THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Towards Automatic Modelling of Volleyball Players' Behavior for Analysis, Feedback and Hybrid Training

Citation for published version:
Salim, FA, Haider, F, Postma, DBW, Van Delden, R, Reidsma, D, Luz, S & Van Beijnum, BJ 2020, 'Towards Automatic Modelling of Volleyball Players' Behavior for Analysis, Feedback and Hybrid Training', *Journal for the Measurement of Physcial Behaviour*, vol. 3, no. 4, pp. 323–330. https://doi.org/10.1123/jmpb.2020-0012

Digital Object Identifier (DOI):
10.1123/jmpb.2020-0012

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Journal for the Measurement of Physcial Behaviour

Towards Automatic Modelling of Volleyball Players' Behavior for Analysis, Feedback and Hybrid Training

Journal:	*Journal for the Measurement of Physical Behaviour*
Manuscript ID	JMPB.2020-0012.R2
Manuscript Type:	Original Research
Keywords:	accelerometers, activity classification, computational methods, machine learning, physical activity

Towards Automatic Modelling of Volleyball Players' Behavior for Analysis, Feedback and Hybrid Training

submitted: 31-03-2020

Correspondence Address:

Running Title: Towards Automatic Modelling of Volleyball Players' Behavior for Analysis, Feedback and Hybrid Training

Abstract

Automatic tagging of video recordings of sports matches and training sessions can be helpful to coaches and players, and provide access to structured data at a scale that would be unfeasible if one were to rely on manual tagging. Recognition of different actions forms an essential part of sports video tagging. In this paper, we employ machine learning techniques to automatically recognise specific types of volleyball actions (i.e. underhand serve, overhead pass, serve, forearm pass, one hand pass, smash and block which are manually annotated) during matches and training sessions (uncontrolled, in the wild data) based on motion data captured by inertial measurement unit (IMU) sensors strapped on the wrists of 8 female volleyball players. Analysis of the results suggests that all sensors in the IMU (i.e. magnetometer, accelerometer, barometer and gyroscope) contribute unique information in the classification of volleyball actions types. We demonstrate that while the accelerometer feature set provides better results than other sensors overall (i.e. gyroscope, magnetometer and barometer) feature fusion of the accelerometer, magnetometer and gyroscope provides the bests results (Unweighted Average Recall (UAR)= 67.87%, Unweighted Average Precision (UAP)= 68.68% and Kappa = 0.727), well above the chance level of 14.28%. Interestingly, it is also demonstrated that the dominant hand (UAR =61.45%, UAP= 65.41% and Kappa = 0.652) provides better results than the non-dominant (UAR = 45.56%, UAP = 55.45 and Kappa = 0.553) hand.

Apart from machine learning models, this paper also discusses a modular architecture for a system to automatically supplement video recording by detecting events of interests in volley-

ball matches and training sessions and to provide tailored and interactive multi-modal feedback by utilizing an html5/JavaScript application. A proof of concept prototype developed based on this architecture is also described.

Introduction

Coaches and players desire and would benefit greatly from easy access to performance data of matches and training sessions[15]. They use this information not only to monitor performance but also to plan training programs and game strategy. According to the assessment of volleyball coaches in Netherlands [1], the two areas which can substantially improve sports training are as follows:

- Interactive exercises and enhanced instructions.
- Providing the trainer with information from live data on player behaviour.

It is because 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 may be redesigned. State of the art data science methods have led to ground breaking changes. Data is from sources such as tracking position and motion of athletes in basketball[32] and baseball and football match statistics[30].

Furthermore, new hardware platforms appear, such as LED displays integrated into a sports court[12] or custom tangible sports interfaces[21]. These offer possibilities for hybrid training with a mix of technological and non-technological elements[12]. This has led to novel kinds of exercises[11,21] including real-time feedback, that can be tailored to the specifics of athletes in a highly controlled way.

These developments are not limited to elite sport. Interaction technologies are also used for youth sports (e.g., the widely used player development system of Dotcomsport.nl), and school sports and Physical Education[15].

Identification and classification of events of interest in sports recordings therefore, is of interest for not only coaches and players but also for sports fans who might, for example, wish to watch all home runs hit by a player during the 2013 baseball season[22], or a coach searching for video recordings related to the intended learning focus for a player or the whole training session[15].

Analysis of videos, displaying different events of interest, may help in getting insightful tactical play and engagement with players[8]. Video edited game analysis is a common method for post-game performance evaluation[15].

---

[1]https://www.volleybal.nl/eredivisie/dames -- last accessed (June, 2020)

However, these examples require events to be manually tagged which not only requires time and effort but would also splits a trainer's attention from training to tagging the events for later viewing and analysis.

A system which could automatically tag such events would help trainers avoid manual effort and has the potential to provide tailored and interactive multi-modal feedback to coaches and players. The approach described in this paper precisely addresses the above issue.

The context of the current paper is the Smart Sports Exercises project in which we aim to use multimodal sensor data and machine learning techniques to enable players and coaches to monitor performance but also to provide interactive feedback[26].

This paper extends our previous research[7,27,28,39] and details the architecture, components and a comprehensive analysis of a machine learning based system which automatically classifies volleyball actions performed by players during their regular training sessions. The presented paper demonstrates the following:

- Description of a proof of concept prototype of a real-time video supplementary system to allow coaches and players to easily search for the information or event of interest (e.g. All the serves by a particular player).

- Description of an annotated and anonymized Dataset of IMUs data of players while playing volleyball in real-life training scenarios.

- A novel and comprehensive analysis to:
  - the evaluation of each sensor data from IMUs (3D acceleration, 3D angular velocity, 3D magneto meter and air pressure) and their fusion for automatically identifying basic volleyball actions such as: under hand serve, overhead pass, serve, forearm pass, one hand pass, smash, block.
  - Evaluate the role of dominant and non-dominant hand for modelling the type of volleyball action.

Related Work

There are many applications of automatically identifying actions in sport activities[1,22,25,33]. Due to their portability and reasonable pricing, Wearable devices such as Inertial Measurement Units (IMUs)[2,31] are becoming increasingly popular for sports related action analysis[25].Researchers have proposed different configurations in terms of number and placement of sensors[36], however it is ideal to keep the number of sensors to minimum due to issues related to cost, setup effort and player's comfort[5,9,35,36].

Inertial Measurement Unit (IMU) sensors [2,31] have been utilized to automatically detect sport activities in numerous sports e.g. soccer[23,29], tennis[17,37], table tennis[3], hockey[23], basketball[20,24] and rugby[14]. Many approaches have been proposed for human activity recognition. They can be categorized into two main categories: wearable sensor-based and vision-based.

Vision-based methods employ cameras to detect and recognize activities using computer vision technologies. While wearable sensor-based methods collect input signals from wearable sensors mounted on human bodies such as accelerometer and gyroscope. For example, Liu et al.[19] identified temporal patterns among actions and used those patterns to represent activities for ~~automatic action~~automatic action recognition. Kautz et al.[13] presented an automatic monitoring system for beach volleyball based on wearable sensor devices which are placed at wrist of dominant hand of players. Beach volleyball *serve* recognition from a wrist-worn gyroscope is proposed in Cuspinera et al.[6] which is placed on the forearm of players. Kos et al.[16] proposed a method for tennis stroke detection. They used a wearable IMU device which is located on the players' wrists. A robust player segmentation algorithm and novel features are extracted from video frames, and finally, classification results for different classes of tennis strokes using Hidden Markov Model are reported[38].

~~Jarit et al.[10] studied college baseball players, in total 88 subjects of two groups. Jamar dynamometer was used to test maximum grip strength (kgf) for both hands. The recording was done for dominant and nondominant hands. The highest measurements were taken for the statistical analysis. Every subject put their maximal effort. 2-factor repeated measures to analyze the variance was used to compare both hands' grip strength ratios of the experimental and control group. Results of the study showed that there is no significant differences of baseball players' dominant and nondominant hands grip strength.~~

Based on the above literature, we have concluded that the most studies take into account the role of dominant hand particularly for volleyball action modelling and the role of non-dominant hand is less explored. It is also noted that none of the studies above evaluated the IMU sensors for volley-ball action recognition. The paper extends our previous work[7,27,28,39] in which we evaluated the IMU sensors for two class problem (action and no-action). However this study evaluates the sensors for type of volley-ball action such as serve or block which is a seven class problem.

By combining machine learning models based on IMUs sensors with a video tagging system, this paper opens up new opportunities for applying sensor technologies such as IMU sensors with interactive system to enhance the training experience.

Approach

The presented paper extends upon the ideas presented in our previous work[7,27,28,39]. Figure 1 shows the overall system architecture. This paper focuses on step 3 of the proposed system. However, this section provides a brief summary of all the steps to provide a full idea of the proposed approach.

Data was collected in a typical volleyball training session. In which 8 female volleyball players wore Inertial Measurements Units (IMU) on both wrists and were encouraged to play naturally step (0) in Figure 1. The details of the data collection protocol and annotation procedure is presented in section "Volleyball Data set".

Time domain features such as mean, standard deviation, median, mode, skewness and kurtosis are extracted over a frame length (i.e. time window) of 0.5 seconds of sensor data with an overlap of 50% with the neighbouring frame. See step(1) of figure 1.

Classification is performed in two stages i.e. step (2) and step (3). In step (2) binary classification is performed to identify if a player is performing action or not, using supervised machine learning with unweighted average recall (UAR) as high as 86.87%. The details of the action vs non-action classification procedure is described in[7,28,39]. Next in step (3) (figure 1), type of volleyball action performed by the players is classified using supervised machine learning algorithms. The details of type of action classification is described in section "Experimentation".

Once the actions are 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 Solr search platform[34]. An example of a Smash indexed by Solr is shown in table 1.

[Table 1 about here.]

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 web application. The server-side is written using asp.net MVC framework. While the front-end is developed using HTML5/Javascript.

Figure 2 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 user. Details of the interactive system are described in previous work[27,28].

[Figure 1 about here.]

[Figure 2 about here.]

Volleyball Data set

In order to collect data for the experimentation, 8 female volleyball players wore Inertial Measurement Units (IMU) on both wrists during their regular training session (see Figure 3). All players were amateur volleyball players and belonged to different age groups. The players were encouraged to play naturally so that the data is representative of real life training scenarios. The video is also recorded using two video cameras. Later the IMU sensors data and video streams are synchronised. No screen-shots of the recorded session are added due to explicit request by players not to publish their pictures or videos. It is done so that the models trained are capable of performing in the wild instead of controlled settings.

It is for this reason the collected data is highly imbalanced, e.g. for the binary classification task of action vs non-action recognition[39], there is 1453 vs 24412 seconds of data respectively.

Similar unbalanced can be seen in the type of volleyball actions performed by players. Table 2 shows the frequency of each volleyball action performed by each player.

[Figure 3 about here.]

[Table 2 about here.]

Three students annotated the video using Elan software[4]. All annotators were the participants of eNTERFACE2019 and the annotation task is not paid. 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 annotator have annotated the same action or if an annotator might have missed or mislabelled an action.

Experimentation

Feature Extraction The feature set for this paper is extracted from the feature set of a previous study conducted to distinguish actions from non-actions in volleyball training sessions[7]. In that study we used time domain features such as mean, standard deviation, median, mode, skewness and kurtosis which are extracted over a frame length of 0.5 seconds of sensor data with an overlap of 50% with the neighbouring frame. For the current study we did not apply frequency domain approaches or deep learning approaches due to fact that the data set is rather small for such approaches. The second reason for not opting to use deep learning methods is to evaluate IMU's sensor information in resource constrained settings such as a mobile application.

For the current study, we calculated an average of frame-level features over the time window length of an action. ~~the mean of each of the features of the starting frame and ending frame of each individual action.~~ It is done so because the current models are intended to be used on the classification performed by the previous model: first a classifier such as the one described in Haider et al.[7,39] would identify the presence of an action (start and end time of an action); subsequently the model trained and reported in this paper would further classify the type of that action.

Classification Methods

The classification experiments were performed using five different methods, namely decision trees (DT, with leaf size of 10), nearest neighbour (KNN with K=5), linear discriminant analysis (LDA), Naive Bayes (NB, with kernel distribution assumption) and support vector machines (SVM, with a linear kernel, box constraint of 0.5, and sequential minimal optimization solver).

The classification methods are implemented in MATLAB using the statistics and machine learning toolbox. A leave-one-subject-out (LOSO) cross-validation setting was adopted, where the training data does not contain any information of the validation subjects. To assess the classification results, we used the Unweighted Average Recall (UAR) as a primary measure as the dataset is imbalanced but we also reported overall accuracy, Unweighted Average Precision (UAP) and Kappa[18] for the best results.

The unweighted average recall is the arithmetic average of recall of all classes and unweighted average precision is the arithmetic average of precision of all classes.

---

[2] http://uk.mathworks.com/products/matlab/ (December 2018)

Results

The UAR of dominant hand and non-dominant hand for all sensors are shown in Table 3 and Table 4 respectively. These results indicate that the dominant hand (UAR= 61.45%, UAP = 65.45 and Kappa = 0.652) provides better results than the non-dominant hand (UAR=45.56%, UAP = 55.45% and Kappa = 0.553). The averaged UAR across sensors indicate that the SVM classifier provides the best average UAR (40.34%) across sensors for dominant hand and NB provides the best averaged UAR (34.85%) across sensors for non-dominant hand for action type detection. It is also noted that the accelerometer provides the best averaged UARs across classifiers for dominant (53.92%) and non-dominant (42.70%) hand. The pressure sensor provides the least UAR across classifiers, and the gyroscope provides better UAR across classifiers than the magnetometer. For further insights, confusion matrices of the best results using dominant hand and non-dominant hand are shown in Figure 4 and Figure 5 along with precision, recall of each class, overall accuracy, UAR, UAP and Kappa[18]. From Figure 4 and Figure 5, it is also noted that the dominant hand provides better kappa (0.652) than non-dominant hand (0.533). It is noted that the dominant hand provides better precision for ‘under hand serve' (78.79%), 'serve' (80.95%), 'over head pass' (74.80%),' one hand pass' (50.00%) and 'forearm pass'(75.12%). However, non-dominant hand provides better recall for 'smash' (76.67%) and 'block' (44.44%). It is also noted that the non-dominant hand (63.30%) provides better recall for 'smash' action than dominant hand (55.05). For all other actions the dominant hand provides better recall than non-dominant hand. It suggests that both hands are important in classifying type of volleyball actions. That is why, we also experimented with combining different sensors and also with using both the dominant and non-dominant hand to see if using both hands instead of only one hand would provide better results.

Table 5 shows the UAR using fusion of different sensors and using d~~d~~ominant hand~~-(DH)~~, n~~n~~on-d~~d~~ominant hand~~ (NDH)~~ and both hands. While the dominant hand gives better results (UAR = 61.79%) compared to the non-dominant hand (UAR= 54.28%). However, using both hands (UAR= 67.87%) provided better results than dominant hand. We also noted that the LDA provides better results than SVM. For further insights, confusion matrix of the best result for both hands is shown in Figure 6. It is noted that the fusion improves precision of 5 volleyball actions but results in a decrease of recall for 'one hand pass' (35.29%) and 'block' (25.00%). However, the overall accuracy (78.17%), UAR (67.87%) and Kappa (0.727) are improved. it is also noted that the fusion improves the recall of five volleyball actions but results in decrease of recall for 'block' (from 41.67% to 37.50%) and 'forearm pass' (from 85.99% to 81.64).

[Table 3 about here.]

[Table 4 about here.]

[Figure 4 about here.]

[Table 5 about here.]

[Figure 5 about here.]

To better understand the relationship between the dominant, non-dominant and both hands, we also drew the Venn diagram ~~depicted~~ shown in Figure 7. In that Figure, the blue area (labelled "Target") represents the annotated labels (i.e. ground truth), the green area represents the predicted labels when the *non-dominant hand* information was used, the red area represents the predicted labels when *dominant hand* information was used and finally the yellow area represents the prediction obtained with the *fusion* of both hands.

The Venn diagram suggests that the information captured by dominant and non-dominant hand is not similar, as only 320 out of 646 instances are detected by all the methods (i.e. dominant, non-dominant and fusion) and there are 74 out 646 instances which have not been captured by any of methods. Those 74 instances contain 8 of 'block', 16 of smash one of 'under hand serve', 12 of 'serve', 9 of 'over head pass', 18 of 'one hand pass' and 10 of 'forearm pass'.

[Figure 6 about here.]

Discussion

The results reported above show that the dominant hand plays an important role in classifying the type of action, compared to the non-dominant hand which provided better results for action vs no-action classification[7]. ~~However~~However, the non-dominant hand certainly plays a useful role in action type classification as the results improved to 67.87% UAR compared to 61.79% using only the dominant hand. The results are highly applicable as they demonstrate the added value of using sensors on both arms for type of action classification compared to using only one arm.

The results are highly encouraging and show the viability of the trained model to be used in a real time system[27]. While the 67.87% UAR does leaves room for improvement, it is our contention that it can be easily achieved by collecting data from a couple of additional training sessions, as the models are currently trained over a single training session in which players were encouraged to play naturally resulting in an unbalanced data set.

This~~e~~ article ~~presented paper~~ focuses~~d~~ on the type of volleyball action recognition. The overall approach works using two steps~~s~~ ~~in multipl~~classification method ~~steps~~ (see Figure 1). First the system classifies start and end times of an action and non-action event[7,39] (i.e. binary class problem see step 2 in Figure 1) and then upon detection of an action event, it further classifies the type of action (the focus of this article). In real life scenario, the system will use the machine learning models for both classification steps i.e. action vs non-action classification[7,39] ~~7, 39~~ -and type of action classification (see section Experimentation). -

Concluding Remarks

This paper has proposed and described an approach to model volleyball player behavior for analysis and feedback. The described system and machine learning models automatically identify volleyball specific actions and automatically tags video footage to enable easy access to relevant information for players and coaches. Apart from saving time and effort on the coach's behalf. By providing real time data the proposed approach opens up new possibilities for coaches to analyze player performance and provide quick and adaptive feedback during the training session.

The presented experiment also demonstrated the role of dominant and non-dominant hand in classification of volleyball action type and presented evaluation results of different sensors and machine learning methods. The results on the relatively small and unbalanced data set are highly encouraging and applicable.

Future Directions

The outcome of the presented paper has the potential to be extended in multiple ways.

In terms of machine learning models, we plan to use frequency domain features such as Scalogram and Spectrogram instead of time domain features currently used to train the models. Apart from extending the machine learning models the aim is to further develop the video tagging system from a proof of concept prototype to a more functional and integrated system.

The following list summarises possible ways to extend the project.

- Further classify actions
- Using frequency domain approaches for feature extraction such as scalogram, spectrogram.
- Using transfer learning approaches such as ResNet, AlexNet, VGGNet.
- Classification based on the above feature set.
- Further integration of Demo system and models.

In terms of further development and testing of the proposed system, we plan to conduct user studies with coaches and participants to understand the ways in which it can enhance their experience while performing their regular tasks. The user studies will be conducted using user centric design approaches and with systematic feedback from the participants to not only understand how the system is being used by them, but what functionalities can be added to the system to further enhance its usability for coaches and player alike.

References

1. Bagautdinov, T., Alahi, A., Fleuret, F., Fua, P., and Savarese, S. (2017). Social scene understanding: End-to-end multi-person action localization and collective activity recognition. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, 2017-Janua:3425–3434.

2. Bellusci, G., Dijkstra, F., and Slycke, P. (2018). Xsens MTw : Miniature Wireless Inertial Motion Tracker for Highly Accurate 3D Kinematic Applications. *Xsens Technologies*, (April):1–9.

3. Blank, P., Ho, J., Schuldhaus, D., and Eskofier, B. M. (2015). Sensor-based stroke detection and stroke type classification in table tennis. In *Proceedings of the 2015 ACM International Symposium on Wearable Computers*, ISWC ’15, pages 93–100, New York, NY, USA. ACM.

4. Brugman, H., Russel, A., and Nijmegen, X. (2004). Annotating multi-media/multi-modal

resources with elan. In *LREC*.

5. Cancela, J., Pastorino, M., Tzallas, A. T., Tsipouras, M. G., Rigas, G., Arredondo, M. T., and Fotiadis, D. I. (2014). Wearability assessment of a wearable system for Parkinson's disease remote monitoring based on a body area network of sensors. *Sensors (Switzerland)*, 14(9):17235–17255.
6. Cuspinera, L. P., Uetsuji, S., Morales, F., and Roggen, D. (2016). Beach volleyball serve type recognition. In *Proceedings of the 2016 ACM International Symposium on Wearable Computers*, pages 44–45. ACM.
7. 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., and Luz, S. (2019). Evaluation of dominant and non-dominant hand movements for volleyball action modelling. In *Adjunct of the 2019 International Conference on Multimodal Interaction*, ICMI '19, pages 8:1–8:6, New York, NY, USA. ACM.
8. Harvey, S. and Gittins, C. (2014). Effects of integrating video-based feedback into a Teaching Games for Understanding soccer unit. *Agora para la educación física y el deporte*, 16(3):271–290.
9. Ismail, S. I., Osman, E., Sulaiman, N., and Adnan, R. (2016). Comparison between Markerless Kinect-based and Conventional 2D Motion Analysis System on Vertical Jump Kinematic Properties Measured from Sagittal View. *Proceedings of the 10th International Symposium on Computer Science in Sports (ISCSS)*, 392(2007):11–17.
10. Jarit, P. (1991). Dominant-hand to nondominant-hand grip-strength ratios of college base-

ball players. *Journal of Hand Therapy*,4(3):123–126.

11. Jensen, M. M., Rasmussen, M. K., Mueller, F. F., and Grønbæk, K. (2015). Keepin' it Real. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*, (April):2003–2012.

12. Kajastila, R. (2015). Motion Games in Real Sports Environments. *Interactions*, (3):44–47.

13. Kautz, T., Groh, B. H., Hannink, J., Jensen, U., Strubberg, H., and Eskofier, B. M.(2017). Activity recognition in beach volleyball using a deep convolutional neural network. *Data Mining and Knowledge Discovery*, 31(6):1678–1705.

14. Kautz, T., thomas. kautz, and benjamin. groh (2015). Sensor fusion for multi-player activity recognition in game sports.

15. Koekoek, J., van der Mars, H., van der Kamp, J., Walinga, W., and van Hilvoorde, I. (2018). Aligning Digital Video Technology with Game Pedagogy in Physical Education. *Journal of Physical Education, Recreation & Dance*, 89(1):12–22.

16. Kos, M., Ženko, J., Vlaj, D., and Kramberger, I. (2016a). Tennis stroke detection and classification using miniature wearable imu device. In *2016 International Conference on Systems, Signals and Image Processing (IWSSIP)*, pages 1–4. IEEE.

17. Kos, M., Ženko, J., Vlaj, D., and Kramberger, I. (2016b). Tennis stroke detection and classification using miniature wearable imu device.

18. Landis, J. R. and Koch, G. G. (1977). The measurement of observer agreement for categorical data. *biometrics*, pages 159–174.

19. Liu, Y., Nie, L., Liu, L., and Rosenblum, D. S. (2016). From action to activity. *Neurocom-*

put., 181(C):108–115.

20. Lu, Y., Wei, Y., Liu, L., Zhong, J., Sun, L., and Liu, Y. (2017). Towards unsupervised physical activity recognition using smartphone accelerometers. *Multimedia Tools and Applications*, 76(8):10701–10719.21. Ludvigsen, M., Fogtmann, M. H., and Grønbæk, K. (2010). TacTowers: an interactive training equipment for elite athletes. *DIS '10 Proceedings of the 6th conference on Designing Interactive Systems*, pages 412–415.

22. Matejka, J., Grossman, T., and Fitzmaurice, G. (2014). Video Lens : Rapid Playback and Exploration of Large Video Collections and Associated Metadata. In *Uist*, pages 541–550.

23. Mitchell, E., Monaghan, D., and O'Connor, N. E. (2013). Classification of sporting activities using smartphone accelerometers. *Sensors (Switzerland)*, 13(4):5317–5337.

24. Nguyen Ngu Nguyen, L., Rodríguez-Martín, D., Català, A., Pérez, C., Samà Monsonís, A., and Cavallaro, A. (2015). Basketball activity recognition using wearable inertial measurement units.

25. Pei, W., Wang, J., Xu, X., Wu, Z., and Du, X. (2017). An embedded 6-axis sensor based recognition for tennis stroke. *2017 IEEE International Conference on Consumer Electronics, ICCE 2017*, pages 55–58.

26. Postma, D., van Delden, R., Walinga, W., Koekoek, J., van Beijnum, B.-J., Salim, F. A., van Hilvoorde, I., and Reidsma, D. Towards smart sports exercises: Firstdesigns. In *CHI PLAY Extended Abstracts '19*.

27. Salim, F., Haider, F., Tasdemir, S. B. Y., Naghashi, V., Tengiz, I., Cengiz, K., Postma, D., Delden, R. v., Reidsma, D., Luz, S., and van Beijnum, B.-J. (2019a). A searching and automatic video tagging tool for events of interest during volleyball training sessions. In *2019 International Conference on Multimodal Interaction*, ICMI '19, pages 501–503, New York, NY, USA. ACM.

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., and van Beijnum, B. J. F. (2019b). Volleyball Action Modelling for Behavior Analysis and Interactive Multi-modal Feedback. In *ENTERFACE'19*, Ankara.

29. Schuldhaus, D., Zwick, C., Körger, H., Dorschky, E., Kirk, R., and Eskofier, B. M. (2015). Inertial Sensor-Based Approach for Shot / Pass Classification During a Soccer Match. *Proc. 21st ACM KDD Workshop on Large-Scale Sports Analytics*, 27:1–4.

30. Stensland, H. K., Landsverk, Ø., Griwodz, C., Halvorsen, P., Stenhaug, M., Johansen, D., Gaddam, V. R., Tennøe, M., Helgedagsrud, E., Næss, M., Alstad, H. K., Mortensen, A., Langseth, R., and Ljødal, S. (2014). Bagadus: An integrated real time system for soccer analytics. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 10(1s):1–21.

31. Technologies, X.-i. (2019). NG-IMU.

32. Thomas, G., Gade, R., Moeslund, T. B., Carr, P., and Hilton, A. (2017). Computer vision for sports: Current applications and research topics. *Computer Vision and Image Understanding*, 159:3–18.

33. Vales-Alonso, J., Chaves-Dieguez, D., Lopez-Matencio, P., Alcaraz, J. J., Parrado-Garcia, F. J., and Gonzalez-Castano, F. J. (2015). SAETA: A Smart Coaching Assistant for Professional Volleyball Training. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 45(8):1138–1150.

34. Velasco, R. (2016). *Apache Solr: For Starters*. CreateSpace Independent Publishing Platform.

35. von Marcard, T., Rosenhahn, B., Black, M. J., and Pons-Moll, G. (2017). Sparse Inertial Poser: Automatic 3D Human Pose Estimation from Sparse IMUs. *Computer Graphics Forum*, 36(2):349–360.

36. Wang, Y., Zhao, Y., Chan, R. H., and Li, W. J. (2018). Volleyball Skill Assessment Using a Single Wearable Micro Inertial Measurement Unit at Wrist. *IEEE Access*, 6:13758–13765.

37. Weiping Pei, Jun Wang, Xubin Xu, Zhengwei Wu, and Xiaorong Du (2017). An embedded 6-axis sensor based recognition for tennis stroke. In *2017 IEEE International Conference on Consumer Electronics (ICCE)*, pages 55–58.

38. Zivkovic, Z., van der Heijden, F., Petkovic, M., and Jonker, W. (2001). Image segmentation and feature extraction for recognizing strokes in tennis game videos. In *Proc. of the ASCI*.

39. Haider, F., Salim, F.A., Postma, D.B., Delden, R.V., Reidsma, D., van Beijnum, B.J. and Luz, S., 2020. A Super-Bagging Method for Volleyball Action Recognition Using Wearable Sensors. Multimodal Technologies and Interaction, 4(2), p.33.

List of Tables

Table 1 Sample Solr structure

"id":"25_06_Player_1_action_2"

"player_id":["25_06_Player_1"],

"action_name":["Smash"],

"timestamp":["00:02:15"],

"_version_":1638860511128846336

Table 2 Data Set Description: number and type of actions performed by each player

ID	# Actions	Forearm Pass	Onehand Pass	Overhead Pass	Serve	Smash	Underhand Serve	Block
1	120	40	3	16	0	29	28	4
2	125	36	2	14	32	15	0	6
3	116	50	3	3	34	25	0	1
5	124	46	2	19	21	28	4	4
6	150	30	1	70	0	12	30	7
7	106	39	4	13	0	14	34	2
8	105	34	4	16	34	17	0	0
9	144	42	1	58	33	4	1	5
total	990	317	20	209	154	144	97	49

Table 3 Dominant Hand: Unweighted Average Recall

Sensor	DT	KNN	NB	SVM	LDA	avg.
Acc.	46.26	54.09	50.29	***61.45***	57.53	***53.92***
Mag.	35.67	34.98	37.72	36.31	40.88	*37.11*
Gyr.	41.61	36.07	35.77	42.09	38.89	**38.89**
Baro.	24.90	15.89	14.39	21.51	22.60	19.86
avg.	37.11	35.26	34.54	**40.34**	39.40	–

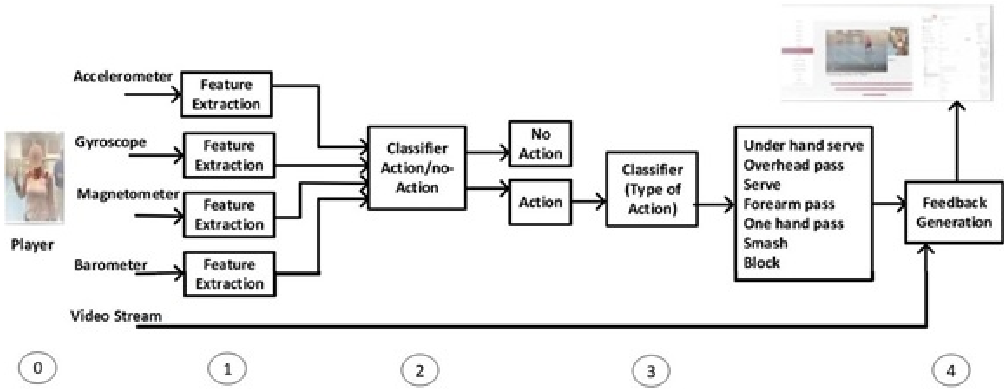
Table 4 Non-Dominant Hand: Unweighted Average Recall

Sensor	DT	KNN	NB	SVM	LDA	avg.
Acc.	39.85	37.67	45.06	45.38	***45.56***	***42.70***
Mag.	35.70	32.40	38.65	29.37	31.36	*33.50*
Gyr.	33.50	32.83	36.85	32.40	31.95	**33.51**
Baro.	16.32	12.77	18.83	14.29	15.42	15.53
avg.	31.34	28.92	**34.85**	30.36	31.07	–

Table 5 Sensor Fusion: Unweighted Average Recall (%) for Dominant Hand (DH), non-Dominant Hand (NDH) and Both Hands (BH)

	SVM			LDA		
Sensor	DH	NDH	BH	DH	NDH	BH
acc	61.45	45.38	57.61	57.53	45.56	62.96
Mag	36.31	29.37	44.50	40.88	31.36	50.12
Gyr	42.09	32.40	42.50	38.89	31.95	47.54
Baro	21.51	14.29	17.40	22.60	15.42	25.76
Acc + Mag	59.08	45.58	60.14	61.28	50.79	65.87
Acc + Gyr.	55.71	45.20	44.99	61.19	49.67	64.14
Acc + Baro.	**61.79**	45.37	54.99	58.34	49.12	63.47
Gyr + Mag	47.36	36.93	43.41	50.71	40.24	61.24
Acc + Mag + Gyr	55.50	43.76	44.06	60.95	**54.28**	***67.87***
Acc +gyr + Baro	55.92	44.54	44.47	61.06	50.54	64.72
All	55.43	43.59	44.22	59.76	53.87	67.78

List of Figures

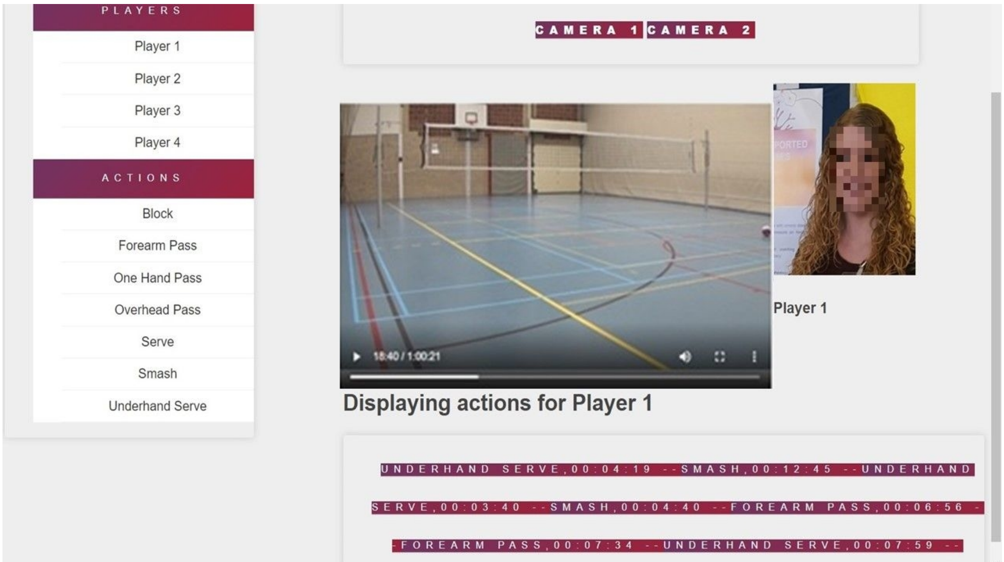


Prototype System Architecture

121x71mm (120 x 120 DPI)

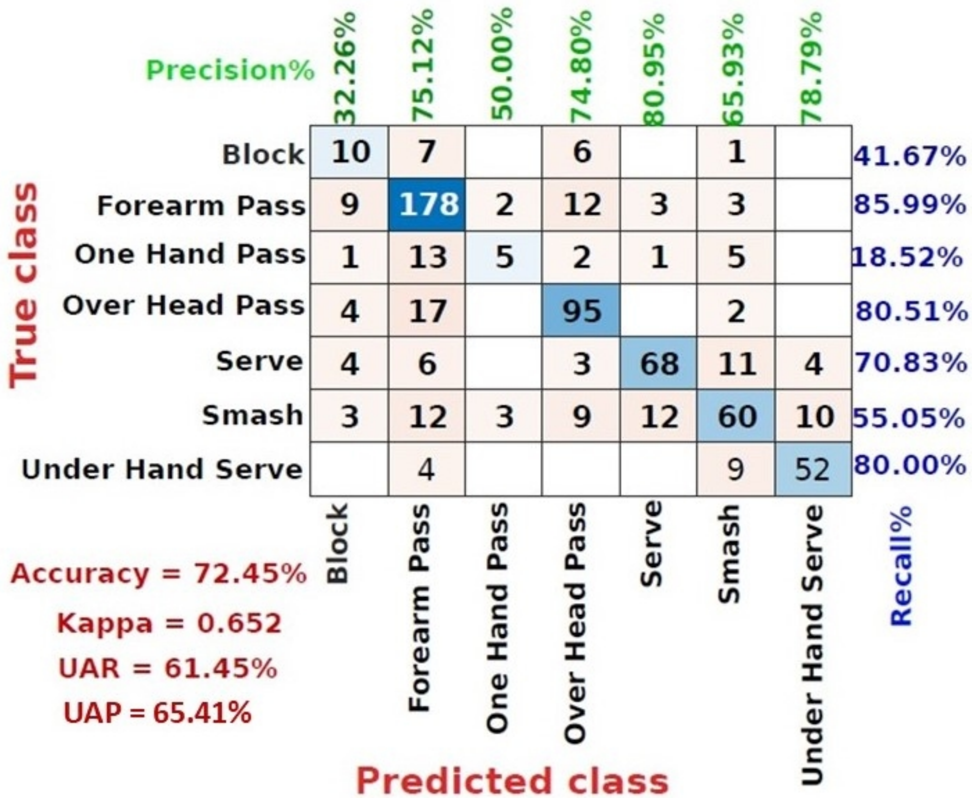
Player wearing 2 IMUs on both wrists

82x71mm (120 x 120 DPI)



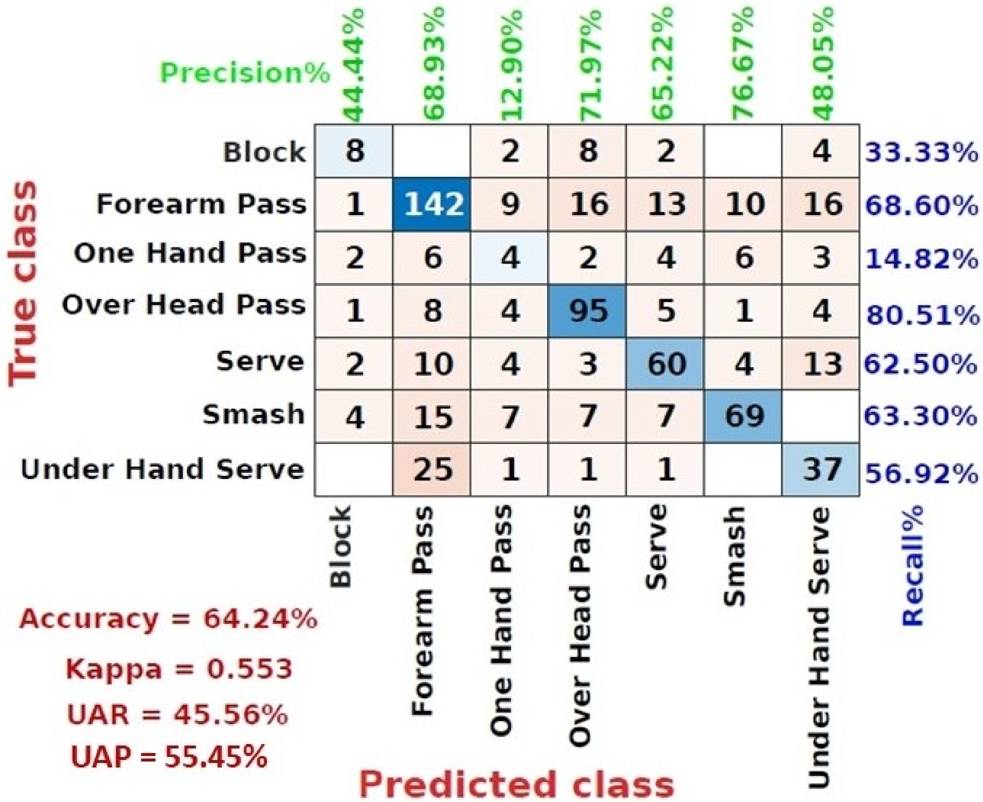
Interactive front-end system

451x254mm (72 x 72 DPI)



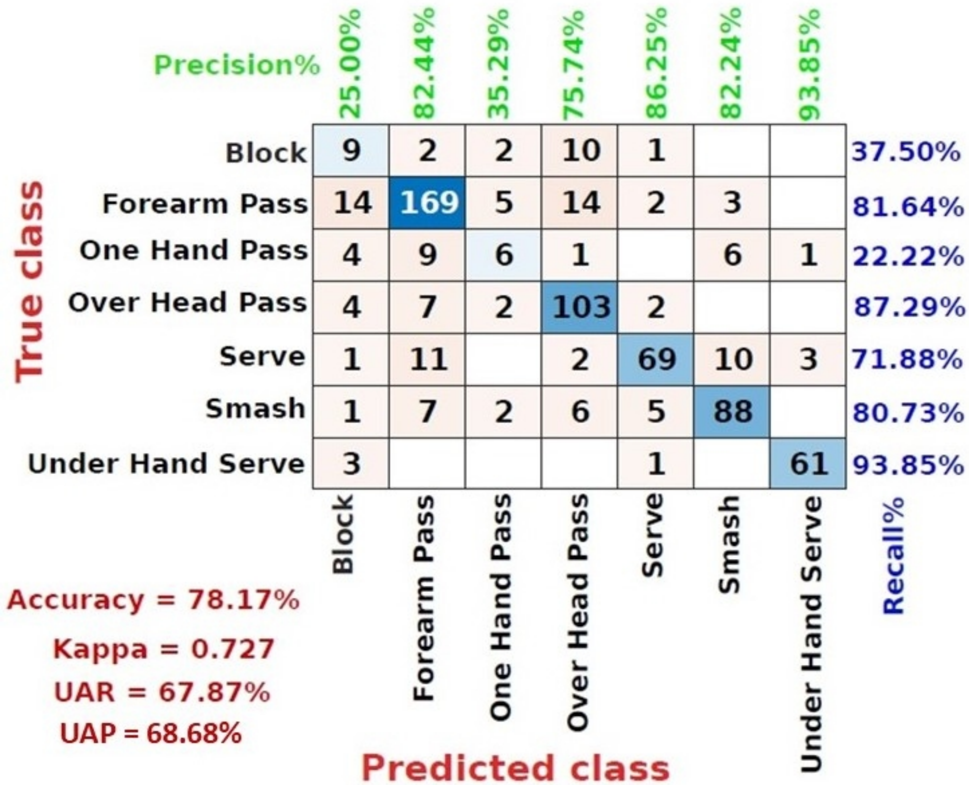
Confusion Matrix for best result using Dominant Hand Accelerometer and Barometer and SVM method

190x155mm (96 x 96 DPI)



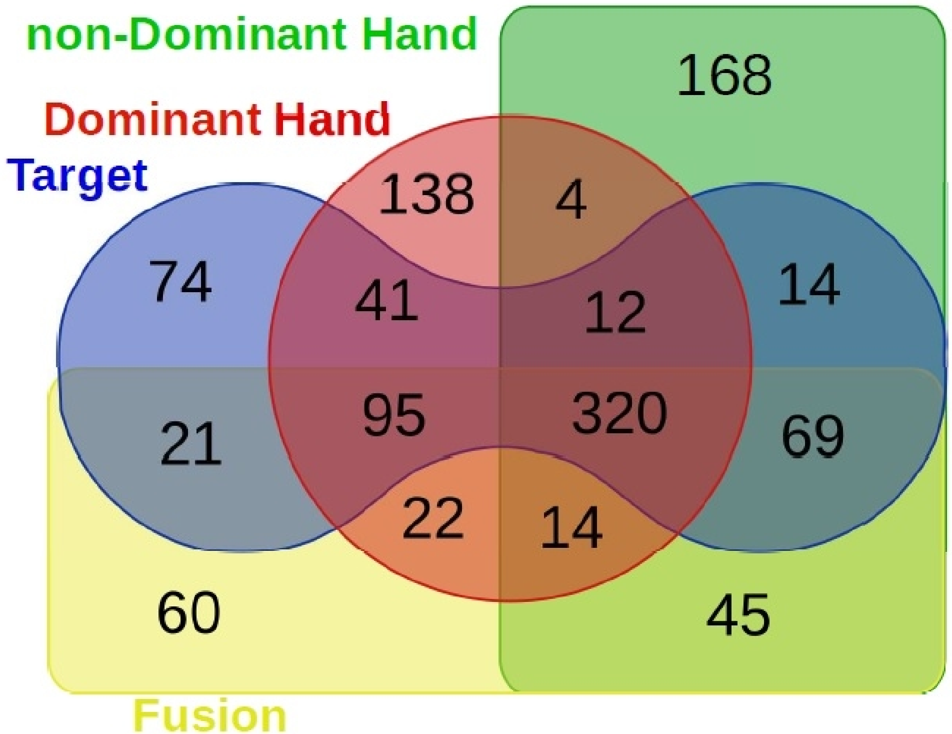
5. Confusion Matrix for best result using Non-Dominant Hand Accelerometer, Gyroscope and Magnetometer and LDA method

182x147mm (96 x 96 DPI)



6. Confusion Matrix for Both Hands and using Accelerometer, Gyroscope and Magnetometer and LDA method

197x158mm (96 x 96 DPI)



Venn diagram of the best results.

144x111mm (120 x 120 DPI)