Formula 1 Analytics

John David Anthony[1], Eric Yuyitung[2], Connor Sparling[3], Brendon Cop[4]
QMIND – Queen's AI Hub
Queen's University, Kingston, Ontario K7L 3N6, Canada.

1 e-mail: johndavid.anthony1@gmail.com
2 e-mail: 17edy@queensu.ca
3 e-mail: connor.sparling@queensu.ca
4 e-mail: 17bsc@queensu.ca

Abstract: Broadcasting companies are continually looking for ways to increase viewer engagement. Can a broadcast be tailored to the wants of viewers based on observing and comparing viewer engagement levels? To answer this question, we created 3 classification models to tag footage of the 2019 Formula 1 racing season. The race was tagged to determine if the footage was a replay, what camera angle is currently shown, and what teams are on screen. We were able to tag replays with 100% accuracy; however, unbalanced data presented an issue for the camera angle and team on screen classifier. We were able to classify teams with a sufficient amount of data correctly 76% of the time.

---

1. INTRODUCTION

1.1 Motivation

In the age of video streaming services such as Netflix, viewers can choose to watch content at a time that is suitable for them. Sporting events, however, are mainly viewed live and because of their long duration, maintaining the viewer's attention is a challenge.

The value of sports broadcasting rights continues to increase year after year. As a result, broadcasters are looking for ways to increase both viewership and engagement from casual and hardcore fans. By increasing fan engagement, broadcasters can attract new viewers and gain the opportunity to display advertisements to further strengthen their bottom line.

The question that now remains is how can sports companies alter their broadcasts to make it more exciting for the viewer? In order to make a sports broadcast more engaging, the broadcaster needs to know what the viewer wants to see.

1.2 Related Works

To increase the engagement of the viewer, there needs to be a way to measure their engagement. Instead of eye trackers, J. Hernandez et al. [1] used RGB cameras to predict the viewer's engagement level. However, challenges arise when the viewer's engagement could not be fully determined from the image captured from the camera. The viewer may be looking away from the sports broadcast and discussing the sport with a friend. In this scenario the system would predict this as a low level of engagement.

Pattern matching is a technique used to identify if a template is present in an image. The work of Y. Hel-Or et al. [2] explores how to accomplish this task while maintaining efficiency and effectiveness.

1.3 Problem Definition

We have decided to analyze footage from the 2019 Formula 1 season to determine what impacts viewer engagement. Deloitte Canada has provided their EmotionPlus platform that predicts facial sentiment from a standard camera. In addition, Deloitte has provided raw data of over 40 viewers. The provided data included the level of Valence for the viewers which is the viewer's general happiness while watching the race. The data also includes the viewer's level of engagement. This is calculated through eye and face movement as well as their expressed emotions.

To determine what effects the viewer's engagement level, we first need to classify the frames of the race. The first property to be tracked is if a frame is a replay or not. Secondly, what camera angle is currently being used. Finally, what teams are on screen.

Three classifiers need to be created that will take a frame from a race as input and will output their predicted label.

These labels will then be used against the emotion data previously collected to identify patterns and insights.

2. METHODOLOGY

To begin with, the data needed to be manually tagged to train the classifiers. Three different races were filmed at 50 frames per second. 1 frame was chosen every 50 frames to be manually tagged, as we did not want to create the world's largest repository of Formula 1 pictures. The individual frames were then tagged with its replay status, camera angle, and the teams on screen in bounding boxes.

Figure 1: Replay Template

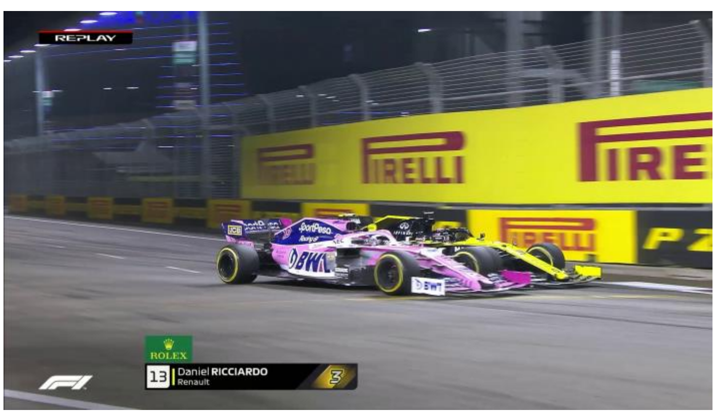

Figure 2: Replay Frame

In Formula 1 broadcasts, there is a replay flag present in every replay sequence. This is incredibly helpful as it reduces the magnitude of the first classifier. By using the pattern matching method explained previously, the template (Fig. 1) is transferred across the frame (Fig. 2) seeking the best match. If a sufficient match is found, the frame is tagged as a replay.

The camera angle classifier uses a pre-trained neural network Inception-V3 (IV3). IV3 is a convolutional neural network used for analyzing images. The last 3 layers of the network were retrained with the tagged data from the races. The camera angles were divided into 5 categories: driver view, pit view, spectator view, track view, and other view. The majority of frames fell within the track view.

The team on screen classifier uses a simplified variant of YOLOV3 – a pre-trained object detection network. Originally, YOLOV3 was planned to generate cropped images of all cars in the frame. The images would then be fed through a function to generate their colour histogram. This histogram would then be used to predict the car's team based on their colour. This was not needed as YOLOV3 displayed exceptional promise in differentiating cars based on their colour and decals.

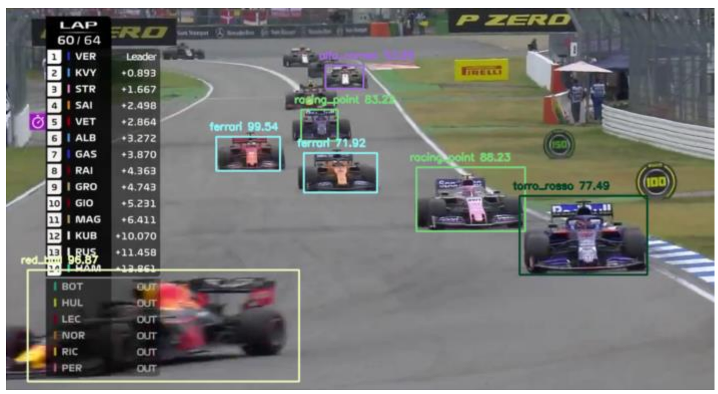

Figure 3: Tagged Teams

3. RESULTS AND DISCUSSION

The replay classifier was able to achieve 99.99% accuracy as a result of the trivial nature of the problem. The classifier's only misclassifications arose when there was a crowd in the area where the replay flag usually is. This caused sporadic replay classifications for only 1 frame. As there will never be a 1 frame replay, this classification can be ignored bringing the total accuracy to 100%.

The camera angle classifier was not able to get accuracy as high as the replay classifier however it was a more difficult problem. In the current state, the classifier is biased towards predicting a track view.

This is the result of unbalanced data, as the majority of the race is viewed on the track.
The team on screen classifier had similar issues to the camera angle classifier. The training and testing sets were unbalanced. When only including the 4 teams with the most data (Ferrari, Racing Point, Renault, and Mercedes) the average precision was 76%.

Additionally, COVID-19 had a negative impact on the ability for the team to continue development as well as we hoped.

4. CONCLUSIONS AND FUTURE WORK

The work completed thus far shows that broadcast footage can be tagged based on various features.

Moving forward, to improve the impact of the system we would like to expand the replay tagging system. Throughout the race, various events occur on the screen such as the fastest lap, warning flags, the team on the radio, etc. These events have an image on the screen similar to the replay flag, thus we can detect these events occurring and further analyze their impact on the viewer’s engagement.

Secondly, further work needs to be done on the camera angle classifier and team on screen classifier. We will use multiple strategies to tackle the issue of unbalanced data. By oversampling the minority classes, we allow the network to learn more about the imbalanced classes. Additionally, cost sensitive learning is an effective way to train the network where the cost of classification mistakes is higher on the minority classes.

Finally, we believe that more emotional data would allow for more in-depth insight into viewer engagement. Especially emotional insight for geographically separated viewers. This would allow for additional awareness into what engages different regions and create the opportunity for more tailored broadcasts.

REFERENCES

[1] J. Hernandez, Zicheng Liu, G. Hulten, D. DeBarr, K. Krum and Z. Zhang, "Measuring the engagement level of TV viewers," 2013 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG), Shanghai, 2013, pp. 1-7.

[2] Y. Hel-Or and H. Hel-Or, "Real-time pattern matching using projection kernels," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, no. 9, pp. 1430-1445, Sept. 2005.