Trading Places – Simulating Goalkeeper Performance using Spatial & Body-Pose Data

1. Introduction

Late in the 2018 Champions League Final between Real Madrid and Liverpool with the score 2-1 in favor of the Spaniards, in the 83rd minute Real Madrid's substitute Gareth Bale took aim at goalkeeper Loris Karius from 35 yards with a powerful yet straight strike. The rest is history, as the ball ended up sailing through Karius' hands effectively giving Real Madrid their third straight title. The reaction to the loss was immediate by Liverpool, with the club breaking the world record fee for a goalkeeper by purchasing Brazilian Alisson for £67m from AS Roma from the Serie A.

While this transfer triggered a flurry of other high-priced goalkeeper transfers between the top European leagues putting the cost of goalkeepers at an all-time high – it begs the questions: i) how can you compare the performance of different goalkeepers across teams and leagues?, and ii) how can you approximate whether or not a goalkeeper will be a success in your specific team? Currently goalkeepers are assessed using coarse metrics, such as "clean-sheets", "total goals conceded" or "shots saved to goals conceded" ratio. More recently "expected metrics" such as expected saves (xS) [1] have been introduced to compare performance to the league average. Problems arise with these methods because goalkeepers could have different types of saves to make depending on the style of the team and the opponents they face.

Instead of using metrics which may not capture all the different situations and contexts, *why can't we go beyond metrics and simply simulate each goal-keeper for every shot – then compare who would concede the least amount of goals?* Imagine if we could ask the question: ***"if Alisson played for Liverpool last year, how many goals would he have saved/conceded based on the shots that Liverpool faced during the season?"*** In this paper, we show how we can do this accurately which we validate across 150,000 shots. The result to the question is given in Figure 1.

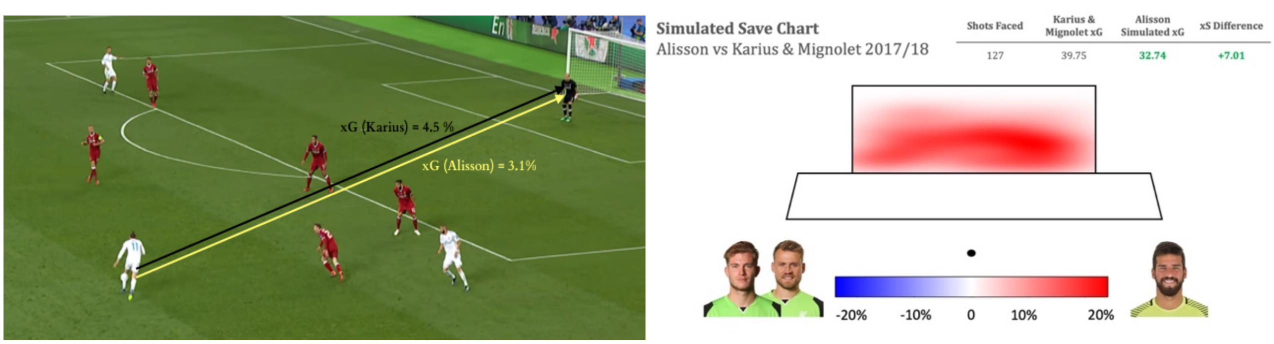

Figure 1: Left – simulated save likelihood of Karius (4.5%) vs Alisson (3.1%) for the goal by Gareth Bale in the 83rd minute. Right – Comparing Alisson vs Liverpool keepers for all shots in season 2017-18 shows that Alisson would have saved 7 more goals than Karius/Mignolet.

Even though simple in concept, in terms of being able to accurately simulate the "swapping" of different goalkeepers for specific situations is challenging due to:

1. **The lack of specific examples for each goalkeeper:** the task would be simple if each goalkeeper faced 1,000,000 shots per season – but given that each goalkeeper on average faces 2 to 5 shots on target per game (70 – 150 shots on target per season for a 38 game season), a goalkeeper may only face a couple of shots per location/context, or may not at all be based on whom they play for. For example, a goalkeeper who plays for a team that generally sits back deeply defensively may not face many counter-attacking shots, or another goalkeeper who plays on a team who is very strong on set-pieces, may not actually face many shots from set-pieces.

2. **The changing form of a goalkeeper:** Due to injury, fatigue, age, confidence, improvement in skill, coaching etc. - a goalkeeper's form may change across the course of a season and/or career, which may result in previous examples of goalkeeper saves being irrelevant (i.e., examples are not predictive of current or future performance).

3. **The data we have is not granular enough**: The observation for each shot we have is restricted to only the *x,y* position of the shot location, the *x,y* goal-keeper location at the time of the strike, the *x,y* final ball position (with the associated player identities). To accurately predict the likelihood of a goalkeeper saving a shot we would ideally know their body-pose position (i.e., whether they crouched, standing up straight/unbalanced, whether they are "big" (arms wide) as well as the strikers body-pose).

To address these challenges, we utilize a ***personalized prediction*** approach using dynamic spatial features within a deep learning framework. Specifically, we employ a feed forward neural network with a combination of fixed (i.e., shot and goal-keeper location) and dynamically updated (e.g. player form, time in game or score-line) embeddings and features to predict the chance of a shot being saved (*expected saves*), where a shot will be placed and critically allow the interchange between goal-keepers to compare performances in the same situation.

The intuition on why the above approach works is as follows. Firstly, the identity of a goalkeeper contains an enormous amount of information. For example, if we think of Gianluigi Buffon who is arguably the best goal-keeper in the world over the last 20 years, some attributes immediately standout: physically large, consistent, always in great position and immaculate technique for shot-stopping and set-pieces. Instead of explicitly having to capture and store those features observed for each save over Buffon's career, we can implicitly use Buffon's identity as a surrogate for all those attributes. This is beneficial as it is both low-dimensional (i.e., one feature) and contains granular information which we do not necessarily have to observe directly.

In Sections 2 and 3, we first describe our method, how it allows us to do more accurate shot prediction both at the save and save location level. We then showcase how it can be used to compare goalkeepers within leagues and across leagues as well as how it can be used to predict how a goalkeeper would perform on specific teams. Specifically, we give a ranked list of goalkeepers for across the top 5 European leagues (English, French, Spanish, Italian and German top divisions). We then highlight specific "what-if" questions: such as how would Alisson perform on Liverpool's team, as well as charting the journey of Joe Hart.

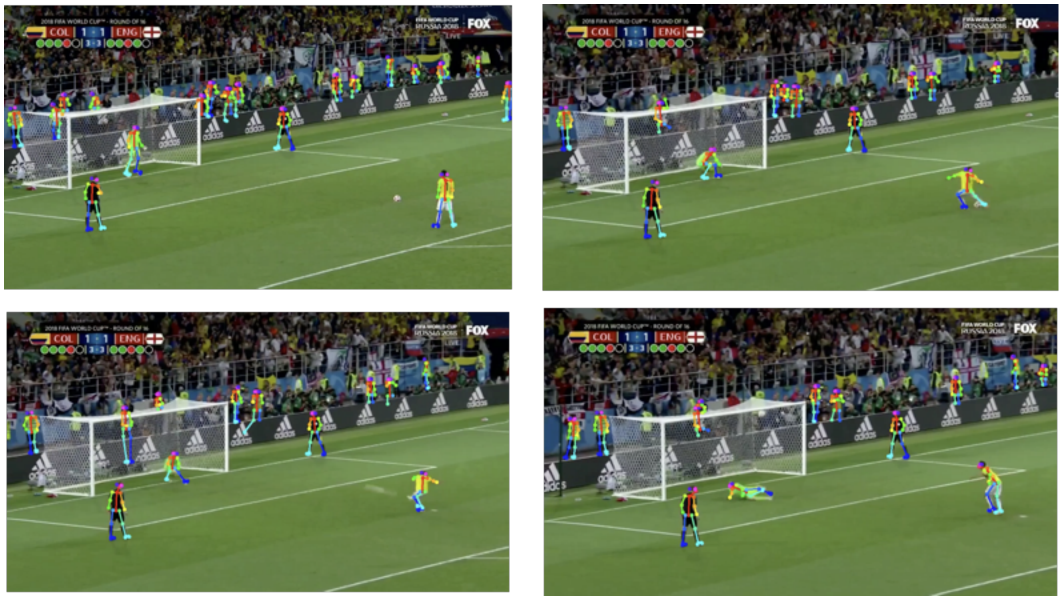

Figure 2: Using our body-pose technology, we can detect skeletons in video to capture the level of granularity required to do accurate prediction of shots. Above is an example of England's Jordan Pickford saving the penalty against Colombia's Carlos Bacca. In this paper we conduct body-pose analysis on all 61 penalties taken during the 2018 World Cup in Russia.

Even though utilizing "identity" information can enable more accurate prediction using this approach, it is suboptimal as it is still very much an approximation (i.e., over large amounts of data, it improves prediction). Ideally however, we would like to explicitly describe each save using fine-grain information such as body-pose. In Section 4, we show how we can utilize the latest computer vision techniques to capture the skeleton of goalkeepers and strikers directly from video. We show how we can capture goalkeeper posture and motion to enhance shot prediction. To highlight the utility of method, we run our analysis on all 61 penalties taken during the 2018 FIFA World Cup (see Figure 2 for an example of our analysis).

2. Personalized Predictions Utilizing Goalkeeper Identity

2.1 Dataset

To do our analysis, we have used 3 seasons worth of data (2016-2018) from 54 different leagues/competitions across the world with a total sample of 147,000 (43,000 goals, 104,000 saved) shots on target faced by over 2000 goal-keepers. We split our data into train/test set (80%/20%). Our raw features included the shot start (x, y) and end location (x, y, z) and goal keeper start position (x, y), time in game, half, score, venue and player identities. We then handcraft additional geometric features from the shot information including striker and goal-keeper angle and distance to the center of the goal and each other. The saves to shots on target ratio was 70.38% for the training set.

2.2 Previous Work

The ability to personalize content for end users has become a critical feature in numerous content generating social networks and search engines such as YouTube [2], Google Search [3] and Facebook [4]. The main issue is that the data available is highly sparse, meaning that while a large number of people may have used a service such as watching a music video on YouTube, the number of interactions that person has had compared to the actual amount of content is vastly inferior. As a result, we end up with a large and highly sparse matrix of interactions. This type of situation is highly challenging but has been solved by introducing the concept of wide and deep networks and embeddings. This approach has also been applied to sport with Alcorn [5] introducing a batter | pitch 2vec model which created pitcher batter embeddings in order to predict at-bat outcomes. Similarly, collaborative filtering techniques have been used in tennis [6] while deep reinforcement learning has been used in basketball to recommend how a defense may react to a drawn-out attacking play [7].

2.3 Method

One of the main reasons it is difficult to assess goalkeepers is the small number of shots a goalkeeper faces in a season. For example, a team pushing to win the league faces considerably fewer shots than a team fighting for relegation. Take Ederson who won the league with Manchester City last season. He faced just 67 shots on target conceding 29 (43%) compared to Ben Foster who was relegated with West Bromwich Albion who faced 127 shots conceding 50 goals (39%). On face value Foster would appear the better keeper but, how do we know which keeper faced the most difficult shots?

In order to provide a standardized unit, we first create an expected save model (xS). Given the shot location of the ball and the position of the goal-keeper, we predict the chance of the shot being scored or saved. We trained a four-layer feed forward neural network with hidden layers of sizes 12 and 8 with ReLU activation function and an output layer with a sigmoid activation function for the prediction task at hand (see Figure A1 in the Appendix for the network architecture). However, this would only give us the save likelihood for the average keeper. In order to make more personalized predictions, we need to capture the identity of the goalkeeper. In this paper, we take the approach of crafting a ***spatial descriptor of a goalkeeper***, to capture the influence of goal-keeper on the shot outcome. By crafting these descriptors, we are able to learn the latent space that best allows us to use a small sample of shots to create accurate personalized predictions.

These descriptors contain a large amount of information about a goalkeeper's strengths and weakness which include: clean sheet percentage, win percentage, save percentage for shots ending in the middle, left and right thirds of the goal and save percentage of shots that are struck directly at them, to the right or to the left of the goalkeeper. This spatial descriptor, which captures the identity of a goalkeeper, is what takes on the form of our embeddings and is dynamic in nature. We capture these embeddings on a season level and as a 10-game rolling window average to capture hot and cold streaks of keepers.

Features	Accuracy	Loss
Feature I	74.14%	0.53
Feature I, II	74.47%	0.52
Feature I, II, III	78.65%	0.46
Feature I, II, III, IV	**80.58%**	**0.43**

Table1: Accuracy and log loss for predicting goal scored using Shot Location (Feature I), keeper position (Feature II), score/time and final ball position before being scored/saved (Feature III) and dynamic keeper embedding (Feature IV).

2.3 Evaluation

From the results shown in Table 1, we can see that adding the keeper position does not significantly improve the loss and accuracy in predicting save likelihood. This is because most keepers adopt very similar starting positions when facing shots. It is obvious from these initial models that we need to incorporate goal-keeper identify in-order to boost prediction. To do this we learn a spatial representation of the keeper based on their historic ability to save shots based on the shot destination (left, middle or right segments of the goal) and their angle to the shot trajectory. These features capture keeper strengths and weaknesses on a high level and act as proxies for factors like handedness (weaker/stronger towards left/right), anticipation, and reaction time for shots of varying angles and destinations. Adding these dynamic keeper embedding features significantly improves loss and accuracy of our model.

3. Goalkeeper Rankings and Analysis

3.1 Goalkeeper Rankings Per League

To demonstrate our ability to accurately simulate goalkeeper skill we pose the question *what if* a *goalkeeper faced every shot taken, how many goals would they prevent compared to everyone else?* To do this we first simulate the number of goals an average keeper would have conceded. We then selected every goalkeeper who faced >60 shots from the *"Big Five Leagues"* in Europe for the 2017/18 season and simulated their xS value by swapping in their dynamic embedding and then subtract the difference.

Top 10 Keepers

Goal-Keeper	Team	Goals +/-
Jan Oblak	Atletico Madrid	0.98
David De Gea	Manchester United	0.74
Samir Handanovic	Inter Milan	0.72
Pau Lopez	Real Betis	0.68
Ron-Robert Zieler	VFB Stuttgart	0.60
Marc-Andre Ter Stegen	Barcelona	0.59
Neto	Valenica	0.59
Jiri Pavlenka	Werber Bremen	0.59
Nick Pope	Burnley	0.43
Regis Gurtner	SC Amiens	0.41

Bottom 10 Keepers

Goal-Keeper	Team	Goals +/-
Raul Lizoain	Las Palmas	-0.49
Bingourou Kamara	RC Strasbourg	-0.52
Eiji Kawashima	RC Strasbourg	-0.54
Vid Belec	Benevento	-0.56
Simon Mignolet	Liverpool	-0.60
Alex McCarthy	Southampton	-0.60
Geronimo Rulli	Real Sociedad	-0.63
Heurelho Gomes	Watform	-0.79
Sergio Rico	Sevilla FC	-0.88
Joe Hart	West Ham Uinited	-1.19

Table 2: Top 10 vs bottom 10 keepers across the Big 5 Leagues in Europe for the 2017/18 season (England, Spain, Italy, Germany, France).

We showcase the top and bottom 10 keepers across the leagues in Table 2 showing the number of additional goals each goalkeeper would have saved or conceded compared to average. We can clearly see that Manchester United's David De Gea (+0.74) and Atletico Madrid's Jan Oblak (+0.98) were the most outstanding keepers whilst Sevilla's Sergio Rico (-0.88) and West Ham United's (on-loan) Joe Hart (-1.19) were the leakiest. Interestingly both Hart and Rico were signed by Burnley and Fulham respectively, who have conceded the most goals in the 2018/19 English Premier League up to game week 15.

3.2 Dynamic Embedding Clusters

As stated earlier, if our dynamic embeddings capture differences between keepers, we should be able to see significant separation in our dataset and more importantly we should see elite shot stoppers in one cluster and poor shot stoppers in another. Due to the high dimensionality of the embeddings we apply t-SNE [8] multi-dimensional reduction technique to discover unique clusters. Figure 3 demonstrates that the embeddings generate two well separated clusters. We visualize where the top and bottom rated keepers from Table 2 sit within the embedding space. We can clearly see that the dynamic embeddings do a good job at separating goalkeepers into different skill groups with the top-rated keepers (green) being in the top cluster and the lower rated goalkeepers (red) in the bottom cluster.

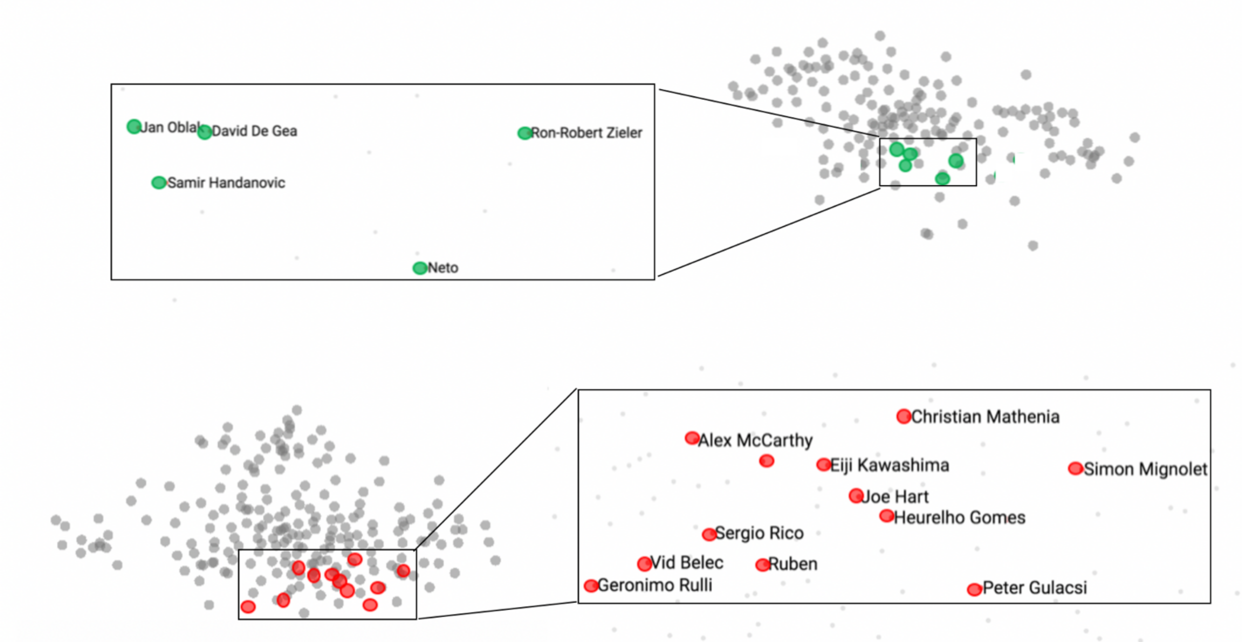

Figure 3: Applying 2D t-SNE to the dynamic embeddings creates two highly separated clusters goal-keeper. Zooming in (top right / bottom left) we demonstrate the embeddings separate goalkeepers into skill groups by visualizing the top and bottom keepers from Table 2.

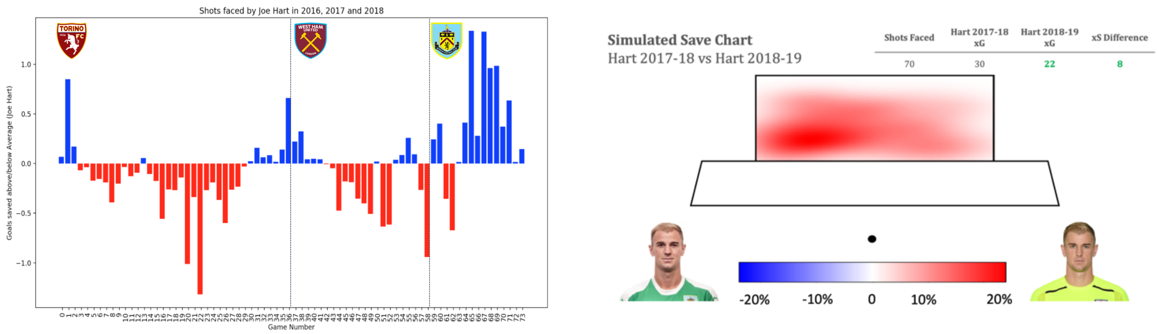

Figure 4: (Left) How Joe Hart has performed over the past 3 seasons compared to average. Time line is split by the time spent at clubs. (Right) Simulated save map comparing how many more goal Hart would have saved the previous season given his current form.

3.2 Trading Places: Simulated Goalkeeper Swapping

Given that are our embedding space enables us to make more personalized xS predictions, we can now swap keeper identities and simulate the behavior of other keepers across any shot within our dataset. To demonstrate this, we use the example of Liverpool signing Alisson from Roma. As previously stated, Liverpool spent a world record fee of £67m to secure Alisson's services but how could they be certain he would be able to transfer his form from Roma to Liverpool?

With the help of our model, we can now the swap the identities, aka the spatial descriptors, of Liverpool keepers, Loris Karius and Simon Mignolet, with Alisson and simulate his behavior for the shots faced by Liverpool. Figure 1 shows the simulated shot map, which is a weighted 2D gaussian distribution of where Liverpool conceded shots for the 2017/18 season. Each shot is weighted by the difference in the xS between the keepers. Red shows where Alisson increased the chance of saving a shot and blue shows where Karius/Mignolet increase the chance. As we can see ***no part of the simulated save chart is blue***. What is truly revealing is that the center area of the goal see's the largest increases in xS (>20% improvement). Taking every shot into account, had Alisson played for Liverpool for the 2017/18 season, they could have expected to conced 7 fewer goals.

3.3 Form is Temporary, but Class is Permanent

We previously highlighted that Joe Hart was one of the lowest performing goalkeepers in the big 5 leagues for the 2017-18 season. However, is this permanent or does form evolve over time? Because we learn an embedding that is dynamic in nature, we are able measure how a goalkeeper changes from season to season. In Figure 4 Left, we show Hart's performance at a game level over the last 3 seasons compared to the average keeper. We can clearly see that his form as varied from season to season where he has had periods of high performance followed low. 2018/19 has seen his most sustained level of high performance with Burnley, indicating that he might be justified to be playing ahead of Pope who was rated 9th in Europe (Table 2) and second in the 2017/18 English Premier League (see Table A1 Appendix).

We can also look and see how much a keeper may have improved by asking '*what if I faced the same shots in my current form?'*. To demonstrate this, we compare Joe Hart in 2017-18 season vs Joe Hart in the 2018-19 season. Our model predicts that if Hart had faced the same shots now that he faced last season, he would have prevented an addition 8 goals. Looking at his simulated save chart in Figure 4 Right we can see that he has improved significantly across the goal but mainly in the bottom left corner.

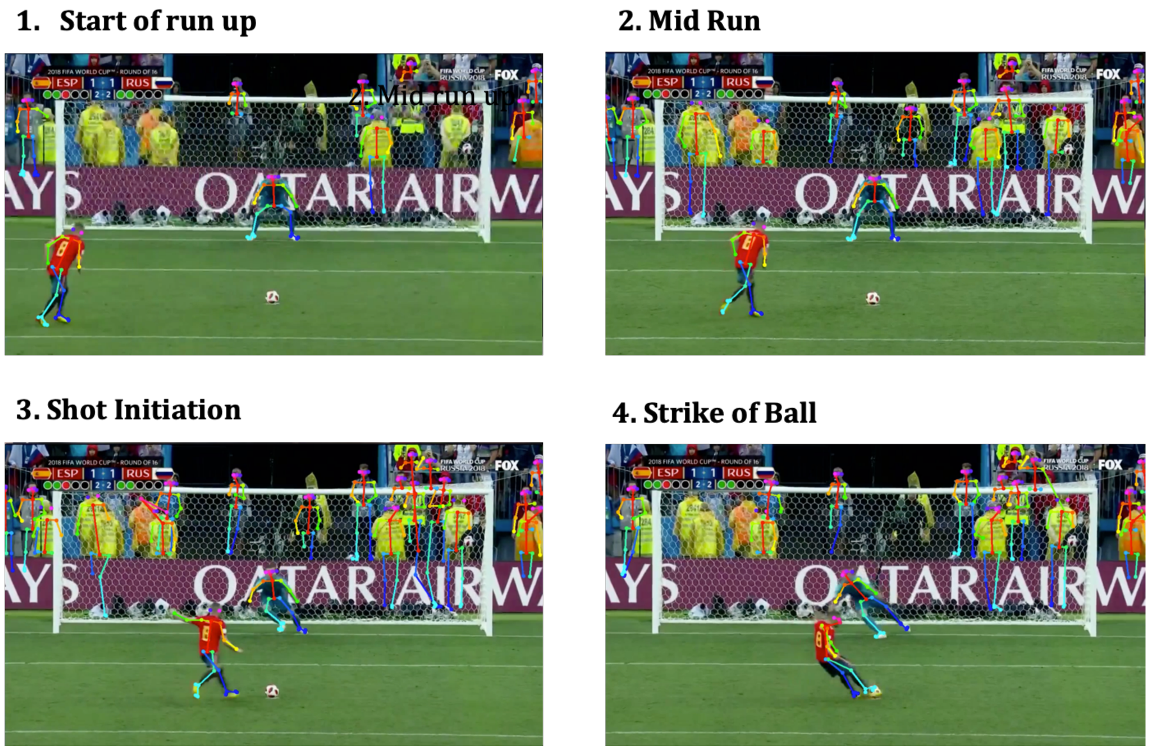

Figure 5: An example of the three phases analysed for the striker run up. We can see that the goalkeeper has initiated his movement at phase 3.

4. Predicting Penalties from Body-Pose

A penalty is the most controlled scoring situation in football which massively favours the striker with penalties being saved 30% of the time (the most converted shot type in football). However, certain keepers show an uncanny ability to save these shots. For example, Real Madrid's Keylor Navas has a 66% save rate, therefore what differentiates keepers from each other?

To be able to determine this we need to go beyond spatial-data and use more fine-grain body pose data. Body-pose has previously been applied to basketball [9] to find physical attributes that differentiate between successful and unsuccessful 3-point shooters. We apply this technique to determine what pose attributes best predict where the striker will place the ball. For simplicity we focus the task on predicting if the ball will be shot to the left, right or centre of the goal.

4.1 Detecting Striker Intent for Body-Pose

On average goal-keepers chose the right direction 47.5% of the time, however, if a goal-keeper was to use the most naïve model of picking the most likely side the ball will be placed given the starting position of the penalty-taker, the keeper would be correct 60.66% of the time. This indicates that most goal-keepers either guess where the shot will be placed as opposed to being able to read the key tells from the striker.

While a striker aims to hide their true intention of where they will place the shot, we hypothesise that there are certain 'tells' a striker gives during the run up and shot phase before they make contact with the ball.

Our data reveals that 88.52% of the time the keeper has started their primary move to save the penalty before the ball strike. This makes sense as while they would gain maximum information to predict where the shot will be place, due to the speed of the shot they would not have enough time to react. Therefore, in order to detect these tells, we split the penalty into three phases (see Figure 5 for details):

1. **Start of Run Up:** Start position and angle
2. **Mid Run:** Run type e.g. stutter and speed
3. **Shot Initiation:** Body lean angle, upper body angle, hip orientation, none kicking arm position and shoulder alignment.

4.2 Won't Get Fooled Again

For our analysis we predict two tasks, which segment the ball be shot at and which direction will the keeper dive. The second task in particularly interesting as it provides evidence as to what goal-keepers may be focusing on and how a striker can fool a goal-keeper.

Our results demonstrate that the most predictive features to where a shot will be placed are; run up location and run up type during the run up phases and body lean angle and hip orientation at shot initiation giving 67.86% accuracy. Conversely, when predicting the most likely direction a keeper will dive, we find that the run up location and type, combined with the body lean angle, none kicking arm position and shoulder alignment give an accuracy of 64.29% (Figure 6). This truly demonstrates that the hips don't lie when coming to predicting penalties and that goalkeepers can be easily tricked by how the striker holds their arm before striking the ball.

Figure 6: Strikers hip angle is open towards the right-hand post indicting a larger probability of shooting to the right. However, the left arm being held wide, which the keeper takes his cue from.

5. Summary

In this work, we introduce a personalized prediction approach to analyze goalkeepers which goes beyond metrics like "clean sheets" and "total goals conceded". We use a dynamic spatial descriptor to capture goalkeeper identity which allows us to simulate the performance *for any goalkeeper given any shot across all teams and leagues*. Using this approach allows us to "swap" goalkeeper identity and accurately compare them despite the lack of specific examples for each goalkeeper. Since the embeddings are dynamic in nature, they can also accurately capture the changing from of a goalkeeper. Finally, we demonstrate the ability to use fine-grained body-pose data to more accurate predict striker intent during penalty kicks.

References

[1] Colin Trainor, Goalkeepers: How repeatable are shot saving performances?, 2014 https://statsbomb.com/2014/10/goalkeepers-how-repeatable-are-shot-saving-performances/"

[2] https://www.theverge.com/2017/8/30/16222850/youtube-google-brain-algorithm-video-recommendation-personalized-feed

[3] Paul Covington, Jay Adams and Emre Sargin, Deep Neural Networks for YouTube Recommendations, RecSys '16 September 15-19, 2016, Boston, MA, USA

[4] Heng-Tze Cheng et al., Wide & Deep Learning for Recommender System, arXiv:1606.07792v1, 2016

[5] Michale Alcorn, (batter|pitcher)2vec: Statistic-Free Talent Modeling With Neural Player Embeddings, 2017. MIT Sloan Sports Analytics Conference.

[6] "The Thin Edge of the Wedge": Accurately Predicting Shot Outcomes in Tennis using Style and Context Priors, 2016. MIT Sloan Sports Analytics Conference.

[7] Bhostgusters: Real-time NBA Play sketching with synthesized NBA defenses, 2017. MIT Sloan Sports Analytics Conference.

[8] L.v.d. Maaten, G. Hinton. Visualizing data using t-SNE, Journal of Machine Learning Research, Vol 9(Nov), pp. 2579—2605. 2008

[9] "BodyShots": Analyzing Shooting Styles in the NBA using 3D Body-Pose Information, 2016. MIT Sloan Sports Analytics Conference.

Appendix

La Liga

Player Name	Team	Goals +/-
Jan Oblak	Atletico Madrid	1.03
Pau Lopez	Real Betis	0.71
Marc-Andre Ter Stegen	Barcelona	0.58
Neto	Valencia	0.54
David Soria	Getafe CF	0.32
Antonio Adan	Real Betis	-0.54
Ruben	CA Osasuna	-0.54
Raul Lizoain	Las Palmas	-0.62
Geronimo Rulli	Real Sociedad	-0.64
Sergio Rico	Sevilla FC	-0.94

Premier League

Player Name	Team	Goals +/-
David De Gea	Manchester United	0.81
Nick Pope	Burnley	0.52
Adrian	West Ham	0.28
Wayne Hennessey	Crystal Palace	0.27
Hugo Lloris	Tottenham Hotspur	0.26
Jordan Pickford	Everton	-0.19
Alex McCarthey	Southampton	-0.42
Simon Mignolet	Liverpool	-0.45
Heurelho Gomes	Watford	-0.62
Joe Hart	West Ham United	-0.99

Bundesliga

Player Name	Team	Goals +/-
Ron-Robert Zieler	VFB Stuttgart	0.63
Jiri Pavlenka	Werder Bremen	0.6
Marwin Hitz	FC Ausburg	0.31
Philipp Tschauner	Hannover 96	0.22
Yann Sommer	Borussia Munchengladbach	0.14
Bernd Leno	Bayer Leverkusen	-0.19
Roman Burki	Borussia Dortmund	-0.2
Timo Horn	FC Koln	-0.21
Peter Gulacsi	RB Leipzig	-0.44
Christian Mathenia	Hamburg SV	-0.48

Ligue 1

Player Name	Team	Goals +/-
Regis Gurtner	SC Amiens	0.4
Anthony Lopes	Olympic Lyon	0.31
Alphonse Areola	Paris Saint-Germain	0.28
Tomas Koubek	Stade Rennais	0.24
Ciprian Tatarusanu	FC Nantes	0.2
Karl-Johan Johnsson	Guingamp	-0.18
Benoit Costil	Bordeaux	-0.19
Baptiste Reynet	Toulouse	-0.27
Eiji Kawashima	RC Strasbourg	-0.43
Bingoutor Kamara	RC Strasbourg	-0.44

Serie A

Player Name	Team	Goals +/-
Samir Handanovic	Inter Milan	0.63
Alisson	Roma	0.33
Mattia Perin	Genoa	0.28
Alfredo Gomis	SPAL	0.26
Antonio Mirate	Bologna	0.22
Albano Bizzarri	Foggia	-0.17
Marco Sportiello	ACF Fiortentina	-0.23
Andrea Consigli	Sassuolo	-0.25
Emiliano Viviano	Sampdoria	-0.3
Vid Belec	Benevento	-0.62

Table A1: Top and Worst 5 goalkeepers across the big five leagues (La Liga, English Premier League, Bundesliga, Ligue 1 and Serie A) in the 2017-18 season.

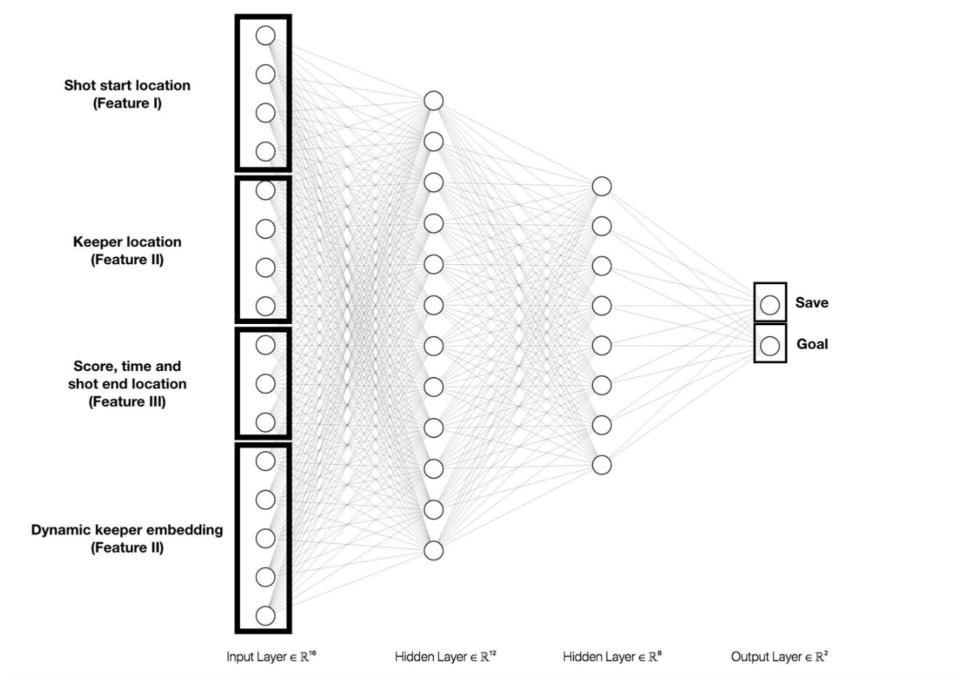

Figure A1: Feedforward neural network architecture used for dynamic embeddings.