Rugby Field Registration with Hough Line Detection

by

Tom Keane

20214537

University of Limerick, 2020

Master of Science

In

Artificial Intelligence

(Department of Computer Science)

Supervised by Dr. Tony Scanlan

Rugby Field Registration with Hough Line Detection

1. Abstract

Sport field registration (SFR) and projection to a template field map can provide unique insights to supplement sport analytics and coaching. SFR and projection have been utilized to various degrees across different sports, including football, American football, basketball, and ice-hockey. The challenges imposed on SFR differ in scope and qualitatively across and within sports due to variations in pitch dimensions, field markings, application of camera pan & zoom, among others. This dissertation proposes a sport field registration model based on Canny edge detection and Hough line transformations of a rugby broadcast footage, and then contextual semantic rugby field line classification. Classified lines are used to provide a homography matrix that projects player positions to a template field map. Player positions are provided prior to the homography transformation by an Ultralytics YOLOv5 pedestrian detector which is run on each broadcast frame.

Keywords: Sport field registration, Hough line transformation, Homography matrix

2. Content

3. Introduction

Professionalism in rugby union has created a competitive environment with demands for sports-analysts to create systems for talent identification, study tactical systems and track player actions. Rugby union was a pioneering sport in the use of GPS tracking, following a study published by (Ishii, et al., 2002). This study simply tracked the distance covered by referees in rugby and football matches. In recent years, GPS has been used in both professional and amateur rugby to provide a wider range of insights on player performance, such as the "high-speed running distance/time" of individual players, "repeated high intensity exercise bout", and "high metabolic load distance", among others (World Rugby, 2022).

Applications of GPS tracking in football and rugby, in conjunction with micro-electromechanical sensors (MEMS) such as accelerometers and gyroscopes technology, tend to concern monitoring and managing player workloads in training and matches. This narrow mode of analysis is reflected in the academic texts on the subject, which concentrate on physical demands put on players in training and matches, but do not discuss in-game tactical formations or predictions.

(De Silva, et al., 2018) – a study which "statistically compares the high-speed running activity demands between training and matches for different positions of play", concludes that "data can be utilized to understand the physical activity demands placed on players during training sessions and in competitive matches".

(Hennessy & Jeffreys, 2018), which provides an extensive analysis of the uses of GPS and MEMS technology in football describes the software as a "physical workload-monitoring system" as opposed to other technical and tactical elements.

Limitations regarding the usefulness of GPS player tracking are found in cases where its usage is unfeasible, too expensive or if access to the GPS data is restricted. It is often the case that GPS cannot provide insights into static in-game situations because users may only have access to data pertaining to their own organizations and not opposing teams.

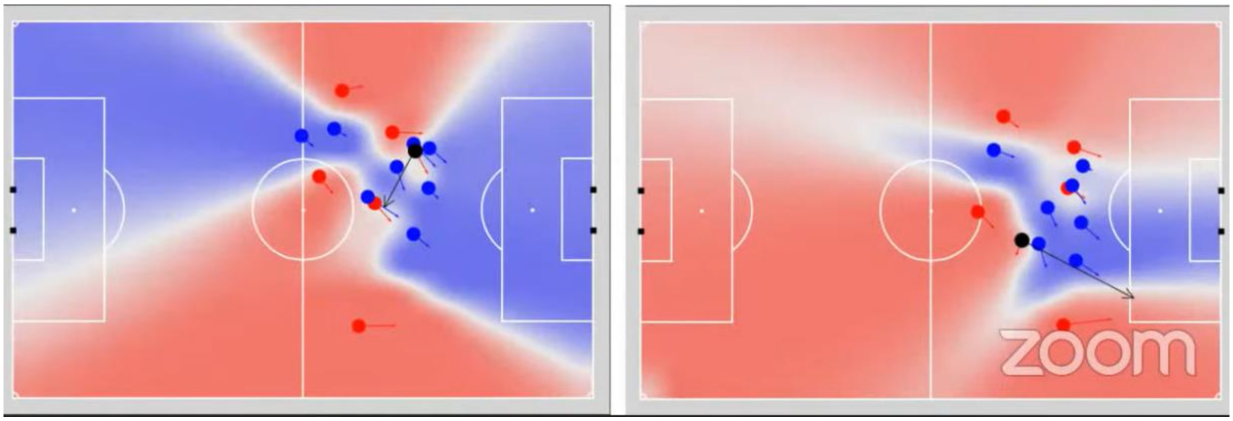

Figure 1 - A pitch control model, achieved through analysis of computer vision player tracking by Twelve Football, (Jernej, et al., 2021)

This dissertation proposes a cost-effective alternative method for player tracking in rugby matches, based solely on the input feed of broadcast footage taken from the halfway line of an international test match. This methodology seeks to evaluate the location of every person detected in-frame, and project the players positions to a top view projection, which may contribute to deeper tactical insights, such as pitch control estimations as seen in Figure 1.

While this model is only adapted for field registration for one test match, there is potential for future development and generalization in terms of adapting to differences in perspective and shade for broadcast footage from other stadia, or varying degrees of lens distortion at the edge of frame. This model is also only capable of field registration when there are four distinguishable major pitch lines, with which a homographic transformation can be performed.

It is the norm that international rugby test matches are filmed from two angles, with a large amount of camera zooming and panning but with minimal trucking – movement left or right without rotation. Large field sports, such as Gaelic games, football, and American football typically feature a large amount of panning and zooming. This poses a challenging hurdle in relation to field registration due to large deviations in pitch homography throughout the course of a video. Therefore, dynamic landmark detection systems are required to overcome these issues.

Formally, this dissertation makes the following contributions:

1. A novel framework for rugby pitch registration which identifies and classifies major pitch lines in each frame in a sequential order to build context and inform line refinement throughout the line classification and merging processes.
2. A less labor-intensive process to gather pitch landmark image datasets through corner detection of major field lines.

Key components for the field registration process performed in this dissertation are **line detection**, **line classification** and **field homography estimation**. This document describes a process for achieving these tasks with just a video input feed from the halfway line of an international rugby stadium. This includes sections on prior work relating to field registration, the applied methodology and results, followed by conclusions and discussions.

- Related work done in field registration focuses on large, outdoor field sports such as football, and describes how the key components outlined above were achieved.
- The methodology description explains the process of image preprocessing for line detection and player detection, followed by an in-depth explanation of how lines are differentiated into 6 classes, and used to provide a homographic estimation. There is also a supplementary section on how players are detected and classified by their team colors.
- The results section provides a series of images that represent the performance of the field registration model, detailing issues causing inaccurate or impossible field registration and poor player classification.
- Conclusions give a review of the method's overall effectiveness and further developments which can be made for field registration in rugby.

4. Related Work

Top view data analysis has grown in prominence throughout the last decade in team field sports. As a result, there have been multiple approaches to solve the problem of field registration across different sports, with varying functionalities and modules regarding feature detection, homography estimation and player tracking.

In order for field registration to occur, a pitch homography must be recognized by the detection of landmarks in-frame which can then be transformed to a field layout model (Equation 1, Figure 2). This was relatively trivial for (Bialkowski, et al., 2013), (Bialkowski, et al., 2013) and (Liang, et al., 2020), who solved this task in hockey, football and basketball respectively by each availing of eight fixed-in-place HD cameras. (Pers & Kovačič, 2000) also made an early contribution to the field which used two fixed-in-place cameras above a handball court. These scenarios are ideal, as their homographies are constant and can be preset.

i. Feature Detection

In field sports, the most generalizable landmarks that are useful for registration are pitch markings, such as lines and circles (Kim & Hong, 2001), (Hayet, et al., 2004, p. 4), (Gupta, et al., 2011).

(Farin, et al., 2004) provided a detailed outline of the line extraction process for registering a tennis court. This filtered white pixels to isolate tennis court lines, before a Hough transform is applied to this filtered image. (Farin, et al., 2004, p. 6) describes a method of overcoming line clusters by drawing a line of best fit in these regions of clusters. This was a useful solution for merging line clusters appearing around concentrated areas, because wide white court lines were interpreted as multiple parallel lines by Canny edge detectors and Hough transforms.

$$s\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \mathbf{H}\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Equation 1 - Homographic transformation equation

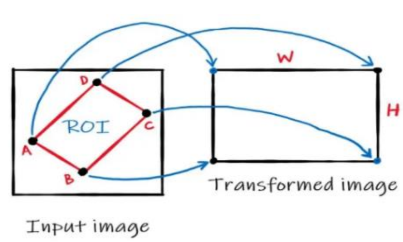

Figure 2 - Homographic transformation (Atul, 2019)

ii. Homography estimation in Large Field Sports – state of the art

Once features are detected, their relationship to each other in-frame must be translated into a homography matrix, which will transform the input pixels to a new perspective – a top-down view.

A large field sport which has received significant contributions to field registration is American football. (Fern & R, 2007) made an early contribution which uses the local distinctiveness of image features to match correspondences with a predefined static model to produce a homographic transform. This method required the creation of a dataset of about 12,000 labelled features as pitch and image landmarks.

Football – soccer, seems to have made the most progress in this area. (Sharma, et al., 2018, p. 3) used k-NN search to match input images to a selection from a dictionary of synthetic edge maps with corresponding homographies. (Jiang, et al., 2020, p. 3) used a two-stage regression framework which makes an initial homography estimate, and then warps and concatenates the input images upon a template pitch map until an error function minimizes based on line equations from the input image and The World Cup homographies dataset (Homayounfar, et al., 2017, p. 1). (Garnier & Gregoir, 2021, p. 5) built a keypoints dataset and a homography dataset from existing 2014 World Cup dataset and images from the Premier League, Ligue 1, and La Liga, which were used to train a Resnet-18 model which estimates the coordinates of four control points in input images, similarly to (Jiang, et al., 2020, p. 4).

The 2014 World Cup homographies dataset, built by (Homayounfar, et al., 2017) has contributed to many of the aforementioned neural-network solutions to homography estimation (Jiang, et al., 2020), (Garnier & Gregoir, 2021). As there are no existing datasets describing the homography of rugby pitches, this dissertation has to resemble the method by (Homayounfar, et al., 2017, p. 5), which classifies lines and pitch landmarks based on line equations.

iii. Sport Field registration without dedicated datasets

While (Homayounfar, et al., 2017) employed a method involving a 16-layer VGG network and extensive labelling of homographies to frames across 20 football games, it also provides quite a detailed description of how lines and shapes were classified when observing broadcast footage of a football pitch. Lines were split into six classes and two side-circles and a center-circle were also potential classes, which were then be used to inform registration. However, the labelling performed in (Homayounfar, et al., 2017) is a task too laborious for consideration in this dissertation

Detected lines were scored based on how much each of them corresponded to the vanishing points, which were in turn predicted by the line voting procedure of (Hedau, et al., 2009). In the absence of time and resources to spend on labelling hundreds of images with homographies and keypoints, an alternative method to the VGG network of (Homayounfar, et al., 2017) had to be chosen for this dissertation.

(Kim & Hong, 2001, p. 4) employed an inter-image homography estimation method based on matching points or lines from one image to the next, and then camera parameters are calculated from the estimated homography and (Hayet, et al., 2004) similarly combined this strategy with two others. To enhance processing speed, a feature tracking model was also employed on a selected area of the frame to quickly predict the homography. If too few features were detected in two successive frames, the entire interframe homography was estimated instead.

5. Material

International test rugby games require home teams to provide 5mb/s 720-pixel match footage from multiple angles, including one from the halfway line. Footage from the halfway line of an Ireland vs England RBS Six Nations match was used in this dissertation.

This dissertation was carried out in with a google colabs runtime environment, running python 3.7. For field registration, pyfld from pypi.org was used for Hough line detections and ultralytics pretrained yolov5 object detector was used to track players in-frame. Image pre-processing and transformations were carried out with opencv-cv2.

6. Methodology

Transforming an input feed to a top-down projection was carried out using a process involving two modules (Figure 5). An input image, (Figure 3) is passed to a preprocessing stage, and then handed to two separate modules. Based on the processed image, one of these modules estimates the pitches homography and the other detects players in-frame. The output of these modules is then used to create a top-down projection of the pitch and players positions, (Figure 4).

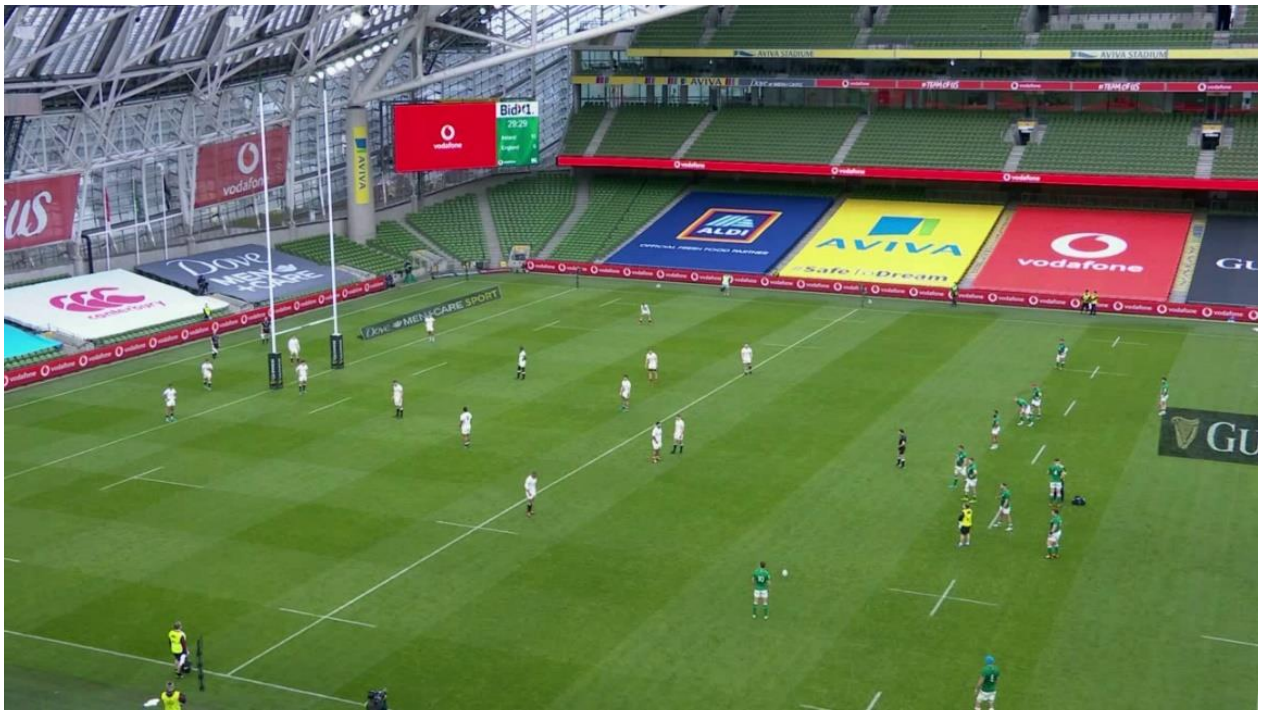

Figure 3 - Input Frame

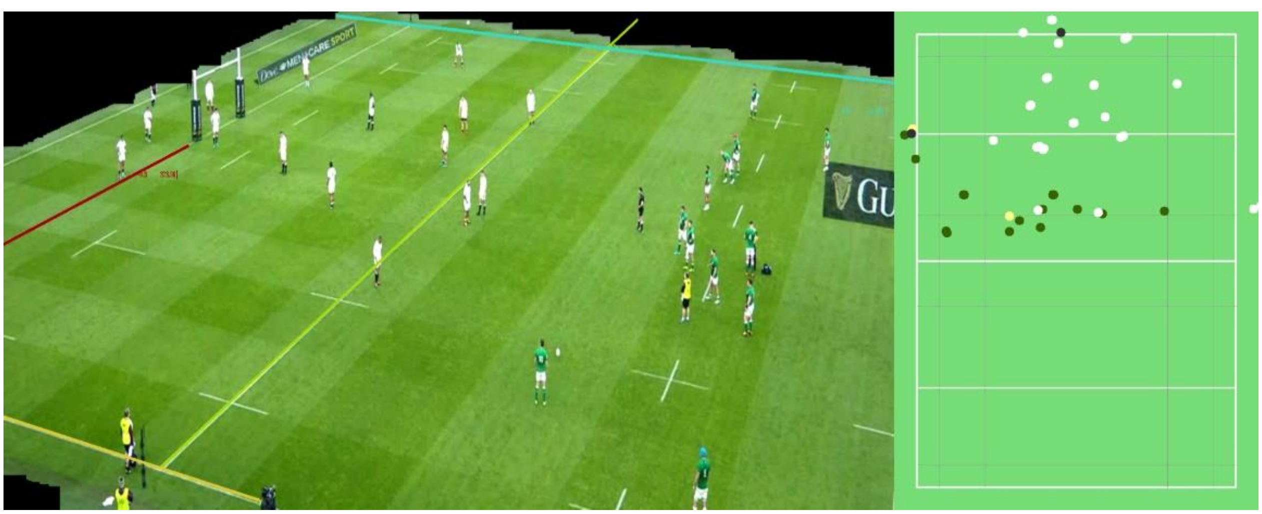

Figure 4 - Rugby Pitch Detection output

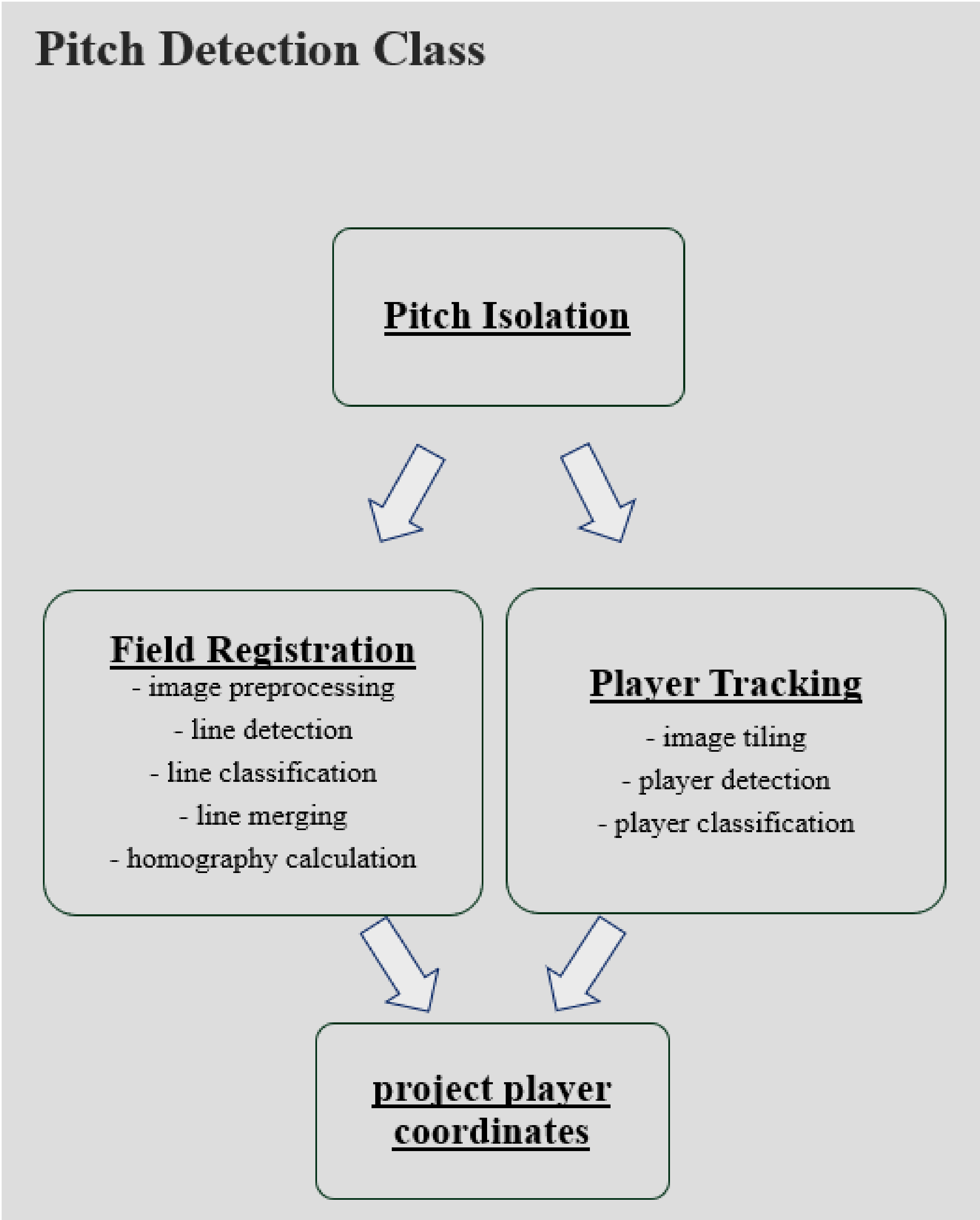

Figure 5 - Rugby Pitch Detection class process

i. Pitch Isolation

Frames must undergo preprocessing stages before field registration and player tracking. This is performed in a Pitch Isolation module. This module is built to create a mask that hides everything besides the area of the pitch. Because player detection and field registration have different requirements regarding what is visible in-frame, different image preprocessing steps are carried out for each.

Prior to isolation, pixel contrasting is increased to reduce probabilities of green colour bands overlapping with other objects in the stadia, such as stands, jerseys, etc.
Isolation was carried out by isolating 3 shades of green with varying brightness, followed by cutting off small ranges of green that are found in stands and advertising boards (Figure 8). Three bands of green were enough to capture the vast majority of play area, while keeping the BGR colour band ratios relatively tight. Care must be taken to exclude shades of green that may be found in other areas, such as stadium seating, etc. (Figure 6).
Then, an initial round of morphological operations is carried out to ensure that all the lines on the pitch are visible (Figure 7).

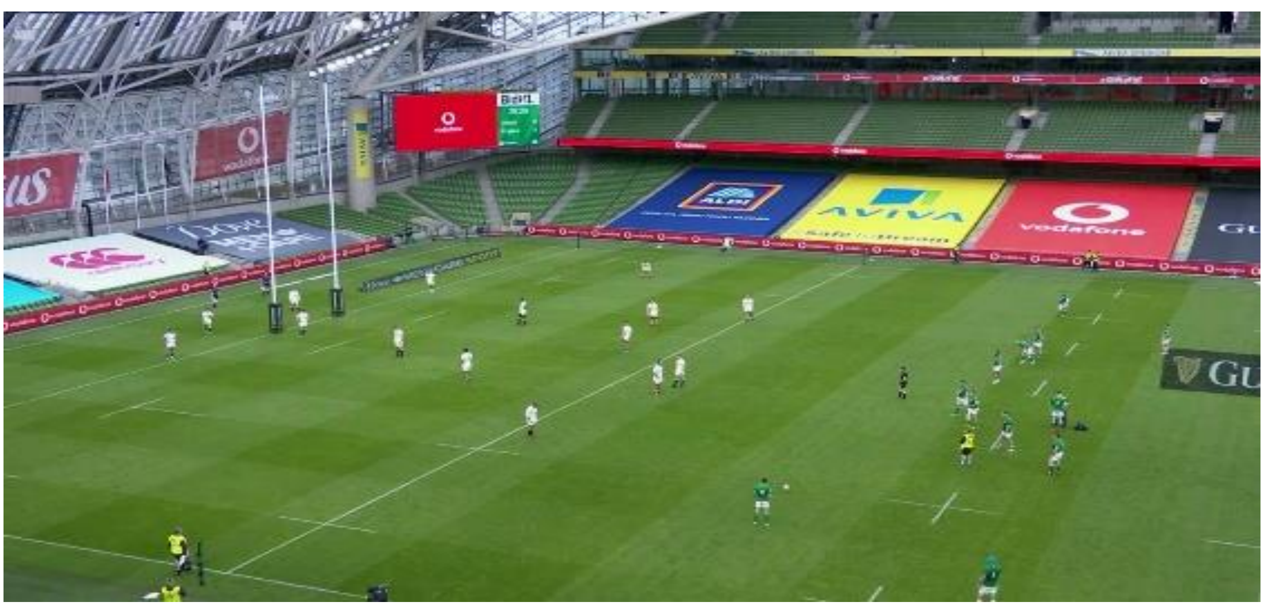

Figure 6 - unprocessed image.

Figure 8 - Grass-green has been isolated, but some off-pitch areas are still visible.

Some morphological closing will be necessary for full pitch isolation.

Figure 7 - This frame is ready to be passed to either one of the field registration or player tracking modules

Next, the image is cropped and resized to a wide landscape 2560x800 shape. Consistent image shape is important for line classification later, as it ensures that line equations keep to a consistent scale throughout the match video.

Next the image is duplicated because the field registration module and the player tracking module have separate requirements for what is visible in-frame, i.e. The player detection frame requires further image closing to ensure that all players are visible, whereas morphological erosion should be kept to a minimum for line detection to avoid interference from objects off the field, such as advertising boards.

For field registration, it was also necessary to mask specific colours after the first round of morphological operations, as advertising boards may still be visible.

ii. Field Registration

As seen in Figure 5, the field registration process is broken down to five key stages.

1) Hough line image pre-processing
2) Hough line detection
3) line classification
4) line merging
5) homography calculation

1) Hough line image pre-processing

Canny edge detection and Hough line detection require grey-scale inputs. This is a short process which converts the image to greyscale and applies some slight contrast modification. Experimentation was done with thresholding after greyscale conversion, but rugby pitch markings vary too much in brightness for consistent results. Bright patches of grass may exceed brightness of lines on the far side of the pitch.

2) *Hough line detection*

This was achieved with the (Tsukada, 2021) fast line detector - pyfld. This carries out Canny edge detection and then line detection by calculating the collinearity of pixels of neighboring edges and merging edges based on curvature and distance thresholds.
Increasing the canny aperture (kernel) size from default 3 to 7 improves detection of faint lines on the far side of the pitch. This is important for cases with blurring due to quick camera rotations.
Immediately after detection, any lines shorter than a low threshold were removed, as these are typically not found along significant pitch lines. Following all this, the line-state of the image should resemble Figure 9.

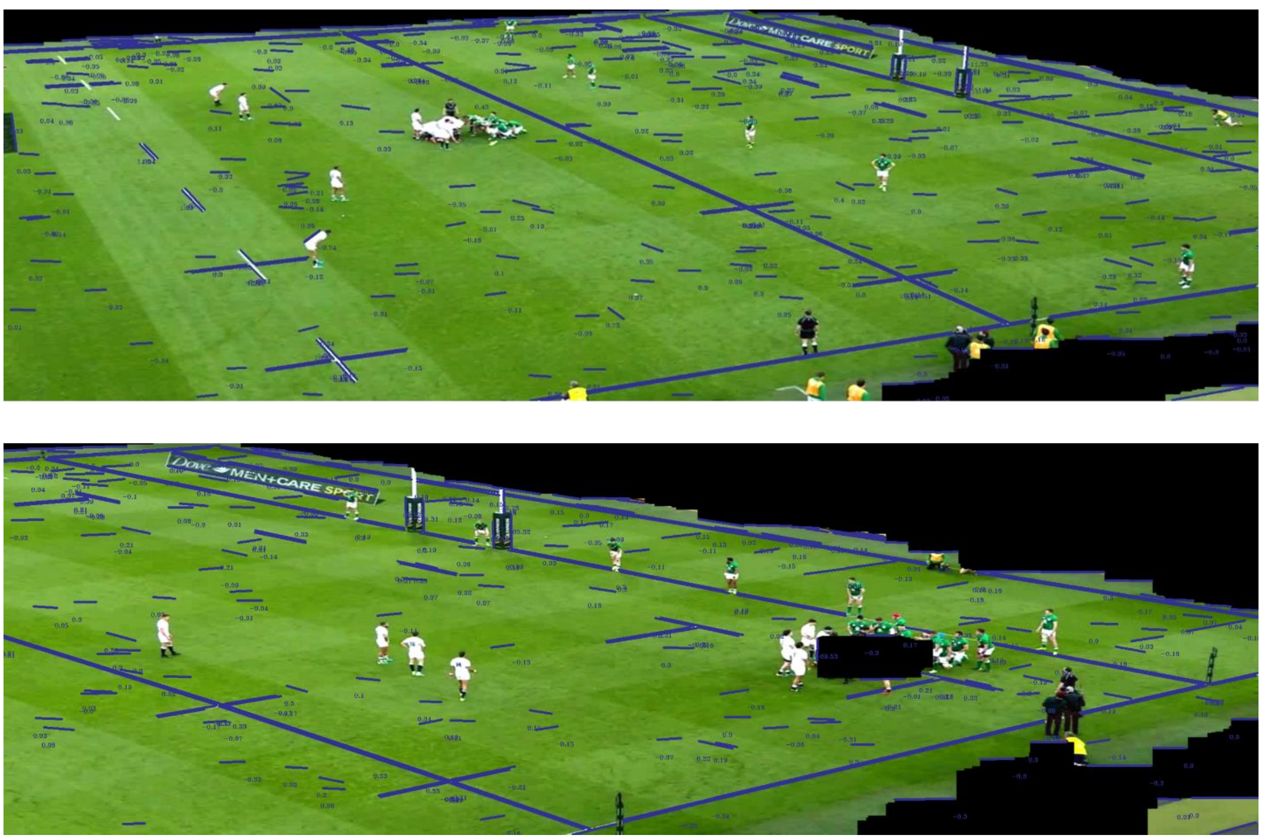

Figure 9 - detected lines after removal of all lines smaller than 0.05 x longestLine. Line slopes are printed beside all lines.

3) Line classification

The following process describes lines found in an image of shape 2560x800 and when there are two cross-pitch lines and two up-pitch lines visible – Up pitch lines are the near side lines (NSL), and the far side lines (FSL). Cross pitch lines are the halfway, twenty-two (22m) meter lines and try lines. The purpose of this stage is to cycle through all of the detected lines and to reduce them to groups of the best candidates approximating all of the major field lines.

Figure 10 - frame nearly bisecting pitch.

Cross pitch lines – halfway, 22m lines, try-lines are near their most vertical and up pitch lines (side-lines) are near their most horizontal.

Polynomials are in the format [slope, x-intercept].

From numpy.polyfit1D()

```
[
       "{
         'class': '22L',
         'polynomial': array([   -0.69315,      462.29])}”
       "{
         'class': '22R',
         'polynomial': array([    0.70358,     -1302.8])}",
       "{
         'class': 'farSideLine',
         'polynomial': array([   0.003163,        15.8])}",
       "{
         'class': 'halfway',
         'polynomial': array([     -183.7,  2.3142e+05])}",
       "{
         'class': 'nearSideLine',
         'Polynomial': array([   0.020972,      758.02])}"
]
```

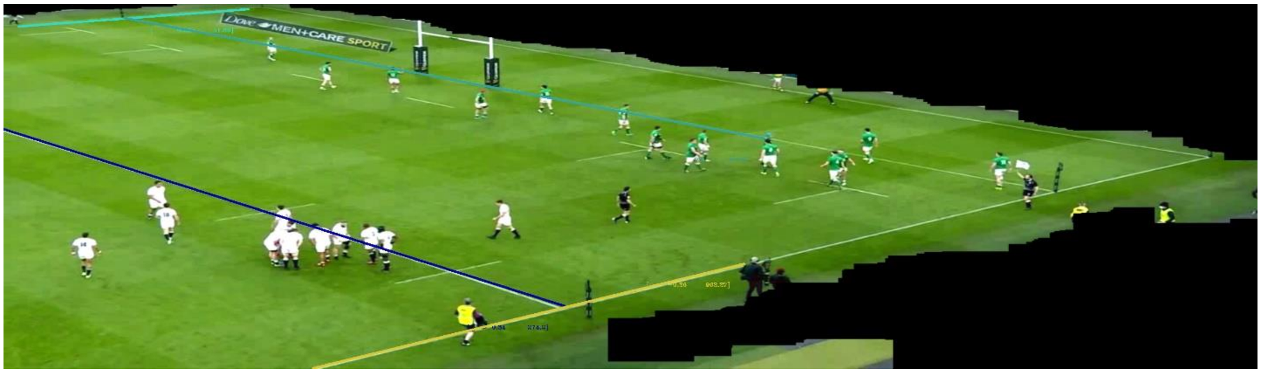

Figure 11- frame with near maximal rotation towards near right corner.

Cross-pitch lines are near their most horizontal and up pitch lines are near their most vertical


```
[
       "{
         'class': '22R',
         'polynomial': array([    0.33852,        274.2])}",
       "{
         'class': 'farSideLine',
         'polynomial': array([  -0.079881,       51.892])}",
       "{
         'class': 'nearSideLine',
         'polynomial': array([    -0.2595,       962.27])}",
       "{
         'class': 'tryLineR',
         'polynomial': array([    0.20122,     -18.435])}"
       ]
```


This is a complex process which begins with creating metrics for all detected lines. Lines were converted to polynomial form and their lengths were extracted from the *x, y* pairs returned by pyfld. Line lengths were taken because they were used for weighting during the merging process.

<u>*Initial classification round*</u>

Lines then underwent an initial classification process, followed by multiple reclassification stages based on the context of the overall state of the lines in the image. This ‘line-state’ of the image is used as context throughout the classification process to reclassify and remove lines.

Figure 12 - frame with near maximum rotation towards near left corner.

note that slopes are approximately reversed when camera is pointing towards the near left corner.


```
[
        "{
          'class': '22L',
          'polynomial': array([   -0.44266,      1128.8])}",
        "{
          'class': 'farSideLine',
          'polynomial': array([   0.088298,     -131.58])}",
        "{
          'class': 'nearSideLine',
          'polynomial': array([    0.25997,      411.84])}",
        "{
          'class': 'tryLineL',
          'polynomial': array([   -0.25499,      476.83])}"
]
```


As the camera is fixed at the halfway line and experiences minimal trucking, the line equations for the twenty-two-meter lines and the try lines stay within certain bounds depending on the image's shape, and the location of the camera. These slopes are symmetrical either side of the halfway line (Figure 10, Figure 11, Figure 12), whose slope is always higher than 3.

Cross pitch line slopes (22m and try lines) ranges were checked manually by inspecting the highest slopes when the camera is pointing directly across the pitch and at their lowest when the camera points towards the corners.

The slopes of the twenty-two-meter lines stay within range ± *(0.27, 0.73)* and try lines have slopes within ± *(0.13, 0.6)*. Twenty-two meter lines and the try lines also always have y-intercepts (where the line meets the top of the image) either less than *1.5 x the image width* for the right 22m line and greater than -*0.5 x the image width* for the left 22m line. These y-intercept bounds are used to stop cross pitch lines being mistaken for NSLs when the camera has turned far towards the near corners of the pitch.

Up-pitch lines are always flatter. Slopes of NSLs are always with absolute values less than 0.29 and found in the bottom 25% of the image and FSLs always have absolute slope values less than 0.11 and are found in the top 15% of the image.

The order in which lines are initially classified (Code Sample 1) impacts the reclassification process afterwards, as it will decide how certain image contexts will lead to misclassifications, as seen in Figure 13.

```
def classifyLine0(self, ln):
  slope = ln[2][0]
  C = ln[2][1]
  if np.abs(slope) > 3: return "halfway"
  elif self.is22L(slope,C,ln) : return "22L"
  elif self.is22R(slope,C,ln): return "22R"
  elif self.isTrylL(slope,C,ln): return "tryLineL"
  elif self.isTrylR(slope,C,ln): return "tryLineR"
  elif self.isFarSideLine(slope, C, ln) :   return "farSideLine"
  elif self.isNearSideLine(slope,C, ln) :   return "nearSideLine"
  elif np.isclose(slope, 0, atol=1.9e-1, equal_nan=False): return "UpPitchLine"
  else: return "other"
```

Code Sample 1 - Initial line classification order

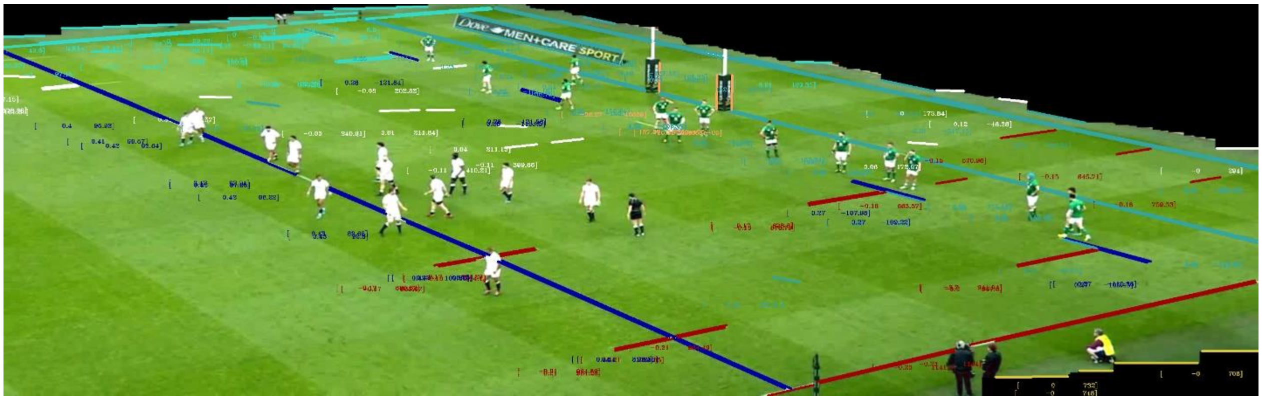

Figure 13 - There will be multiple cases of misclassification immediately following the initial round of classifications.

This image shows the near-side line (which should be yellow) being misclassified as the left try-line (red). The goalposts are also being mistaken for the halfway line (orange).

Five-meter lines are being mistaken for the right 22m line and similarly, the dead-ball lines are being mistaken for the try-lines.

For example, when the camera is pointing towards the near corners of the pitch, near side lines will be initially classified as the try-lines or 22m lines at the opposite end of the pitch.

This is because they will have similar slope and x-intercept ranges and because in Code Sample 1, the status of *isTryLineR()* and is*TryLineL()* is checked before the side-lines. Similarly, other lines such as the goalposts will be mistaken for halfway lines, so analysis of the image's lines-state is used to decide if these lines should be reclassified.

Cross-pitch line reclassifications

The first group of lines to undergo this reclassification are the side-lines which have been misclassified as try-lines or 22m lines at the pitch's corners.

Cross-pitch lines and near-side lines

Because these misclassifications occur on opposite ends of the pitch to the misclassified lines, the presence of these misclassified lines can be evaluated if the cameras perspective can be surmised.

The camera's state is treated as a pair of 'lookingLeft' and 'lookingRight' binaries, dependent on either the sum lengths of cross-pitch lines on the left-hand side of the pitch being greater than those on the right-hand side, or vice-versa. Once the perspective is obtained, then any side-lines that are misclassified as cross-pitch lines can be reclassified based on their position in the image.

The next lines to undergo this are the 22m lines. These reclassifications tend to be necessary because there is a slope range of ± (0.6, 0.27) where 22m lines overlap with try-lines, so all lines (particularly try-lines and sidelines) that are within this range will be initially classified as 22m lines.

22m line reclassifications: removing try-lines and end-lines

Prior to reclassification of the 22m lines, any of these lines that are shorter than the line groups' mean length are removed. This is because lines that are correctly classified tend to be longer than the mean of any given group.
This is due to the actual presence of a white pitch line at these areas, where most smaller lines found throughout the pitch are due to shades of grass or players silhouettes that were picked up by the Canny edge detector.

Once this is done, misclassified 22m lines are reclassified to the most appropriate try-line or side-line based on the lines' parameters. Otherwise, they will be classified as 'other'. This reclassification process is started by finding the median y-intercept for all initial 22m lines. If the camera is facing left, any 22m line with a y-intercept lower than median is reclassified and any line with a y-intercept higher than the median is reclassified when the camera is facing right. This process is repeated until the range of y-intercepts for 22m lines meets a minimum threshold of about 50 pixels, which hopefully is over-represented by legitimate 22m line samples.

Try-line reclassifications: removing end-lines

The try-lines are then reclassified because the dead-ball lines and any borders found beyond the pitch are often mistaken for try-lines. This is conducted in an identical fashion to how 22m lines are reclassified. A wider threshold for minimum y-intercept range is used here, because of the presence of the 5-meter lines.

As the reclassification process is done by removing try-line candidates found ‘beyond’ the median value, a wider minimum y-intercept range allows for the presence of both the try-lines and halfway lines. 5-meter line to try-line misclassifications tend not to be an issue because of short line removals and line-length weighting during merging.

Halfway lines and goalposts

The final cross-pitch line group to go through reclassification checks are any lines classed as ‘halfway’. All near vertical lines in the image will be classified as ‘halfway’, which means that there will be multiple misclassifications in all images, which is why this group has any lines shorter than the mean removed prior to reclassification. In images focused towards either end of the pitch, the goalposts will be routinely classed as ‘halfway’ (Figure 14); however the distribution of these line segments throughout the image is almost guaranteed to be more dispersed than the actual halfway line. Using this rule, if the standard deviation for the ‘halfway’ line groups x-coordinate is above 200 pixels, the entire group is reclassed as ‘other’.

Following this, no more cross-pitch-line reclassifications were necessary (Figure 15).

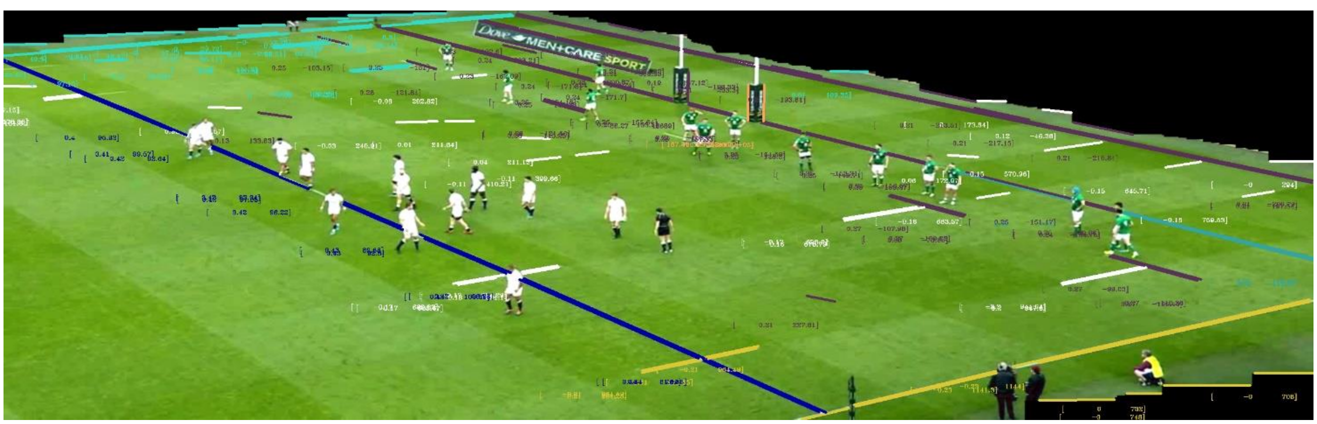

Figure 14 - The image's line state after reclassification of 22m lines, try-lines and near side-lines.

Up-pitch line reclassifications

Most reclassifications necessary for up-pitch line isolation are due to lines being created alongside the image's black spaces on the edges of the image. These black space edges tend to be horizontal due to the morphological operations which created the black spaces. Smaller up-pitch lines, such as the line-out 5m and 15m lines, may also be misclassified as side-lines. These can usually be dealt with by the aforementioned short line removal method. The influence of remaining short lines will be diminished by length weighting in the merging process.

Up-pitch lines undergo two reclassification stages combined with removals of short lines (similarly to the cross-pitch lines), and length weighting is also applied. The two reclassification stages are split into one regarding the far-side line and one for the near-side line.
The purpose of this stage is to find any lines that are wrongly classified as side-lines and reclassify them to 'other'.

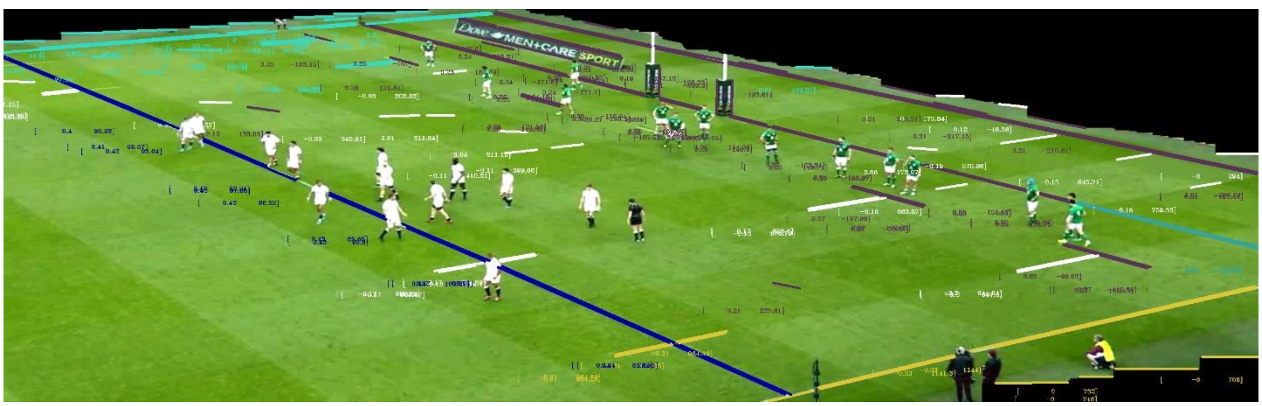

Figure 15 - Cross-pitch lines are fully reclassified.

In contrast with Figure 13*, the near-side line in the right foreground has been reclassified from 'tryLineL' (red) to 'nearSideLine' (yellow), the goalposts have been reclassified from 'halfway' to 'other' and the 22m lines and try-lines have been reduced to accurately portray their relevant respective lines*

Near-side line deletions

This reclassification step applies the rule that cross-pitch lines and up-pitch lines must always have slopes of opposite signs.
Using the images perspective parameters (‘lookingLeft’ and ‘lookingRight’), which were set during the cross-pitch to side-line reclassification stage, if the camera is deemed to be in a ‘lookingLeft’ or ‘lookingRight’ state, any side-line that has a slope value with the same sign as the image’s mean 22m line must be misclassified and removed.

Far-side line deletions

This reclassification stage also removes side-line candidates based on their slopes. A side-line approximation is created by finding appropriate points along the edge of the image’s black space (Figure 17). This is done using the image’s perspective parameters again and by iterating across the image’s black space from the top-down and in a direction leading away from the halfway line.

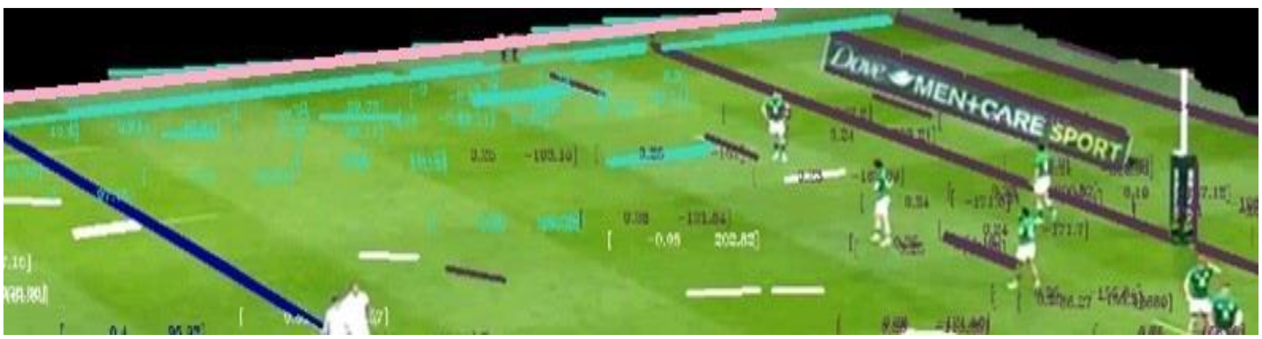

Figure 17 - The pink line illustrates a far-side line approximation, using the image's black space.

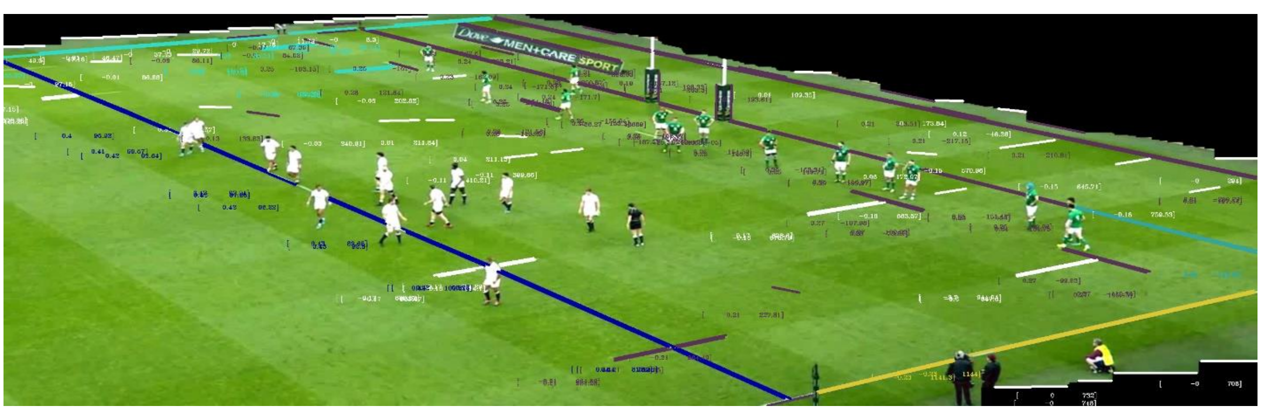

Figure 16 - following line reclassification, the line state of each group is roughly approximating at least four major pitch lines.

Once this approximation line is created, the slopes of all ‘farSideLine’ candidates are compared with it, and any whose slopes are less than half of the approximation line are removed. In addition to these steps, if the frame is focused near the halfway line, any lines found bordering black space were reclassified as ‘other’.
Following this step, the line reclassification process is complete, (Figure 16).

4) Line merging

A minimum of four points are required to perform a homographic transformation, which in this case must be four intersections of two cross-pitch lines and two up-pitch lines.

Now that all of the line groups have been reduced their best candidates, four groups are merged to form an average of each (Figure 18). The two cross-pitch line groups with the longest accumulative lengths were chosen, along with the two side-lines.

Line merging is performed by taking the mean slope and x-intercept of each line group after removing any outliers along each variable.
Prior to this, line-length weighting is performed by appending the line group with each line in proportion to their respective lengths and only if their length is beyond a threshold of 95% of the longest length in the group.

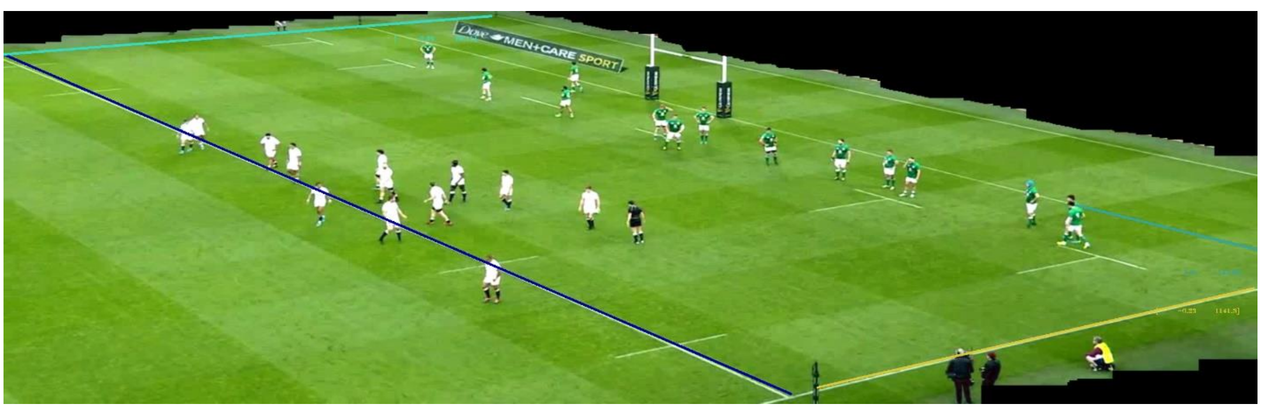

Figure 18 - Each line group has been merged.

5) *Homography Calculation*

Now that the lines have been merged and are approximating the pitches markings, a perspective transformation (homography) matrix was created from the four intersection points of the merged pitch lines (Figure 19).

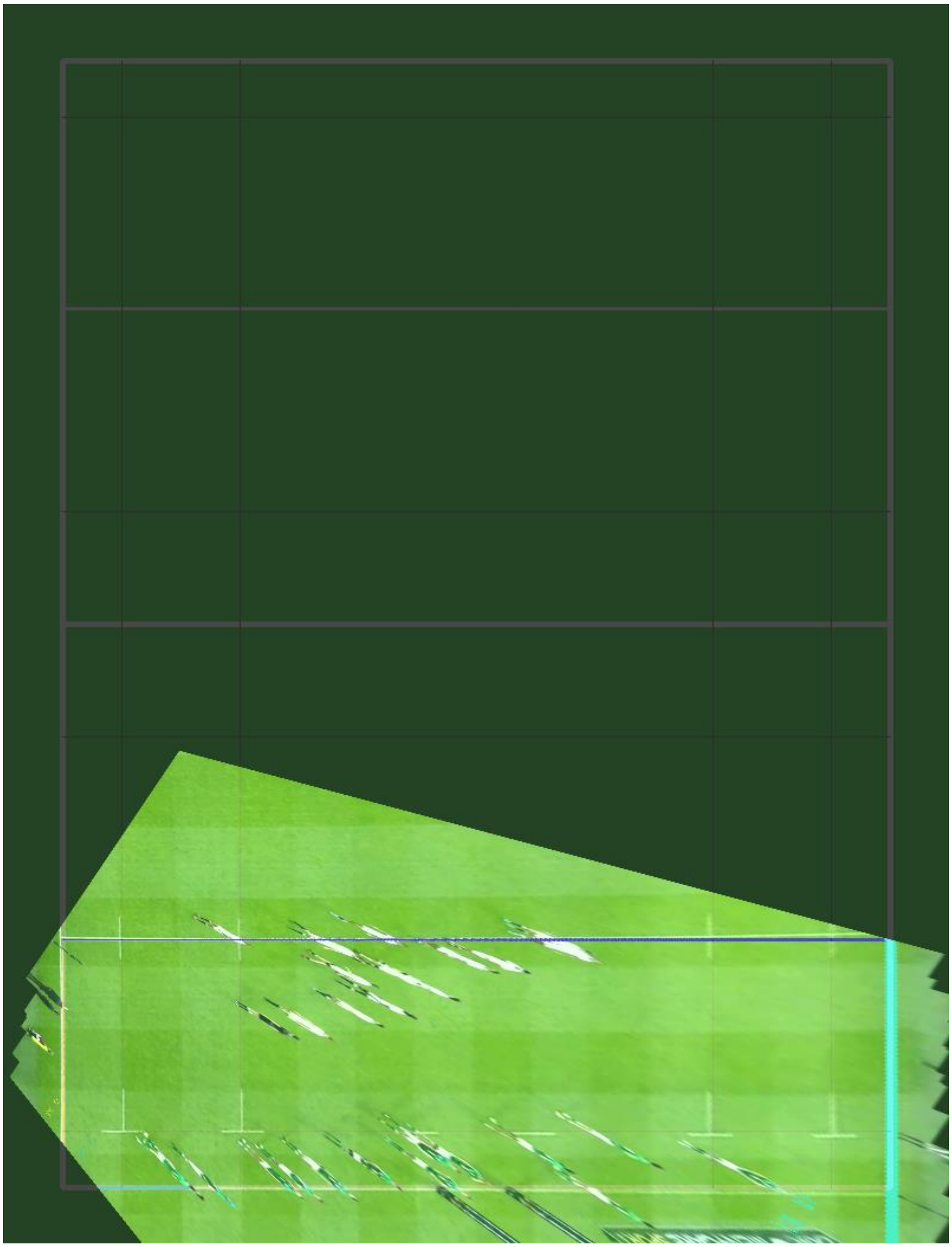

Figure 19 - The image input warped to fit a top-down rugby pitch overlay.

iii. Player Tracking

1) Image tiling

Like field registration, player tracking was performed on each input frame in a separate module (Figure 20) and then their positions were transformed to a top-down projection view. Prior to object detection, the image was tiled to improve detectability of players on the far side of the pitch.

2) Player detection

A yolo-v5 object detector, pretrained on the COCO dataset (Jocher, 2022) was used to detect players, with a minimum confidence threshold of zero.

3) Player Classification

Player classification was carried out by isolating the bounding boxes of detected players and then isolating a narrower box inside this. Then the pixels were counted in color ranges matching the jersey colors of the teams involved, the referee and players wearing yellow bibs.

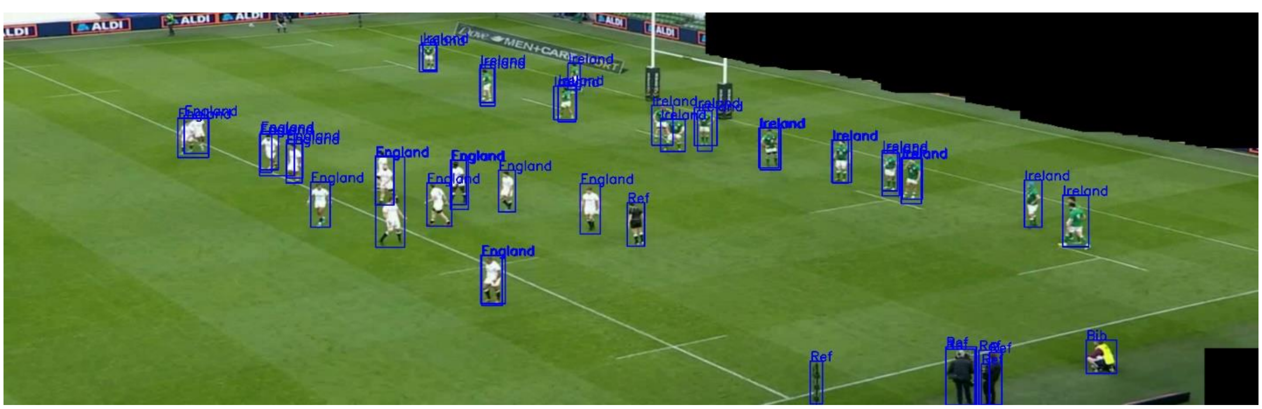

Figure 20 - Player detection module classifies players by jersey colour.

iv. Top-Down View Projection

Finally, by transforming the detected players' positions in-frame with the perspective transformation created in the field registration module, a top-down view match projection can be created (Figure 21).

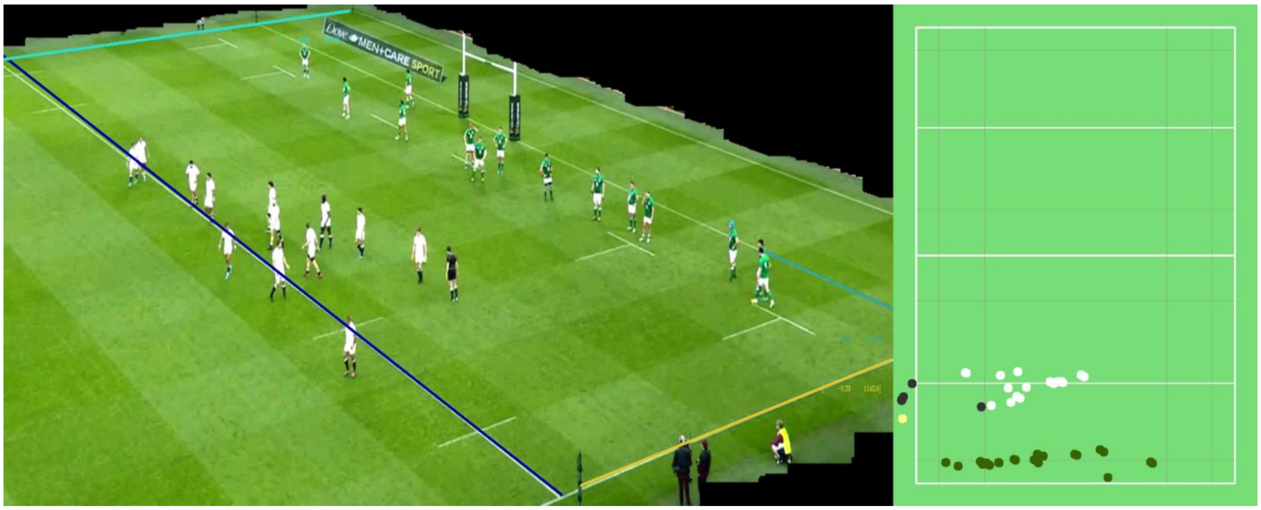

Figure 21 - Input image, alongside top-down projection of player positions.

7. Results

To test this method, multiple video sequences of Ireland vs England from the 2021 RBS Six Nations were used. 80 equally spaced frames were taken from this video and used to review the model's efficacy.

It should be noted that the absence of a labeled image-to-homography dataset meant that a standard loss analysis for the model could not be performed.

i. Field Registration

Field Registration is only possible when at least four major pitch lines are visible. After manual inspection of the 80 frames under review, 52 did not feature enough major pitch lines to create a suitable homography, leaving 28 that were suitable. These suitable frames tended to be between phases of play, prior to lineouts, scrums, kickoffs, etc.

Issues that cause poor homography estimation differ depending on the frame perspective, which can be broken into groups – Those centered at halfway, (Figure 10) and those rotated towards the ends of the pitch (Figure 11, Figure 12).

The pitch lines that cause most issues are the far-side lines. Frames focusing on the halfway line tend to have wide black borders above the far-side line, following pre-processing morphological operations. These are often misclassed as far-side lines (Figure 23), and due to the length weighting applied during line merging, these borders can cause the far-side line to be estimated above it’s true position in-frame, as seen in Figure 22.

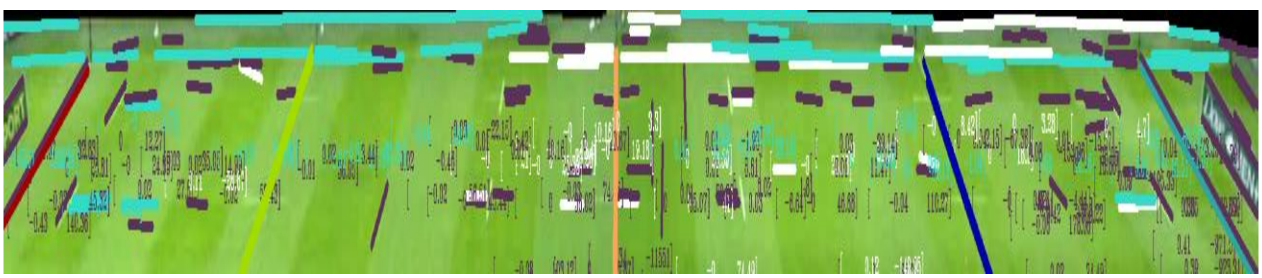

Figure 23 - line classification frame displaying multiple misclassifications along black borders at top

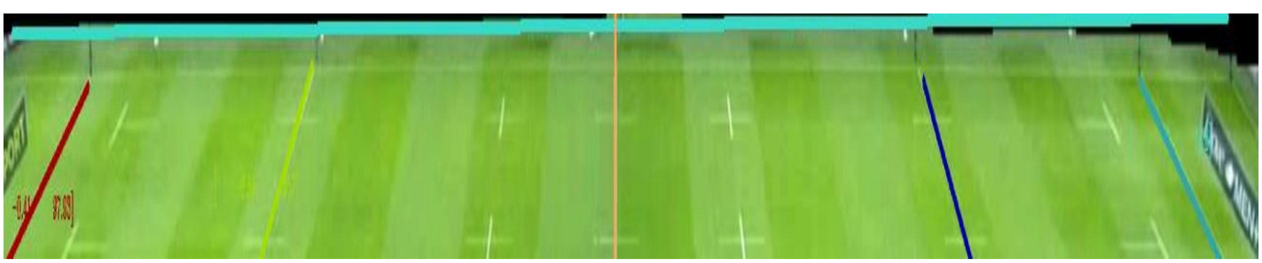

Figure 22 – The far-side line has merged along black border instead of along true sideline

15 of the 28 useable frames were centered near the halfway line. Following reclassifications of lines found along black spaces and removals of short lines in these 15 frames, the number of frames effected by the far-side line merging along the black-space border dropped from 3 to 0.

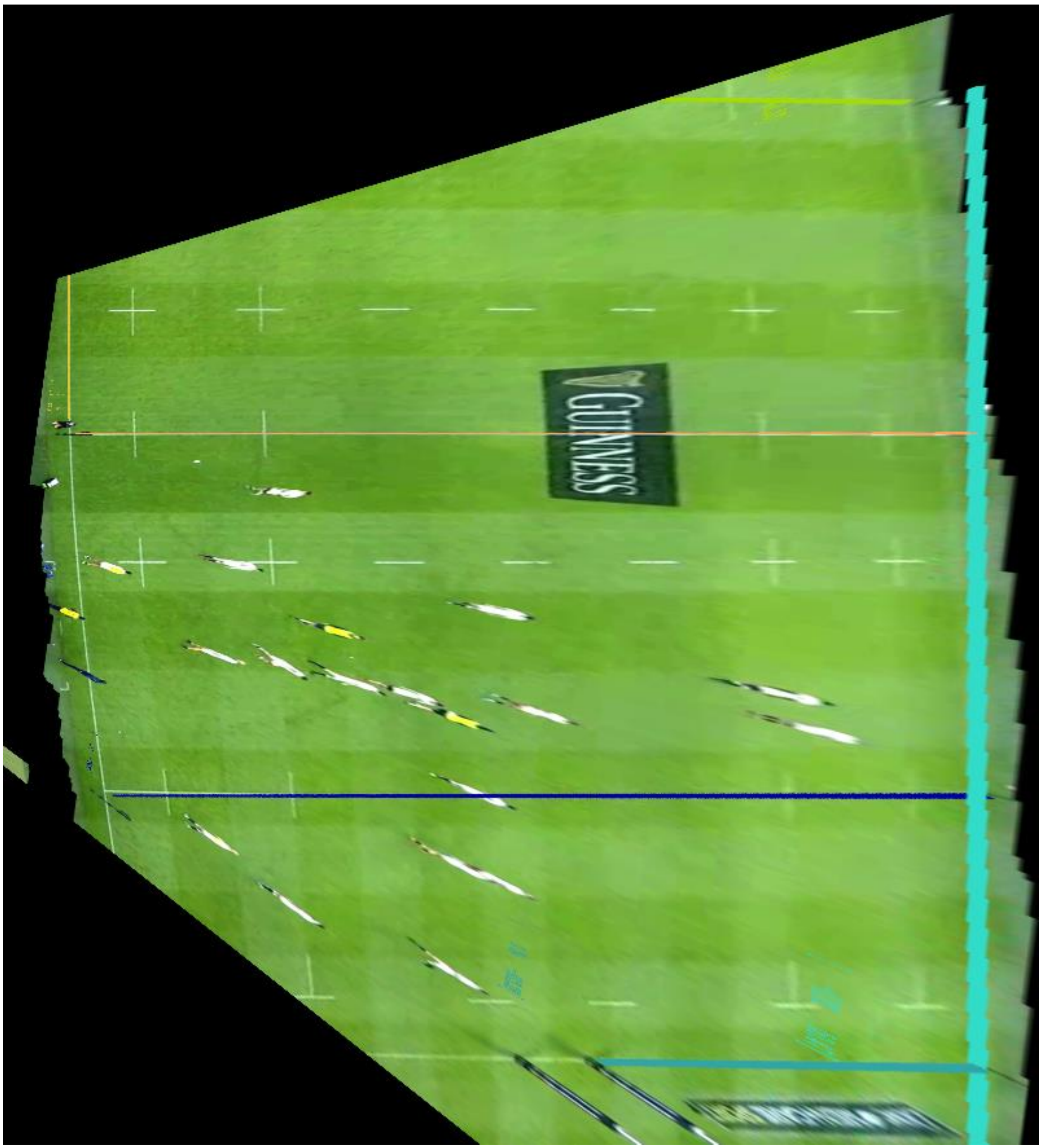

Figure 24 - large deviations of far sideline at top right of image and the near-side line at the bottom left of the image.

This frame is from the half-time period, so the zoom perspective may not reflect match footage.

Another issue effecting the near and far-side lines is due to radial lens distortion near the corners of frames (Figure 25, Figure 24, Figure 26). This is only evident when the camera is nearly maximally zoomed out. The effect that this has in these cases is that the estimated position of players near the corners of the frame will be displaced towards the sidelines. This is relatively minor, as the camera is rarely zoomed out to this degree and only small areas of the frame are affected. Typically, this is only influential in cases where the camera frame spans a large area of the pitch, such as in Figure 24, because otherwise lens distortion is affecting areas off the pitch instead.

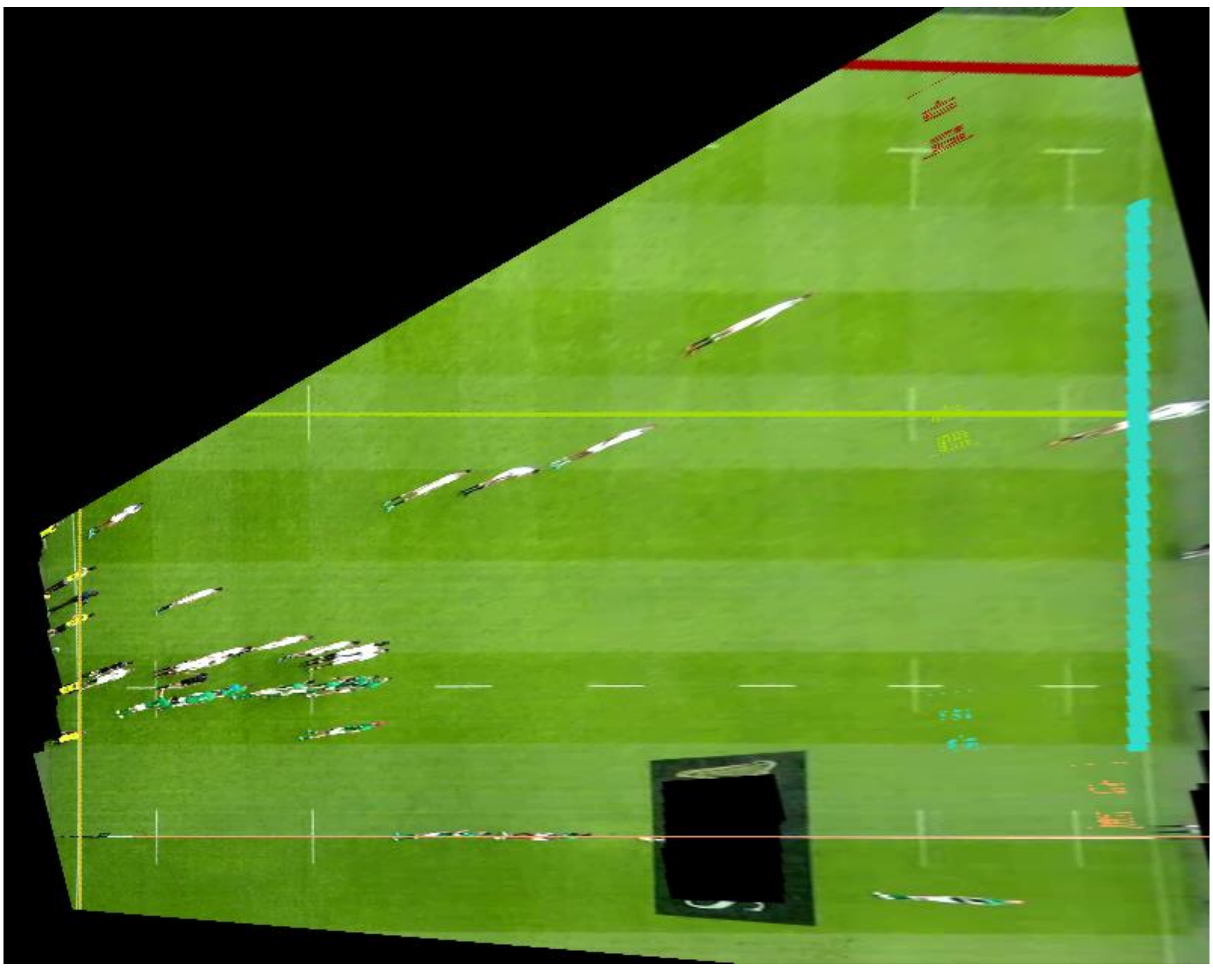

Figure 25 - radial lens distortion affecting the far-side line estimation.

Note the change in angle at the bottom right of the image.

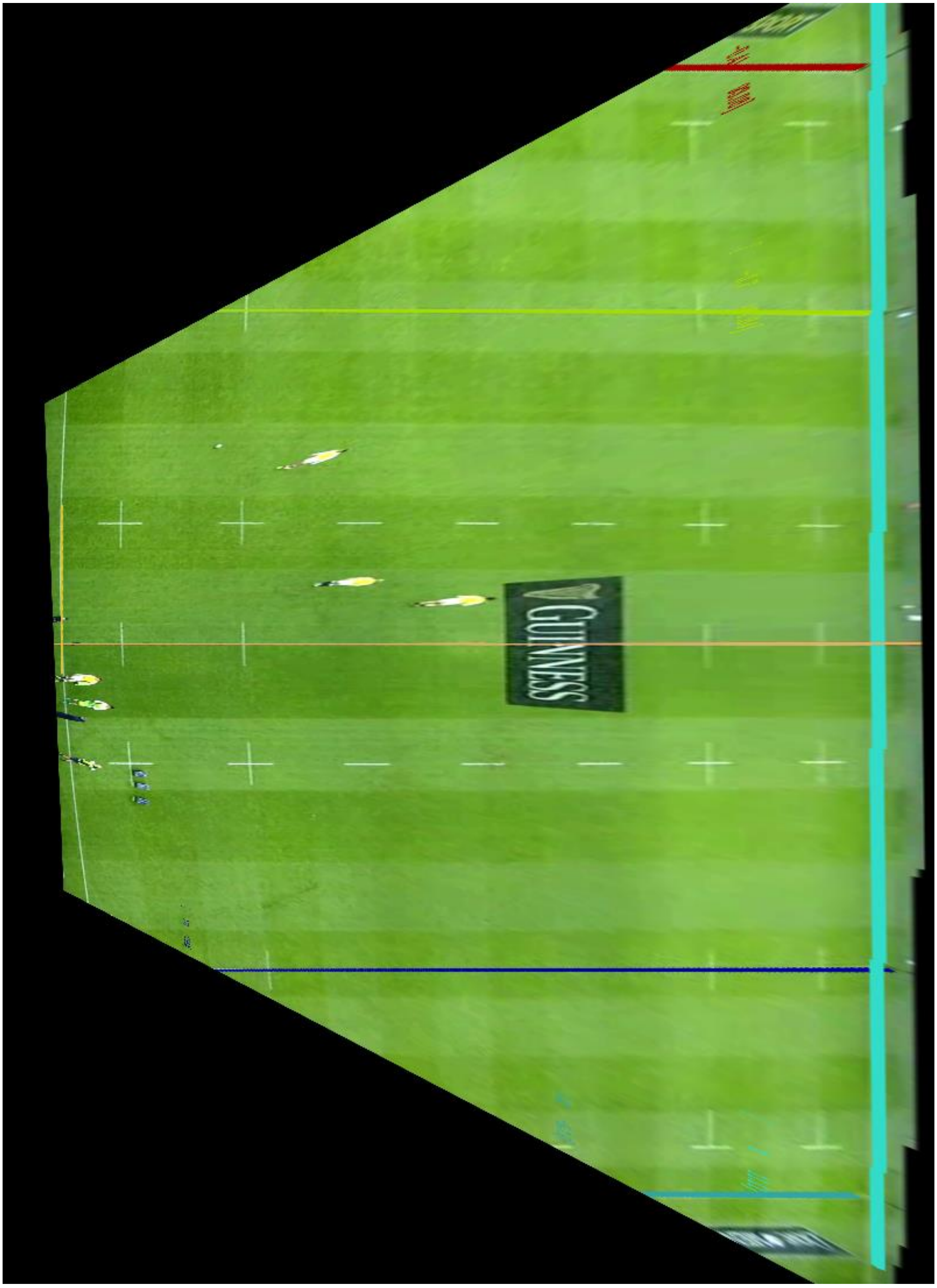

Figure 26 - radial lens distortion effecting the near-side line

In all of the 28 suitable frames inspected, all cross-pitch lines were reliably detected and merged. There are cases of minor deviations of cross pitch lines at either near or far-side lines (Figure 27). These deviations are usually less than a foot at either side-line, so their effects are negligible.

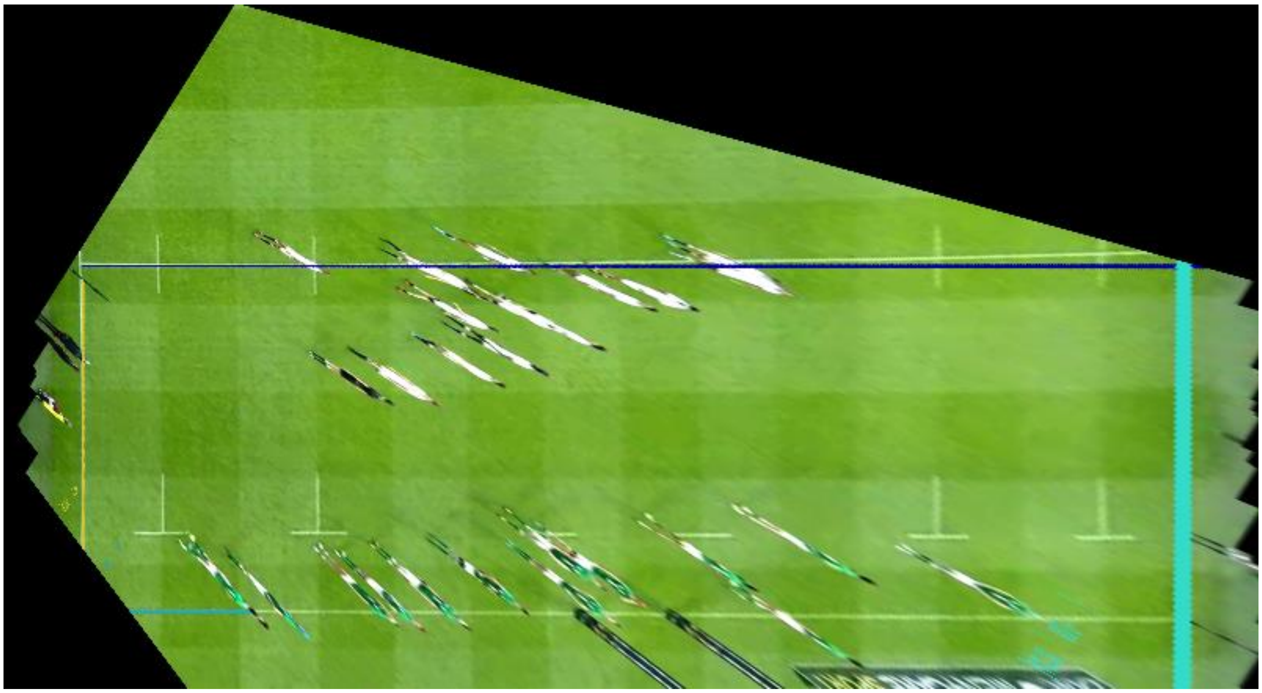

Figure 27 - minor deviation of the 22m line on far side of pitch.

Figure 28 - The neat 15m line has been classified as the near-side line, leading to an inaccurate homography estimation.

This model begins to break down when there are not four major pitch lines with which to perform a homography estimation. The 5m or 15m lines along the side-line may be misclassified as the near-side line when the actual near-side line is not visible (Figure 28). This happens often during passages of play when the camera zooms towards the center of the pitch

ii. Player Tracking

The performance of the ultralytics yolov5 object detector is not the focus of this dissertation, but it was a necessary component to evaluate the effectiveness of the field registration module.

The object detector's minimum confidence threshold was set to 0. This led to some false positive detections of the goalposts and corner flags as 'persons'.

There was a noticeable drop in the object detectors performance with regards to detecting Ireland players vs detecting England players. This was most likely due to the Irish green jersey clashing with the green grass. Increasing color contrasting markedly increased the ability of the object detector to find Ireland players.

Another issue affecting the object detectors performance was blurring as players are moving quickly in-frame. Again, this issue effected Ireland players more than England players due to the Irish green jersey colour clashing with the green grass.

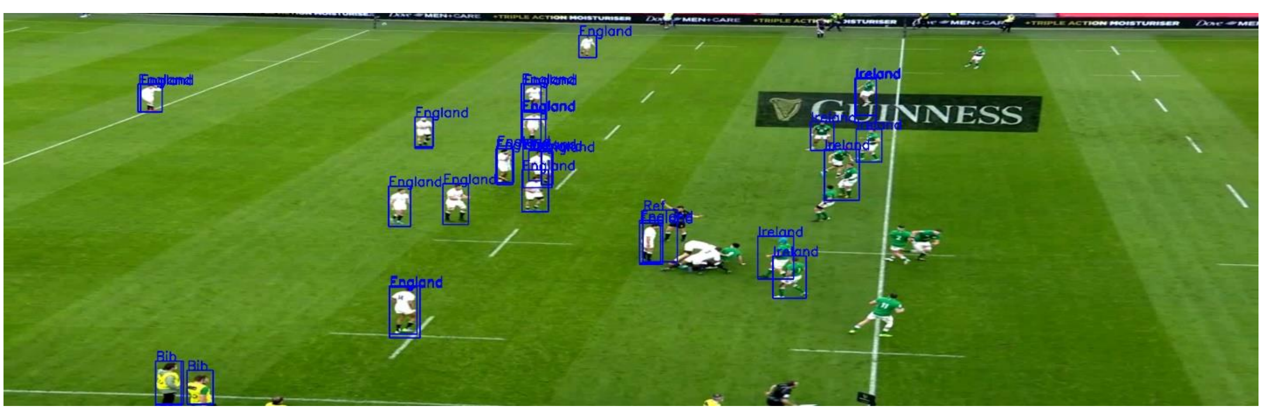

Figure 29 - The object detector is more successful at detecting England Players than Ireland players. These failures are more common when Ireland players are in motion.

Issues causing misclassification of players to their teams are entirely due to the crudeness of the method employed. Irish players have white shorts, which may sometimes be more numerous than the green of their jersey being counted during classification. For this reason, Ireland players are occasionally misclassified as being England players. Players for both teams are also occasionally mistaken as referees if there are enough dark pixels counted.

iii. Performance

In a google-colabs notebook with a GPU hardware accelerator, this model performs a homography estimation, detects and project players to a top-down projection at approximately 31s per frame. The pyfld FastLineDetector takes approximately 20 seconds to run per iteration with a 7x7 Canny aperture.

8. Conclusion

This model successfully achieves field registration of a rugby pitch with Hough line detection under limited circumstances with a high degree of reliability. Most efforts for sports field registration have been achieved with access to labeled landmark datasets and corresponding homography datasets. In the absence of these resources, a semantics-based model was created to detect and classify pitch lines based on Hough detection.

Under the assumption that a match action frame is sufficiently zoomed out and taken from a vantage point at halfway, it has been shown that following a Hough image transform, the perspective (looking left, looking right, looking towards center) of a camera filming a rugby game can be reliably estimated based on the ‘line-state’ of the image – by grouping lines based on their polynomial descriptions and accumulated length of each group. Having estimated this perspective, it has also been shown that rugby pitch lines can be accurately found by classifying and merging the lines of an images Hough transform, and then be used to achieve field homography estimation.

Rare cases of unacceptable homography estimation are typically either due to misclassifications of the far-side line or radial lens distortions affecting lines when the frame spans large a portion of the pitch.

Inaccurate identification of cross-pitch lines is rare. The halfway and 22m lines seem to be correctly identified in most cases. Due to the presence of multiple parallel lines found near the try-lines, there are occasional cases when these are estimated to run a few feet either side of the expected line value.
This model only performs Euclidean (rigid) homography transformations, not accounting for lens distortion (scikit-image, 2022). This causes the models accuracy to drop as points go farther from the points used to transform the pitch in-frame. Model accuracy across the whole pitch could potentially be improved if multiple homographies were estimated using several combinations of pitch lines.

The most useful implementation of this methodology proposed by this dissertation may be estimating player positions in game periods prior to set-pieces such as scrums, lineouts and kickoffs, as these periods are when the camera is most likely to be sufficiently zoomed out.

9. Future Work

An absence of rugby field registration datasets meant that a Hough detection model was necessary to achieve homography estimation. This model may provide a useful framework for feature extraction in future work.
If the homography estimation created by the field registration model is accurate, the location of the 5m line pitch markings and other significant landmarks should be predictable, (Figure 30).

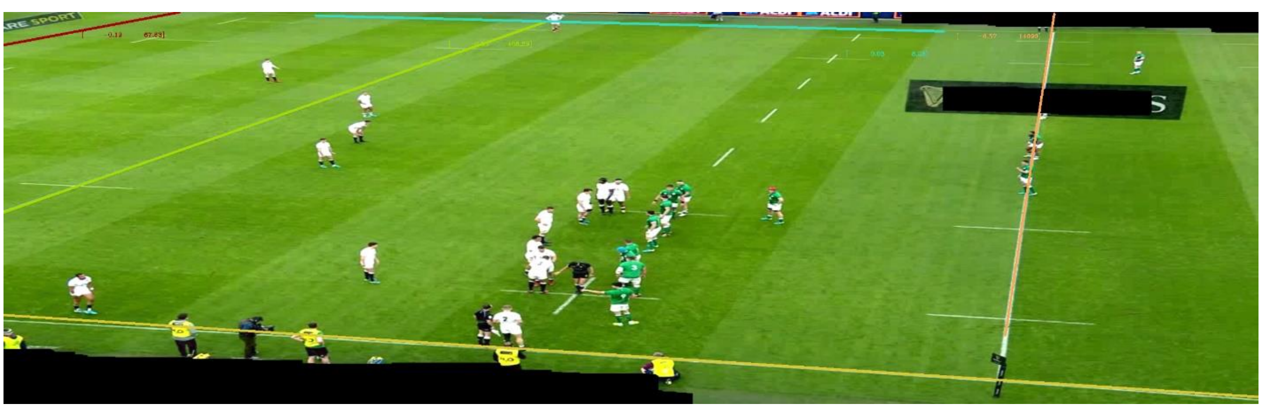

Figure 30 - The 5m lines in this frame will be extracted

By isolating these features in the top-down projection and then transforming them back to the original perspective seen in Figure 30, it may be possible to develop a pitch marking feature dataset, with images similar to those in Figure 32.

Figure 31 - The 5m lines have been isolated within a top-down pitch projection by isolating their predicted location in the top-down projection.

Figure 32 - The 5m lines have been transformed back to their original perspective.

10. References

Atul, K., 2019. *theailearner.com.* [Online]
Available at: https://theailearner.com/tag/cv2-getperspectivetransform/
[Accessed 25 July 2022].

Bialkowski, A. et al., 2013. Representing and Discovering Adversarial Team Behaviors using Player Roles. *2013 IEEE Conference on Computer Vision and Pattern Recognition,* pp. 2706-2713.

Bialkowski, A. et al., 2013. Recognising Team Activities from Noisy Data. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops,* June.pp. 984-990.

De Silva, V. et al., 2018. Player Tracking Data Analytics as a Tool for Physical Performance Management in Football: A Case Study from Chelsea Football Club Academy. Sports (Basel). *Sports,* 6(4), p. 130.

Farin, D., Krabbe, S., With, P. & Effelsberg, W., 2004. Robust camera calibration for sport videos using court models. Volume 5307, pp. 80-91.

Fern, A. & R, H., 2007. Improved Video Registration using Non-Distinctive Local Image Features. *2007 IEEE Conference on Computer Vision and Pattern Recognition,* pp. 1-8.

Garnier, P. & Gregoir, T., 2021. Evaluating Soccer Player: from Live Camera to Deep Reinforcement Learning.

Gupta, A., Little, J. J. & Woodham, R. J., 2011. Using Line and Ellipse Features for Rectification of Broadcast Hockey Video. *2011 Canadian Conference on Computer and Robot Vision,* pp. 32-39.

Hayet, J.-B., Piater, J. & Verly, J., 2004. Robust incremental rectification of sport video sequences.

Hedau, V., Hoiem, D. & Forsyth, D., 2009. Recovering the spatial layout of cluttered rooms. *2009 IEEE 12th International Conference on Computer Vision,* pp. 1849-1856.

Hennessy, L. & Jeffreys, I., 2018. The Current Use of GPS, Its Potential, and Limitations in Soccer. *Strength and Conditioning Journal,* Volume 40, p. 1.

Homayounfar, N., Fidler, S. & Urtasun, R., 2017. Sports Field Localization via Deep Structured Models. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* pp. 4012-4020.

Ishii, T. et al., 2002. THE DISTANCE COVERED OF SOCCER AND RUGBY REFEREES DURING THE MATCH USING A MOBILE "GPS". *Scientific Proceedings of the X X Ih International Symposium on Biomechanics in Sports, Spain,* pp. 322-325.

Jernej, F., David, S. & Matthias, K., 2021. *Friends of Tracking - The future of possession value models.* [Online]
Available at: https://youtu.be/GEp0TJoramM?t=3477
[Accessed 25 July 2022].

Jiang, W. et al., 2020. Optimizing Through Learned Errors for Accurate Sports Field Registration. March.pp. 201-210.

Jocher, G., 2022. *yolov5 documentation.* [Online]
Available at: https://github.com/ultralytics/yolov5
[Accessed 05 08 2022].

Kim, H. & Hong, K.-S., 2001. Robust Image Mosaicing of Soccer Videos using Self-Calibration and Line Tracking. *Pattern Anal. Appl..,* pp. 9-19.

Liang, Q. et al., 2020. Multi-Player Tracking for Multi-View Sports Videos with Improved K-Shortest Path Algorithm. *Applied Sciences,* 10(3).

Pers, J. & Kovačič, S., 2000. Computer Vision System for Tracking Players in Sports Games. pp. 177 - 182.

Ruiheng, Z. et al., 2020. Multi-camera multi-player tracking with deep player identification in sports video. *Pattern Recognition,* 102(107260), p. 107260.

scikit-image, 2022. *scikit-image.org - Types of homographies.* [Online]
Available at: https://scikit-image.org/docs/stable/auto_examples/transform/plot_transform_types.html
[Accessed 09 08 2022].

Sharma, R. A., Bhat, B., Gandhi, V. & Jawahar, C. V., 2018. Automated Top View Registration of Broadcast Football Videos. *2018 IEEE Winter Conference on Applications of Computer Vision (WACV),* pp. 305-313.

Tsukada, T., 2021. *pyfld.* [Online]
Available at: https://pypi.org/project/pyfld/
[Accessed 01 08 2022].

World Rugby, 2022. *conditioning-for-rugby/advanced-conditioning-for-rugby-pre-level-2/programme-management-and-player-monitoring/monitoring-workload/gps-metrics.* [Online]
Available at: https://passport.world.rugby/conditioning-for-rugby/advanced-conditioning-for-rugby-pre-level-2/programme-management-and-player-monitoring/monitoring-workload/gps-metrics/