Development and validity of the Playermaker ball touch classification

Moran Gad[a], Shai Rosenblit[b], Leo Herszenhaut[b] , Amir Zviran[c], Eran Amit[d] and Steve Barrett[e]

[a] Chief Technology Officer

[b] Data Scientist

[c] System Engineer

[d] Head of Algorithms

[e] Director of Sport Science and Research Innovation

1. Introduction

Kicking and ball interactions are a key component of some team sports such as football (soccer), Australian Rules Football (AFL), American Football (NFL) and both Rugby Union and League (Lees et al., 2010). The interaction with the ball is dependent on the sport but is representative of a method to score more goals/points than an opponent. In football (soccer), the technical and tactical statistics of kicking are regularly used by to measure performance indices including a breakdown of the number of passes, shots and crossing performed by one team against another (Barnes et al., 2014; Liu et al., 2015). These actions are currently used by coaches, performance analysts and scouting/recruitment departments, in order to assess a team and individual players performance within a given situation (Yi et al., 2019). These statistics are now openly available via media streams which allow the fans to engage with these statistics during competition. Semi-automated and manual coding methods are currently used in order to extract this data from match play and training using systems such as Opta, Stats, StatDNA, Hudl and many more (Liu et al., 2013). These systems require a number of operators in order to 'tag' the kicking events within sports and can be subjected to potential human error during these methods of tagging (Liu et al., 2016). When analyzing a professional football match in Spain (Real Madrid vs. Valencia), Liu and colleagues (2013), found that there was a very good agreement between operators (0.86-0.94). However, 4 operators were required for this method during one match in order to provide the data within 24 hours to the clubs, using multiple HD cameras situated around the stadium. The expense of these systems means that data collection tends to be restricted to within match situations and limited technical and tactical data is collected within a training environment as a result (Rein & Memmert, 2016).

Reliable and valid methods of automatically monitoring the kicking actions within both training and match environments, would provide coaches and practitioners further insight into the performance of their players (Hopkins, 2000). Inertial measurement devices (IMU's) containing accelerometers and gyroscopes have been suggested as an alternative method to monitor certain physical activity within different sports (Barrett et al., 2014; Boyd et al., 2011; McGregor et al., 2009). The use of accelerometer-based metrics have been used in order to quantify the external training loads of players in multiple sports providing a global measure of training load (Vanrenterghem et al., 2017). However, with an IMU placed between the scapulae, some of the insightful data to monitor actions such as kicking may be missed (Barrett et al., 2014; Needergaard et al., 2018; Vanrenterghem et al., 2017). Placing the device on the feet provides greater insight and allows practitioners to further monitor and quantify the kicking actions within different team sports. Therefore, the aim of this study was to assess the concurrent validity and development of a ball touch classification as measured by an IMU device placed on the feet (Playermaker).

2. Methods

Data was collected on 10 teams and age groups (U9-Seniors) as part of a development process for the company. Verbal consent was gained from each player as part of the protocol.

Experimental Design

Training (n= 40) and match (n= 24) sessions were recorded using an HD camera. Alongside the video, each player was fitted with two Playermaker devices, one on the left foot and one on the right foot, containing a 1000Hz IMU with both an accelerometer and a gyroscope (Invensense 6050, USA). These devices were fitted into silicone straps that were placed around the boot, with the device fitting on the lateral malleolus of the ankle (See **Figure 1**).

Figure 1. A single Playermaker device mounted on a football boot

In order to activate the devices, an operator would press 'start training' on an iPad which was connected via Bluetooth to all the devices. Devices would stay in range of the box until they were registered as 'Started Training'. Participants were instructed prior to each session on how to fit the device before they commenced the session. In order to sync up the video with the data, each participant would stand in front of the camera and double tap their feet on the floor. This defined a peak acceleration and was repeated at the end of the session. During the session, the operator would make a note of the time stamps associated with the start and end time of the individual training drills or match time periods. At the end of the training session, players returned to devices to the operator. Once all participants returned the devices, the operator pressed 'Stop Training'. Devices were then allowed time to sync to a cloud based software for further analysis.

Development of an algorithm to detect ball interactions

The data set contains 323 player hours of various scenarios, including unopposed drills, opposed drills and matches with different age groups and different professional levels of players. In total, 68,306 touches were labelled, with 47,251 touches in opposed drills and 21,055 touches in unopposed drills. This data was split into a development set (70% of the events), and a validation set (30% of the events). The validation set was kept aside and was not used at all during the development of the algorithm.

Raw accelerometer data and gyroscope data were extracted from the Playermaker software and synced up with the video using a custom-made labelling tool. This internally written MATLAB program allowed the events to be tagged throughout the video footage for each individual player, for each individual touch. In order to assess the data set, multiple labellers would categorise the touches of the ball by a receive or release for the individual player. Further internal tags, like touch foot zone, were performed as part of the labelling but aren't shown in this current study. Once labellers had defined the point and data was synced with the IMU data during the contact with the ball, this was recorded (See **Figure 2**) . These individual recordings would then be divided into "training data set" and "validation data set". The training data set would be used to allow the development of an algorithm using machine learning for automated detection of the action. These algorithms were then tested against the validation data set. The algorithm results on the validation data set is used to evaluate its performance. Data is presented as the percentage of actions that the algorithm was able to detect within both opposed and unopposed drills. For an example of the precision recall method used, please see **Appendix 1.**

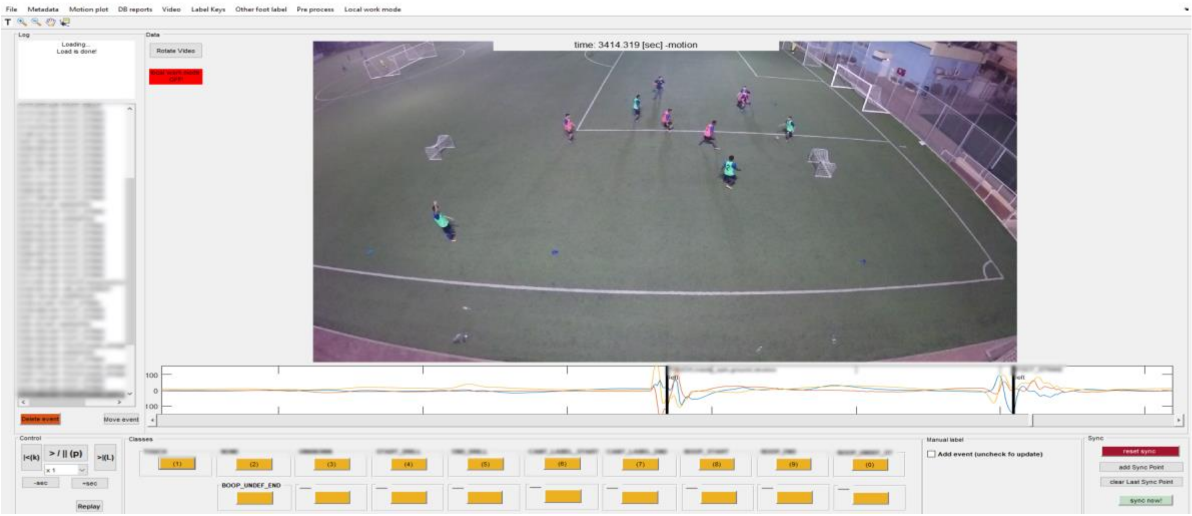

Figure 2. Playermaker's labelling tool

3. Results

On top of the touch classifier additional algorithms may be applied to improve the performance, pending the specific context. The metrics presented in the following tables were found to be the most useful for the analysis of football players' performance. The performance of the algorithms for these four metrics are presented in **Table 1**. Unopposed training drills were close to 100% across all measures. Opposed drills had a lower performance in comparison.

Table 1. The performance of the ball touch classification algorithm for an unopposed scenario

	Performance	Definition
Release Recall	94.9%	Out of all releases, what percentage is detected correctly
Release Precision	95.6%	Out of all reported releases, what percentage is really a release
Receive Recall	94.0%	Out of all receives, what percentage is detected correctly
Touch precision	95.9%	Out of all reported touches, what percentage is really a touch

Table 2. The performance of the ball touch classification algorithm, for opposed drills

	Performance	Definition
Release Recall	88.0%	Out of all releases, what percentage is detected correctly
Release Precision	92.6%	Out of all reported releases, what percentage is really a release
Receive Recall	89.2%	Out of all receives, what percentage is detected correctly
Touch Precision	88.6%	Out of all reported touches, what percentage is really a touch

4. Discussion

Playermaker was able to identify the technical actions associated with ball touch interactions, using machine learning algorithms. In overall training and match sessions the performance of the algorithm varied (See **Table 1**), depending on the action. During the validation process, the 8-10% of missing data was accounted for by miscellaneous touches which occur between multiple players (for example, during a tackle), however, it was still deemed as an acceptable level of validity (Hopkins, 2000). While not reported in the current results, the reliability of the algorithm was 1.00 (100%), when running the same algorithm for the same technical event. It is recommended that practitioners should always test and assess the validity and reliability of any system themselves specific to within their environment (Hopkins, 2000). While further research is required to examine the algorithm within different unique scenarios, the current findings present an overall view of what practitioners and coaches can expect from the performance of the system.

Practitioners and coaches analyse the technical kicking actions as they tend to be associated with the ability to score more goals/ points than an opponent (Yi et al., 2019). The current study presents an option to monitor kicking actions of both match play and training automatically, without the reliance upon manual/semi-automated coding systems, which can require multiple analysts/ cameras (Liu et al., 2013) and take hours to code with the potential of human error (Rein & Memmert, 2016). With the use of automated tagging through machine learning algorithms showing >90% agreement, similar to that observed in human operators (Liu et al., 2015), automated tagging would provide a time saving solution for practitioners. The use of automated tagging systems would allow more time for coaches/ practitioners to gain insight into their data sets and action them, as opposed to spending time collating the data. However, practitioners should be aware that some scenarios performed better using these algorithms than others. Within unopposed drills, such as technical passing, the precision and recall statistics were close to 100%, whereas within opposed drills, that was reduced. Potential explanations for the opposed drill performance come during activities such as tackling and disguised touches (such as dribbling/ skills), which during the development process the labellers weren't able to clearly identify the associated touch or release. This is further apparent within previous research associated with semi-automated camera systems and seems to be a common issue within performance analysis (Liu et al., 2016). Playermaker further uses the receive release touches data to extract ball possession with a recall of 92% and precision of 88%, further study should look to examine these variables to assess individual ball possession and other technical activities during team and remote soccer activities.

5. Conclusion

Playermaker ball touch interactions have been shown as a valid method to monitor the technical kicking components during football (soccer) training and match play at different age groups. Automated machine learning algorithms performed equal to those previously observed within different human operators (Liu, 2015). The precision and recall of the algorithm do vary between opposed and unopposed drills, however, still showed an acceptable level of performance.

6. References

1. Lees et al., 2010 - https://www.ncbi.nlm.nih.gov/pubmed/20509089
2. Barnes et al., 2014 - https://www.ncbi.nlm.nih.gov/pubmed/25009969
3. Liu et al., 2015 - https://www.ncbi.nlm.nih.gov/pubmed/25793661
4. Yi et al., 2019 - https://www.ncbi.nlm.nih.gov/pubmed/31866914
5. Liu et al., 2016 - https://www.ncbi.nlm.nih.gov/pubmed/26613399
6. Rein & Memmert, 2016 - https://www.ncbi.nlm.nih.gov/pubmed/?term=6.%09Rein+%26+Memmert%2C+2016
7. Hopkins, 2000 - https://www.ncbi.nlm.nih.gov/pubmed/10907753
8. Barrett et al., 2014 - https://www.ncbi.nlm.nih.gov/pubmed/24622625
9. Boyd et al., 2011 - https://www.ncbi.nlm.nih.gov/pubmed/?term=The+Reliability+of+MinimaxX+Accelerometers+for+Measuring+Physical+Activity+in+Australian+Football
10. McGregor et al., 2009 - https://www.ncbi.nlm.nih.gov/pubmed/19806216
11. Vanrenterghem et al., 2017 - https://www.ncbi.nlm.nih.gov/pubmed/28283992
12. Needergaard et al., 2018 - https://www.ncbi.nlm.nih.gov/pubmed/30595981

A. **Appendix**

1) Precision Recall Example