# Automating the Annotation of American Football

BRANDON WINGARD, Clemson University, USA

A series of algorithms were developed in order to automatically extract information from a football play using computer vision techniques. Extraction of play information required two tasks: field reconstruction and player recognition. The field markers of yard-lines and hash-marks were extracted using Hough transforms. The player tracking algorithms used blob tracking techniques that focused on the unique color features of the teams that were involved in the play. Despite challenges in reconstructing field coordinates as the play evolved, a team formation was interpreted from a sample still image, proving that these algorithms can indeed work in tandem to extract relevant football information from NFL film, showing high promise for future capability of this technology.

Additional Key Words and Phrases: datasets, NFL, Hough Filter, OpenCV, neural networks, Contour Mapping

**ACM Reference Format:**
Brandon Wingard. 2018. Automating the Annotation of American Football. *Proc. ACM Meas. Anal. Comput. Syst.* 37, 4, Article 111 (August 2018), 4 pages. https://doi.org/XXXXXXX.XXXXXXX

## 1 INTRODUCTION

As the increase in interest in leveraging software innovation into the NFL process, there are handful of necessary processes in the American football analytical community that must still done by hand. One such process is the diagramming and annotation of each play in an entire game.

The current investigation utilizes the computer vision techniques to extract player information from video of an American football play. This investigation is not meant to enhance the NFL Viewer experience; it is instead meant to augment and automate the process of annotating a heavily time consuming process of diagramming play designs for NFL staff. By leveraging this technology, it would enable more efficient use of resources during the player scouting and film review for NFL teams.

The capability of determining meaningful data from film of a football play leads to the formation of datasets for entire games that can then be quickly analyzed. Such an advancement would allow for Scouting and Personnel staff to expedite decision making processes using a standardized annotation process.

## 2 PREVIOUS WORKS REVIEW

### 2.1 Review of Previous Work

The approach that was being utilized in the previous investigation was using the color gradient of the player's uniform with the turf to identify player's locations. Identification of the the yard-lines and the hash-marks is also done by locating the endzone (denoted by a contrasting color along the edge of the field of play) and then using a Sobel Edge Filter and Hough Transform to locate the remainder of the yard-line markers.

It then utilized an color-space inversion to the image and Hough Transformation to determine the location of the hash-marks.

Once done, it utilized a dominant color to track given the identifier color in a player's uniform to track the player's location.

### 2.2 Contributions of the Current Work

As noted in the introduction, the purpose of this investigation is to expand on the previous works by investigating other possible methodologies (including a Kalman Filter) of identifying players and further expanding on previous topics of interest including player motion tracking and annotation of a given player's movement through the course of a play.

## 3 METHODOLOGY

### 3.1 Methodology Summary

In order to properly annotate a football play from a video, it will need to be decomposed into 4 distinct tasks. As discussed in the Reference Literature, the first two tasks will carry over.

- 1. ***Field Reconstruction***. This task is intended to capture the stationary objects on the football field that describe the orientation of the field. These objects include the vertical yard-lines and horizontal hash-marks of the football field. These objects can be used to estimate the field coordinates of any object on the field from that object's pixel value in the video.
- 2. ***Player Recognition***. This task is designed to identify the 11 players on each team.
- 3. ***Movement Tracking***. This task is designed to track moving objects in the football play.
- 4. ***Annotating Movement***. This task is designed to then take the path of the specific player and annotate a path that follows their movement.

### 3.2 Detailed Presentation of Methodology

*3.2.1 Dataset Description.* The choice of NFL film for this investigation is paramount to its success. To ensure that all 22 players are visible, this investigation will utilize 'All-22" film that is made available to scouts, analysts, players, and coaches, to specifically study game footage.

The data set that we will use is from Superbowl LI between the New England Patriots and the Atlanta Falcons. This video was acquired at the behest of a coaching staff member who has requested to remain anonymous. Specifically, the first data set that will be used is an Atlanta Passing Play. By nature of Atlanta's Passing system, the players will be spread out in the field of play. Such play will be less difficult to track the players as it develops, unlike a running

Author's address: Brandon Wingard, bmwinga@clemson.edu, Clemson University, 201 Sikes, Clemson, South Carolina, USA, 29634.


2476-1249/2018/8-ART111 $15.00
https://doi.org/XXXXXXX.XXXXXXX

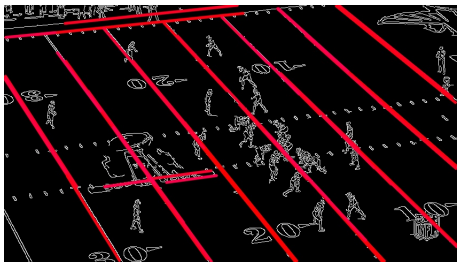

Fig. 1. Results after applying the Hough Filter to the initial frame

Fig. 2. Result from the Original Methodology used in previous Experiments when attempting to isolate New England Players

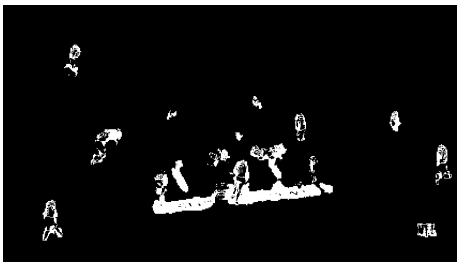

Fig. 3. Result from the Original Methodology used in previous Experiments when attempting to isolate Atlanta Players

play, where most of the players would be grouped closely together and more difficult to distinguish individuals. The second data set is a New England Rushing Play. This particular play is used to demonstrate the ability to track players when they are much closer together in the field of play.

An important component that will need to be taken into account is the change in size of the players as the camera zooms out to continue to have all 22 players in frame. This can be adjusted as the player themselves should be tracked from frame to frame for annotation purposes later. With this, it becomes important to mask out any players on the sideline as they will cause issues with player tracking. This will be done by finding the white boundary of the Sideline and masking anything outside the boundary.

*3.2.2 Field Reconstruction Methodology.* As with previous literature, it is important that a basis is established before approaching field reconstruction. Firstly we must define the coordinate system and the corresponding image anchors. For the sake of this experiment, we will be using a 2D coordinate plane as a 3rd dimension offers little valuable insight for our intentions. Following tradition with image processing, the Origin of our coordinate system will coincide with the origin point of the photograph. As such, X will increase with respect to rightward movement column-wise while Y values increase with downward movement row-wise. This will allow us to then establish a corresponding Coordinate system starting with X at the back edge of the Left Endzone and increase until the Back edge of the Right Endzone is reached. With Y Starting along the "Home Team" (the side in which the camera is located) side of the field of play and increasing as it approaches the "Visiting Team" sideline opposite our point of reference.

Now that we have established our basis, we must discuss specifics about the field of play and regulating dimensions that must be accounted for.

- End zones are 10 yards wide
- The entire width of the field, including both End Zones, is 120 yards
- The field is horizontally separated by 10 yard intervales by vertical yardline markers
- The entire height of the field is 159 feet or 53 yards
- The lower hashmarks are located 70.75 feet or 23.5 yards from the lower boundary line
- The upper hashmarks are located 70.75 feet or 23.5 yards from the upper boundary line
- Therefore, the distance between hashmarks is 18.5 feet or 6.15 yards apart
- The lower numerical markers are placed 13 yards away from the lower boundary line
- The upper numerical markers are placed 13 yards away from the upper boundary line
- Therefore, the distance between the numerical markers is 27 yards
- The crossbar of each goal post is 10 feet above the field located directly over the back edge of the endzone
- The width of the cross bar of each goalpost is identical to the distance between the hashmarks (18.5 feet or 6.15 yards)

The vertical and horizontal hashmarks are critical components to field reconstruction as they provide a reference point for the relative location of players and events that happen in the field of play. Vertical yardlines give an identifier along our X axis while horizontal hashmarks give reference for where the football is with respect to our Y axis on the football field. For reconstruction, a Hough Transformation will allow us to identify the location of these yardline markers.

As shown in (figure), the Hough Transformation Filter accurately identifies the yardline markers.

Once the pixel coordinates of the yardlines and hashmarks are determined, it is possible to then utilize a transformation that will transform pixel coordinates into field coordinates.This is then utilized in the annotation of player movement later in the experiment.

*3.2.3 Original Player Tracking Methodology.* Following the inspiration for this paper, the original tracking methodology will follow color blotting based on the jerseys of the respective teams. Using

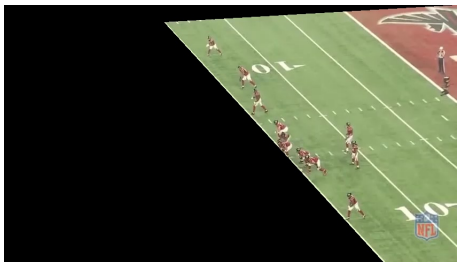

Fig. 4. The Area of Interest for Atlanta players that was segmented by the Line of Scrimmage

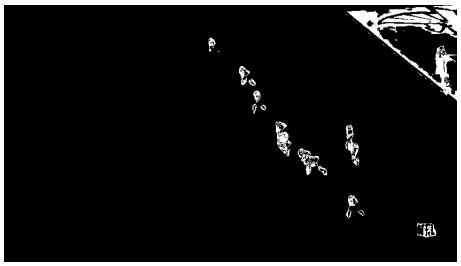

Fig. 5. After Segmentation, a HSV Filter is applied to isolate the Scarlet Color of Atlanta's Jerseys, and then any Value that was not within the specified threshold is decimated to 0

this, it will be important to use a high contrast filter to isolate the individual players from the background. Specifically for this example, both teams have high contrast team colors that allow them to stand out from the field, and contrast from each other. As such, a color detection filter will be used to filter out the players.

Atlanta's primary jersey colors are Scarlet Jerseys and White pants. Following along with the original paper, as white is a dominant color in the video, Scarlet (RGB: (167, 25, 48)) will be used to track these players. However, it will be noted that scarlet is a color whose perception could be influenced by the relative luminance of the video. As such when looking for Atlanta's players, it will be important to signify the exact luminance value for scarlet.

Likewise, the New England Patriots are wearing White Jerseys with Navy colored pants. For the same reason as choosing scarlet for the Falcons, the New England Patriots will have their Navy (RGB: (0, 34, 68)) pants tracked this game. Based on how dark of a color Navy is, its perception should not be affected by the relative luminance of the video.

Once a player has been located, they will then be assigned a player id and have their relative position tracked. Extending further, skill position players on passing plays will have their starting point, the turning point in their route called a route stem, and their ending point for annotation.

Given the sideline angle of the camera it is not possible to track the numbers on the jersey of the players otherwise, the numbers themselves would be used to identify and label each player.

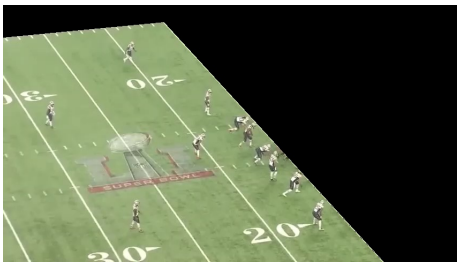

Fig. 6. The Area of Interest for New England players that was segmented by the Line of Scrimmage

Fig. 7. After Segmentation, a HSV Filter is applied to isolate the Navy Color of New England's Jerseys, and then any Value that was not within the specified threshold is decimated to 0

*3.2.4 Updated Player Tracking Methodology.* After further experimentation, a modified methodology was developed in order to significantly improve static player detection. This modified methodology uses a combination of masking based on the Line of Scrimmage, Color Masking, Threshold Obliteration, and Contour Mapping. By splitting the Pre-snap frame in half along the Line of Scrimmage, it isolates half of the players for each color detection sequence. This, in turn, then generates a map of colors associated with each team that then is mapped into a Threshold Obliteration Map to isolate out the colored areas from the background. Finally, a Contour Mapping filter is then applied to create "Areas of Interest" that correspond to predicted player positions. By isolating each side of the field it allowed to wider HSV regions to be tested resulting in improved player recognition. However, when it came to tracking players post-snap, isolating the field became useless while the remainder of the process was still viable. To remedy this, Color Based Masking, Threshold Obliteration, and Contour Mapping were utilized for the remainder of the frames within the video.

## 4 EXPERIMENTAL RESULTS

To evaluate the success of our Tracking Methodologies, there needed to be established parameters that we can evaluate from. As such the following criteria were used to evaluate model performance:

- ***Skill Position Recognition***: The ability to recognize all Receivers, Running Backs, the Quarterback, Corners, Safeties, and the Middle Linebacker. This resulted in a total of 11 players.
- ***Total Player Count Recognition***.:The ability to recognize all 22 players on the field.

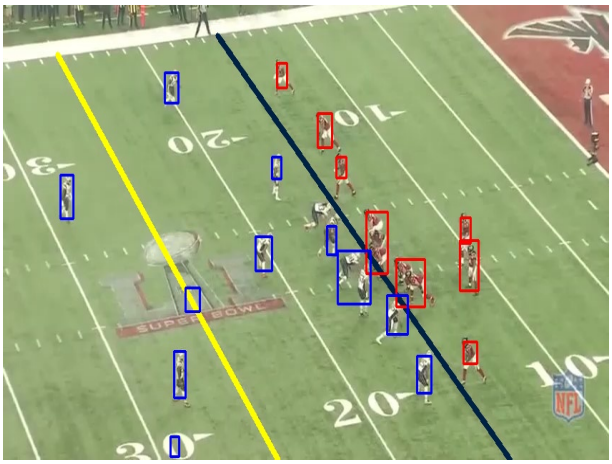

Fig. 8. The End Result of both iterations of Player Detection. Line of Scrimmage is Marked in Navy, the 1st Down marker is Yellow

- ***False Positive Rate***: The number of falsely identified objects within the Frame that were not players.

***Original Methodology Results*** :

| Metric | Raw Score | Percentage |
|---|---|---|
| Skill Position Recognition | 5/11 | 0.455 |
| Total Player Recognition | 5/22 | 0.227 |
| False Positive Rate | 2/7 | 0.286 |

Table 1. Results from layering Color Masking and Contour Mapping

***Modified Methodology Results*** :

| Metric | Raw Score | Percentage |
|---|---|---|
| Skill Position Recognition | 11/11 | 1.00 |
| Total Player Recognition | 17/22 | 0.773 |
| False Positive Rate | 3/20 | 0.15 |

Table 2. Results from layering Image Segmentation, Color Masking, Threshold Obliteration, and Contour Mapping

### 4.1 Analyzing Results

The transformation between pixel and field coordinates was done successfully by hand (though error was introduced) for the sake of annotating the play. Unfortunately, the original goal of a completely automated play annotation was not realized despite best efforts. However, it is possible to show that such software is possible with a specially tuned DNN model. Such results reconfirm that there is significant room for further investigation into this area once a model is ready.

Discussing the change in methodology comes at a change from the previously used methodology to use of a personally curated model. This change had an overwhelmingly positive affect during pre-snap player recognition as it improved from 45.4% to 63.6% in Skill Position Detection. There was also a substantial increase in Total Player Recognition from 22.7% to 77.3% leaving only offensive and defensive lineman not properly identified.

## 5 CONCLUSION

In Conclusion, it was important to continue the research into player tracking methodologies to expand the possibilities for all vested parties of the NFL. The combination of pre-snap player detection using Line of Scrimmage image separation, Color HSV Detection, and Threshold Obliteration drastically improved the overall results of the pre-snap detection while also reducing the false positive rate.While the finalized result of play annotation was not achieved, the experiment did reinforce methodologies of pre-snap player recognition which in itself can be invaluable to teams for labeling and categorizing play types by their pre-snap formation.

There is sufficient evidence to suggest that with higher resolution base data, that the post-snap player recognition would improve and could ultimately lead to realized automated play annotation.

## 6 AREAS FOR FURTHER INVESTIGATION

Despite the challenges with the publicly available data, there is significant promise for future developments in this space. Future work would be focused on automating the pixel to coordinate system conversion. From there, it would be possible to then extrapolate the exact locations of each player during the play. Further research could be done into the segmentation of each offensive and defensive lineman to finalize the correct identification of all 22 players on the field at a given time.

## 7 ONLINE RESOURCES

## REFERENCES


[1] lee, title=Automating NFL Film Study: Using Computer Vision to Analyze All-22 Film, url=https://cvgl.stanford.edu/teaching/cs231a$_w inter1415/, journal = CS231A - ComputerVision : From3Dreconstructiontorecognition, publisher = StanfordUniversity, author = Lee, TimothyE$

[2] techvidvan$_2$021, $title = Detectobjectsofsimilarcolorusingopencvinpython, url = https://techvidvan.com/tutorials/detect - objects - of - similar - color - using - op$

[3] hynes convention center$_2$017, $title = UsingComputerVisionandMachineLearningtoAutoma https://global - uploads.webflow.com/5f1af76ed86d6771ad48324b/6005a1646d2198b4$

[4]

[5] misccheng$_f engl u_m ahasseni_f iez_f ern_t odorovic, title = PlayTypeRecognitioninReal - WorldFootballVideo, url = https://web.engr.oregonstate.edu/ sinisa/research/publications/wacv14.pdf, publish OregonStateUniversity, author = Cheng, ShengandFeng, Zhongyuanand Lu, Qingkaiand$

[6] team color codes$_2$021, $title = AtlantaFalconsHexColorCodes, url = https://teamcolorcodes.com/atlanta - falcons - color - codes/, journal = TeamColorCodes, year = 2021, month = Jun$

[8] dayala$_d ayala_2$020, $title = Colorfiltering/segmentation/detection˜HSV, url = https://cvexplained.wordpress.com/2020/04/28/color - detection - hsv/, journal = ComputerVision, author = Dayala, RaghunathandDayala, Raghunath, year = 2020, month = Aug$