing Platform: Scotts Valley, CA, USA, 2016.

42. Postma, D.; van Delden, R.; Walinga, W.; Koekoek, J.; van Beijnum, B.J.; Salim, F.A.; van Hilvoorde, I.; Reidsma, D. Towards Smart Sports Exercises: FirstDesigns. In Proceedings of the Annual Symposium on Computer-Human Interaction in Play (CHI PLAY '19), Barcelona, Spain, 22–25 October 2019. [CrossRef]

43. Haider, F.; Salim, F.A.; Busra, S.; Tasdemir, Y.; Naghashi, V.; Cengiz, K.; Postma, D.B.W.; Delden, R.V.; Reidsma, D. Evaluation of Dominant and Non-Dominant Hand Movements For Volleyball Action Modelling. In Proceedings of the 21st ACM International Conference on Multimodal Interaction (ICMI 2019), Suzhou, China, 14–18 October 2019.