What Are You Looking At? Team Fight Prediction Through Player Camera

Marko Tot[1]*, Michelangelo Conserva[1]*, Alan Pedrassoli Chitayat[2]†, Athanasios Kokkinakis[2]‡, Sagarika Patra[2]†, Simon Demediuk[2]‡, Alvaro Caceres Munoz[2]‡, Oluseji Olarewaju[2]‡, Marian Ursu[2]‡, Ben Kirmann[2]‡, Jonathan Hook[2]‡, Florian Block[2]‡, Anders Drachen[2]‡, Diego Perez-Liebana[1]*

[1]*Queen Mary University London, London, United Kingdom*

[2]*Weavr, Arena Research Cluster (ARC), University of York, York, United Kingdom*

*{*m.tot, m.conserva, diego.perez*}@qmul.ac.uk

†{*alan.pchitayat, moni.patra*}@york.ac.uk

‡{*firstname.surename*}@york.ac.uk

***Abstract*—Esport is a large and still growing industry with vast audiences. Multiplayer Online Battle Arenas (MOBAs), a sub-genre of esports, possess a very complex environment, which often leads to experts missing important coverage while broadcasting live competitions. One common game event that holds significant importance for broadcasting is referred to as a team fight engagement. Professional player's own knowledge and understanding of the game may provide a solution to this problem. This paper suggests a model that predicts and detects ongoing team fights in a live scenario. This approach outlines a novel technique of deriving representations of a complex game environment by relying on player knowledge. This is done by analysing the positions of the in-game characters and their associated cameras, utilising this data to train a neural network. The proposed model is able to both assist in the production of live esport coverage as well as provide a live, expert-derived, analysis of the game without the need of relying on outside sources.**

***Index Terms*—neural network, team fight, engagement, MOBA, esports, player analytics**

I. Introduction

The esport industry has been rapidly growing and generating vast interest, both through large audiences [1] as well as economically [2]. For this reason, academic interest has also developed to meet the demands of this growing industry [3]. One popular branch of esports is the field of Multiplayer Online Battle Arena (MOBA), including titles such as Dota 2 [4] and League of Legends [5]. Esport tournaments for these titles are commonly organised with community-funded prize pools (either partly or entirely) that reach over 34 million dollars - as of the Dota 2 International 2019 [6]. Due to this active and engaged community, broadcast entities are often looking for different ways to enhance the audience experience to maintain and increase their viewership numbers [7].

Tournament organisers play a fundamental role in directing audiences focus during the match through their streaming and similar delivery platforms. If engagements, such as team fights are not timely detected, important parts of the game can be missed by their audience. Due to the overwhelming accessibility of esport, and indeed regular sports, through several streaming and broadcasting options [8], broadcasters are driven to provide as complete and comprehensive coverage of important game events as possible. However, esports are characterised by their fast pace which can present a challenge even for experts [9] in the field. This can be observed during quick events, such as team fights that typically start abruptly. Thus, a key challenge identified by researchers and the industry is to losslessly transmit the highly complex flow of information from the game to the audience.

This paper provides a novel methodology in the identification and prediction of game events that are prevalent in the broadcast of MOBAs. Specifically, ways to identify and predict team fights, which are in-game events where both teams engage in a confrontation using their abilities and resources in an attempt to gain the advantage in the match while penalising the opposing team. As identified in the literature, these events are established as some of the most enthralling moments for audiences [7], as they are typically very active and can often change the course of the game [10]. However, despite their importance, team fights can often be entirely or partly missed by audiences or tournament broadcasters due to the complexity of the game and the fast-paced tendencies of engagements.

By utilising player input, the work presented in this paper attempts to predict team fight engagements. The suggested model is designed to utilise players' knowledge and understanding of the game to address this problem and predict team fights shortly prior to their start. This is achieved by studying the camera position of players at any given time to train a neural network to identify patterns that can be used to predict and detect ongoing team fights.

Camera locations have been selected as they represent the information that a player can visually derive from the environment at any given time. As this is closely related to player vision [11], meaningful information about the player's

The first three authors of this paper have contributed equally, and thus it should be referenced as Tot, Conserva, Chitayat, et al.

978-1-6654-3886-5/21/$31.00 ©2021 IEEE

intent could be inferred. Player vision is particularly relevant for understanding player's informed decision making. Patterns in player behaviour could be observed and used to make game events predictions. By studying the effects of player vision and how it may connect to team fights, this paper proposes a model that attempts to understand this complex environment by relying on the player's understanding of the game.

This paper shows that, despite the complexity of MOBA games, it is possible to successfully predict player engagements such as team fights, by only considering the player character and camera position. The model is designed to interpret player intent which, as observed in the literature, can be driven by vision and game-state data. The hypothesis proposed by this paper is that when a team is planning to start a fight, their cameras should converge together before their characters move into the position[1].

This paper focuses on Dota 2 as a domain in order to study this hypothesis. The game description, with an emphasis on core game mechanics and relevant game-specific terminology is provided in Section II. Section III reviews the literature about event prediction in the esport domain. A description of the employed methodology, including data acquisition and the training process, is presented in Section IV. Section V displays the outline of the performance, for test and train data sets, as well as the ecosystem data, representing a real use case scenario. The analysis of the obtained results, with a comparison to their respective in-game events, is shown in Section VI. Section VII revisits the proposed hypothesis and evaluates it in regards to the observed results. Finally, Section VIII considers potential avenues of research and proposes several paths of extending this methodology for future projects.

II. Dota 2

Dota 2 is a top-down perspective game with a diagonally symmetric map. In this game, two teams (Radiant and Dire) of five players each attempt to attack and destroy the opponent's base. Players choose from a wide number of characters (heroes), each with their own unique set of abilities and skills, allowing for different roles to be taken.

The two bases are connected through three lanes (Top, Middle and Bottom), each containing buildings (towers) that attack the opponents, dealing a large amount of damage when in close proximity. The map, showing the lanes and both bases, can be seen in Figure 1.

In order to reach the enemy's base, teams must destroy each of the towers in order, which requires a large amount of in-game resources and teamwork to achieve. For this reason, teams often engage in confrontations with each other, which are referred to as engagements. Large engagements are further categorised as team fights, which provide opportunities for inflicting a severe detriment to your opponent while providing the winning team with a large number of in-game resources. The game is won once the main building in the opponent's base is destroyed. Due to the large commitment required to achieve this, teams often engage in large scale team fights to penalise the opponent and progress towards that goal.

[1]In MOBAs players can move their cameras independently of their characters

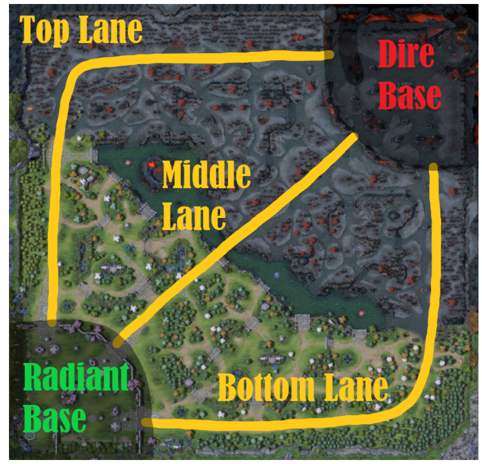

Fig. 1. Dota 2 map with marked lanes and bases.

III. Related Work

Through the esport literature, many attempts to identify and predict different events and aspects within MOBAs have been explored. A common focus in the literature is on the game outcome prediction, where authors model the game environment to discern the game state in order to predict the winner of the match. This is done throughout several stages of the game. Some authors have suggested models that predict the game prior to match start [12], while others have used game state data to predict the outcome during live matches [13]. The authors outline the potential of utilising machine learning techniques, such as logistic regression and random forests, to achieve varying degrees of accuracy. Their results also highlight the difference in performance that can be achieved depending on the time period of the prediction, where shorter intervals typically achieve higher accuracy and reliability compared to the ones looking further into the future.

Furthermore, predicting game events is not limited to their outcome. The use of game-state data is often employed in making predictions. Some authors have utilised this technique [10] to determine the danger level encountered by a player at any given time, performing a death prediction. This was achieved by employing a deep neural network with a vast amount of labelled data to train a model. The resulting network is able to make short term predictions on how likely an in-game character is to die within the next five seconds, displaying the capability of making data-derived predictions that are reliable.

Other authors have attempted to use data to identify aspects of the game which are not easily ascertained, such as player roles [14]. By utilising clustering techniques, and historic performance from professional players, the authors were able to identify the role that an individual takes in the team. This allows for a more in-depth understanding of the game state. This study also showcases the potential of utilising player decision making to detect and classify performance. This is particularly reliable for professional esport tournament data, where player performance is of a high degree of proficiency due to the competitive nature of those tournaments [9].

Similarly, some authors have used game state data to define events such as engagements, also known as encounters [15]. By analysing the capability of each character, the authors were able to determine a minimum distance between characters that allows for an encounter to happen. By observing the utilisation of the character's active abilities as well as the transfer of damage or healing, the authors formally defined an encounter. Team fights - which is the focus of this paper - is a type of encounter that was later labelled by the game developers through the OpenDota platform [16] within Dota 2 games. A formal description of what the game developers define as a team fight is not available. However, it can be inferred that it involves an encounter in which more than two characters die within a set amount of time. Those labels can be retrieved through the OpenDota platform although, as described in Section VI, those labels have inconsistent standards for start and end time.

Lastly, some authors have identified the importance of players acquiring information [11] - which is referred to as *player vision*. By studying the amount of vision available to the team, due to the imperfect information aspects of the game, the authors were able to observe a direct link between vision and in-game advantage. This highlights the importance of information available to a player and how it impacts their strategy and decision making.

As noted in the literature, modelling the game state is difficult, but it can be achieved through several means of interpreting and acquiring information about the game. For this reason, this paper proposes the utilisation of player knowledge to direct a model to simplify the complex game state and make predictions. Player camera positions can be used to infer player vision, which has been shown to be connected to decision making [11].

IV. Methodology

A. Data

In this study, a total of *1,457* professional Dota 2 matches were gathered from the game patch *7.27* using the OpenDota API. This data set consisted of team fight labels as well as match replay files. Using the Clarity Analyzer library [17] - a free Java library for reading Dota 2 replay files - camera and in-game character position data were extracted from those files at one-second intervals. This data was then encoded into four heatmaps, one for each of the following:

- Radiant players position.
- Radiant cameras position.
- Dire players position.
- Dire cameras position.

The heatmaps were generated by aggregating the data from five consecutive snapshots, matching five seconds of game time. A single snapshot contains information about each player's camera position and their in-game character position. The five-second interval was chosen for the purpose of capturing enough information about camera position while reducing the impact of the short term flicking of the camera, moving to a random location due to various events, which players occasionally perform in the game. Combining a series of snapshots provides an indication of where the heroes in a single team were, as well as their associated player vision - i.e. what they were looking at.

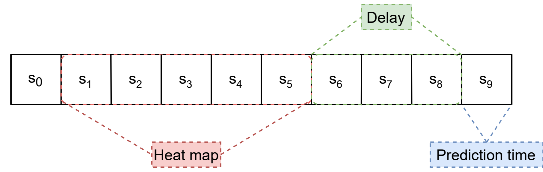

Fig. 2. Heat map generation. S_i denotes information gathered in a single snapshot.

The generated heatmaps split the Dota 2 map into cells. The original range of the coordinates spanned between $[-8472, 9198]$ on the x-axis and $[-8579, 8845]$ on the y-axis. In order to balance between maintaining the precision of the input data, with keeping the input to the neural network as small as possible the dimensions of the heatmap was set to 50×50. Each cell on the heatmap representing 300×350 pixels of the Dota 2 map. The 4 heatmaps were encoded in one of the 4 image channels. The upper limit value of 255 corresponds to the maximum amount of convergence of data (i.e. all players or camera positions converged to the same pixel area in the world map for the entire 5-second interval), while a lower limit value of 0 representing an empty area.

This information makes the input for the neural network, and it is used to predict if a team fight is going to happen. In order to ensure that the model was predicting future team fights instead of only detecting currently ongoing ones, an additional 3 seconds were added to the prediction time labels. The selected delay would give sufficient time to the broadcaster to focus their attention on the incoming event. Those 3 seconds were not used as a part of the heatmap. To reflect what is expected of a prediction using live data, if the snapshots were taken for the in-game time of 1-5 seconds, then the prediction is registered at the start of second 9 without using the snapshots for seconds 6-8. The information stored in each heatmap, delay and prediction time can be seen in Figure 2.

For each of the 1457 professional matches, team fight start and end times are collected timestamps data through the OpenDota platform. The data was processed into a form matching the previously acquired snapshots. For each second of the match, a label was created. That label represented whether there was a team fight at that particular time or not, with the addition of the three seconds delay. This suggests that the training labels values were set to true for any snapshot that contained a team fight. It is important to note that those labels contained inconsistencies with the formal engagement definition as defined by the literature, as described in Section VI.

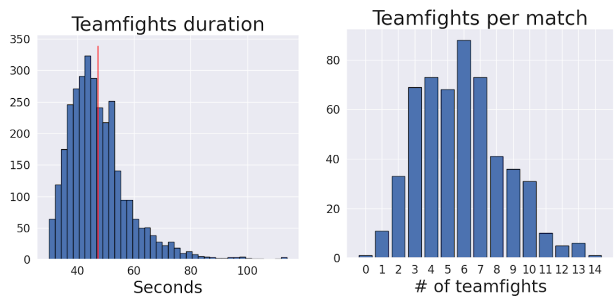

Fig. 3. Team fight duration and the number of team fights per game.

In a typical Dota 2 game, on average, team fights account for only 12% of the total game length. This causes a notable disparity in the amount of team and non-team fight labels. Figure 3 displays the aggregated data for the labels retrieved from OpenDota. In order to address the much lower frequency of team fights, the data set was artificially balanced. This was done by selecting all of the labels for each individual team fight. An equal amount of non-team fight data was also included per team fight, half of this was retrieved prior to the team fight start, and the other half post the conclusion of the fight. As Figure 3 outlines, on average there are six team fights per game with each of them lasting about 48 seconds in duration. After applying this balancing filter, in total there were approximately $800,000$ data points of balanced data.

Furthermore, symmetries were explored to augment the training data. The training data was inverted to generate a greater variance in the heatmaps. Due to the Dota 2 map being mostly symmetrical, inverting the X, Y, or both axes of the heatmap provided additional variety in the training data. This step was done to reduce over-fitting.

B. Training procedure

The data was split $80\% - 20\%$ for training and testing purposes. Each data point contained all four heatmaps, and a label, marking if there was a team fight three seconds into the future. Because of the nature of the heatmaps used as inputs, a two-part network was employed to allow for feature extraction. The first part consists of a convolutional network, and the second part is represented by a deep sequential network. Different architectures were trained and their performance compared. These variations included:

- Different non-linearities.
- Adjusting the parameters for the convolutional layers.
- Batch normalization.
- Adaptive max pooling.
- Different amount of linear layers.
- Multiple dropout rates.

Multiple changes have been enacted. The sequential module of the network was initially expanded from 1 to 4 layers, containing [512, 256, 128, 1] neurons per layer, and subsequently increased to 8 containing [2048, 1024, 768, 512, 256, 128, 64, 1] neurons in each layer. The dropout rate was also included to treat the over-fitting issue. Dropout rates of 0.1, 0.2, 0.4 and 0.5 after each layer have been tested. Leaky ReLU non-linearity has also been introduced instead of the standard ReLU with in-place enabled. Different values for Batch normalization ranging from 4 to 64 have been tested, and adaptive max pooling introduced although it was not adopted for the final model. The summary of the different parameters used for the final model is presented in Table I.

TABLE I
EXPLORED NETWORK PARAMETERS.

Non-Linearity	ReLU, ReLU + inplace, Leaky ReLU
Number of filters	4, 32, 64
Batch normalization	No normalization, 4, 32, 64
Max pooling	[2, 2], [5, 5], Adaptive Max Pool
Number of linear layers	1, 4, 8
Dropout rate	0.0, 0.1, 0.2, 0.4, 0.5

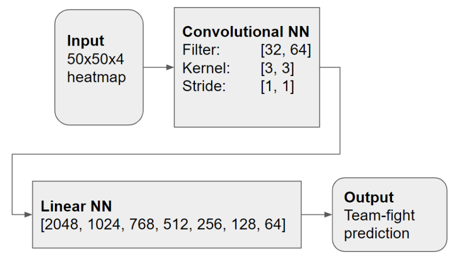

Fig. 4. Final network architecture.

All of the architectures showed similar performance, with the biggest difference manifesting with the increase in the number of linear layers. The final architecture employed two convolutional layers, connected by Leaky ReLU activation functions [18]. Parameters for the number of filters in the convolutional layers was set to 32 and 64 respectively, kernel size to 3, and stride to 1. The network architecture can be seen in Figure 4. Batch normalization was done after each layer. Max pooling with kernel size 2, stride value of 2, was added after both layers as well. Lastly, the output of the convolutional module was flattened and fed into a series of fully connected linear layers.

The sequential part of the network consists of 8 linear layers, including the final output layer, with Leaky ReLU as the selected activation function, all connected by dropout layers in between. The final layer was the classification layer with one output, representing the network prediction. During the training procedure, the network achieved the best performance with a heavy dropout rate of 0.5. The full network architecture, including the sizes for each of the layers, can be seen in Figure 4. The model was trained with the Adam optimizer [19]. The initial learning rate was set to $1e^{-5}$ with the weight decay of $1e^{-4}$. Binary cross-entropy loss was selected for the training process.

V. RESULTS

In this section, the train and validation accuracy is used to report on the model's performance. The result of the training process was a model that reached the accuracy of 84% on

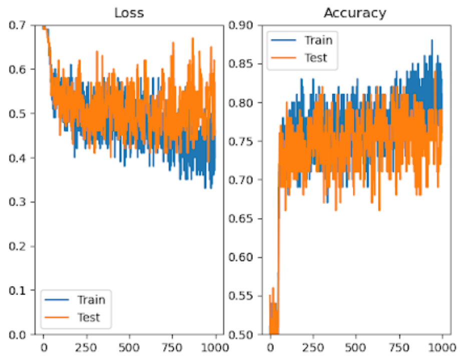

Fig. 5. Training loss and accuracy

TABLE II
FALSE-POSITIVE PREDICTIONS AND THE CORRESPONDING IN GAME EVENTS.

Game Time	Description
6:20	Two engagements at the same time at last for 15 seconds. 6 characters present in the engagement in the top lane, 4 characters on the bottom one. Two heroes die, one in each engagement.
20:35	A small engagement that starts a few seconds before (20:28) turns into a bigger one after several characters join the fight. 5 heroes die.
36:18	All 10 characters present. Relatively quick engagement on the Radiant side of the map, two Dire characters die.
46:29	Last engagement in the game. Starts with a chase, and ends with the Dire team winning the game. The duration of the fight was 58 seconds, all of the characters on the Radiant side die.

both train and validation set. Figure 5 displays the entire training performance, including training and validation for both accuracy and loss. However, this accuracy was achieved on an artificially balanced and augmented data set. For this reason, additional testing was needed.

In order to test the model further, a previously unseen match is used to analyse the performance in an actual use case environment. This method best simulates the performance during its use in the ecosystem, reproducing the behaviour of the model as in a real Dota 2 game.

Figure 6 reports an example of the results obtained for a concrete match *ID:5492227432*[2]. In this Figure, the ‘Time to next team fight’ graph displays the amount of time until the next labelled team fight. The following graphs indicate the desired (in blue) and achieved (in orange) behaviours of the prediction model. The Blue line represents obtained labels, while the orange represents the model’s outputs, with the corresponding confidence thresholds. If the output has reached or surpassed the confidence threshold, the orange line is set to high. Otherwise, it is set to low. A total match between the orange and the blue line would indicate a perfect prediction model. For demonstration purposes, four different levels of the prediction confidence threshold of the network classifier are used, ranging from 0.5 to 0.9.

Additionally, Figure 7 summarises the classification performance of the model using confusion matrices for the same confidence thresholds as displayed in Figure 6. Using the data from the confusion matrices, f1 scores were calculated. Base f1 score, for confidence threshold of 0.50, was 0.49, at 0.63 it increased to 0.51. The maximum f1 score of 0.55 was reached at 0.77, and it dropped to 0.50 at the threshold of 0.90.

VI. DISCUSSION

In this study, a model, and more importantly, a new methodology for predicting and detecting ongoing team fights are proposed. Using only in-game character and camera positions the proposed model was able to reach similar levels of accuracy as encountered in the literature for other similar event predictions [10], [11]. However, it is evident that the high amount of false-positives has impacted the performance of the neural network, as they are the main reason for the f1 scores not reaching similar values.

A parameter search done on the confidence threshold of the model offers a way of reducing the number of false-positives. Modifying the threshold amount from 0.5 to a different value, showed improvement in the performance of the model. Despite visible improvement, the obtained f1 score remained lower than observed accuracy. This indicated a possible issue with the labels. As a formal definition of the labelled team fights is not available, manual evaluation was required to determine the cause for the low f1 scores.

Upon closer examination of the events happening in a match, it was discovered that the large majority of false-positive predictions (91.4%) correspond to an engagement in the game as defined in the literature [15]. The analysis of the positive prediction was done for the match presented in Figure 6 and a sample of the analysis and their corresponding game events is presented in Table II. All of the analysis was done with the confidence threshold of 0.77, with the exact criteria for an engagement taken from the existing literature.

The full description of all positive predictions for the entire match and their corresponding game events can be seen in Appendix A. Most of the cases had all of the preconditions needed to be classified as an engagement. The only differentiating factor that can be inferred from observing the game events was that the number of characters that died during the engagement did not reach the estimated threshold of three and thus the event did not get labelled as a team fight. Additionally, an edge case could be observed where fights that meet the presumed team fight criteria were not labelled as such in the OpenDota data due to it happening just before the game ends. Furthermore, it can be noted that in many instances the OpenDota labels were set to true either too early (i.e. several seconds prior to the engagement start) or too late (i.e. several seconds into the engagement, even when heroes have already been killed). This inconsistency with the labels could be a major factor that negatively impacted the performance of the model. In many of these cases, the proposed model was able

[2] A replay of the match can be obtained the through OpenDota website at: https://www.opendota.com/matches/5492227432

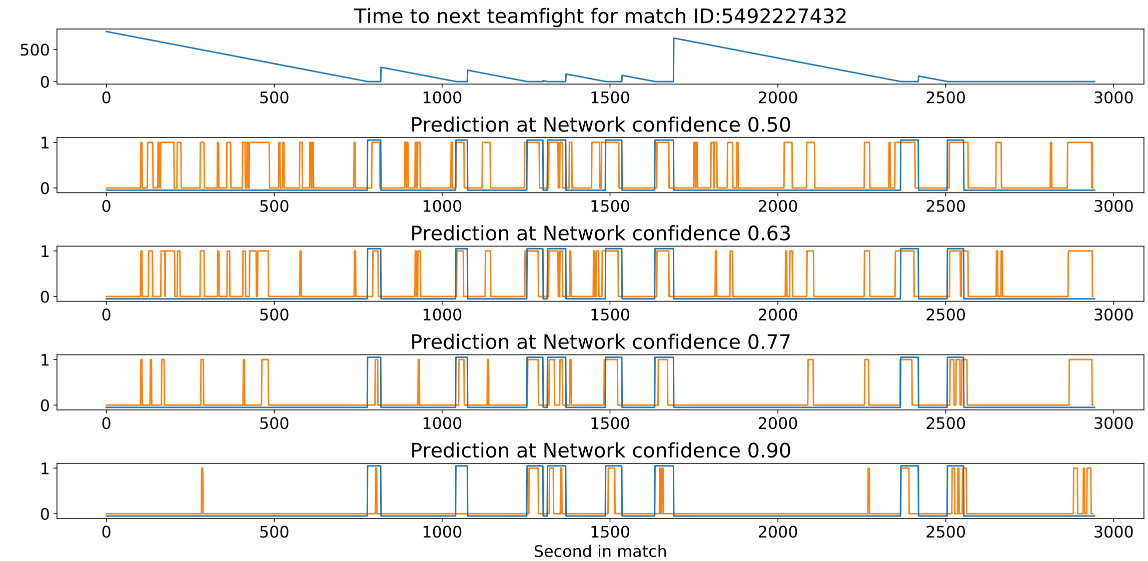

Fig. 6. Testing on an entire match with different confidence thresholds.

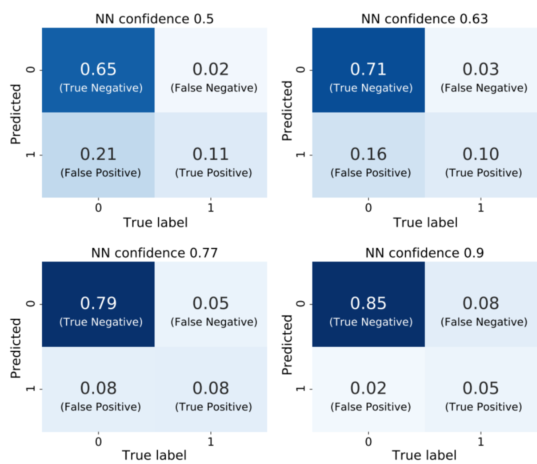

Fig. 7. Confusion matrices for different confidence thresholds.

to provide more precise labelling for the start and end of the conflict based on the literature definition of engagements, despite being trained on these inconsistent labels. This is further evidence of the potential of the methodology which uses and benefits from player expertise.

Due to the fact that the input data consists only of in-game character and camera positions, the model does not receive any information about character state, and is not able to estimate how many players would die in an encounter, and thus make a distinction between an encounter and a team fight. However, the small quantity of false-negative predictions suggests that even without detecting the difference between team fights and encounters, using player and camera position can be a reliable way of predicting and detecting both team fights and similar encounters. This supports the assumption that the camera position reflects aspects of player knowledge as well as their strategy. The proposed network can detect patterns in this data and make relevant predictions without the need to explicitly model this complex process of human decision making.

VII. Conclusion

This paper address the predictability of a team fight using a small number of input parameters. In-game character position, in addition to their player cameras, are used to make observations and conclusions to predict and detect those encounters. Using this data to train a two-part neural network, the model was able to achieve 84% accuracy. However, despite the high achieved accuracy, when evaluated on an entire match, the model possessed lower f1 scores. Through further investigation, inconsistencies and potential issues with existing industry labels for team fights were identified. These inconsistencies influenced both the training and the evaluation process. As the obtained predictions outperform the industry labels, they could be used in conjunction with other existing formal definitions in the literature to provide a better definition for the concept of a team fight, allowing for more meaningful and reliable labels for future work in the area.

Moreover, despite the inaccuracies encountered with labelling, it is clear that the model possesses predictive capabilities. This indicates that the employed methodology, which takes advantage of players expertise to derive conclusions is a reliable way of extracting meaningful information. Due to the high complexity of the environment, models that attempt to interpret the game entirely from raw data may be replicating already available human understanding. In some cases, such as player planned encounters, this understanding could be derived from available or otherwise existing data sources.

This paper presents a novel approach for extracting player intent and utilising this data. No studies could be found - at present - that utilise player camera to replicate or derive player intent and decision making. The obtained results provide a clear indication of the potential of this technique.

The model described in this paper could be used in the industry to enhance game coverage for audiences. One promising example is implementing this technique in existing apps that are used in the broadcast of such titles. Another possible application is to make the results available for camera operators in the production of esport events, to assist in directing the focus of the coverage.

VIII. Future Work

The study described in this paper reveals several avenues of research. Considering other game information, in addition to the existing camera and character position, could be relevant in differentiating between encounters and team fights. In a game such as Dota 2, there are many factors effecting the likelihood of a specific team winning an engagement and the number of deaths that are going to happen in a team fight. Character levels, character roles, gold and experience difference between teams, and game time may all have an impact on the decision of whether a team commits to a team fight or not. Adding the information about the character state could allow the neural network to predict deaths, which could make it suitable for distinguishing between specific types of engagements.

Furthermore, the techniques employed in this paper could be used in different models to improve their performance by relying on players expertise to reduce the complexity of the problem space. Detecting and predicting when a player is planning an engagement could be used as a factor for determining the potential winner of an encounter, and indeed, the outcome of the game. This could serve to aid in the win prediction domain.

Finally, some games may include player-made markers, i.e. pings, to aid rapid communication between teammates. In these cases, a similar approach, as described in this paper, could be used to derive other forms of meaningful information through player intent. Looking into the correlation between pings and team fights could provide additional insight into player's expert knowledge. Combined with the camera positions a model could attain more in-depth modelling of player behaviour and take greater advantage of their expertise.

Acknowledgement

This paper is part-funded by EPSRC CDT in Intelligent Games and Game Intelligence (IGGI) EP/L015846/1 and the Audience of the Future program by UK Research and Innovation through the Industrial Strategy Challenge Fund (grant no.104775) as part of the Weavr project (weavr.tv).

Ethics Approval

Ethics approval was granted by the computer science department at the University of York. Data was collected through readily and freely available means and no personal or otherwise identifiable information was collected, stored or utilised at any stage beyond the publicly available replay game files. No identifiable information was extracted from those files and data was stored in an aggregate format to prevent deanonymisation.

References

[1] R. Turtiainen, U. Friman, and M. Ruotsalainen, ""not only for a celebration of competitive overwatch but also for national pride": Sportificating the overwatch world cup 2016," *Games and Culture*, vol. 15, no. 4, p. 351–371, Jun. 2020. [Online]. Available: https://doi.org/10.1177/1555412018795791

[2] J. Gawrysiak, R. Burton, S. Jenny, and D. Williams, "Using esports efficiently to enhance and extend brand Perceptions–A literature review," *Physical Culture and Sport. Studies and Research*, vol. 86, no. 1, p. 1–14, 2020. [Online]. Available: https://content.sciendo.com/view/journals/pcssr/86/1/article-p1.xml

[3] J. G. Reitman, M. J. Anderson-Coto, M. Wu, J. S. Lee, and C. Steinkuehler, "Esports research: A literature review," *Games and Culture*, vol. 15, no. 1, p. 32–50, Jan. 2020. [Online]. Available: https://doi.org/10.1177/1555412019840892

[4] J. Elam and N. Taylor, "Above the action: The cultural politics of watching dota 2," *Commun Cult Crit*, Feb. 2020. [Online]. Available: https://academic.oup.com/ccc/article-abstract/doi/10.1093/ccc/tcz033/5730340

[5] M. Mora-Cantallops and M.-A. Sicilia, "Team efficiency and network structure: The case of professional league of legends," *Social Networks*, vol. 58, p. 105–115, Jul. 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0378873319300693

[6] cyborgmatt. Dota 2 prize pool tracker. [Online]. Available: https://dota2.prizetrac.kr/

[7] A. Kokkinakis, S. P. Demediuk, I. Nolle, O. Olarewaju, S. Patra, J. Robertson, P. York, A. Pedrassoli Chitayat, A. Coates, D. Slawson *et al.*, "Dax: Data-driven audience experiences in esports," in *Proceedings of ACM International Conference on Interactive Media Experience (IMX)*. Association for Computing Machinery (ACM), 2020, p. 1–12.

[8] B. Hutchins, B. Li, and D. Rowe, "Over-the-top sport: live streaming services, changing coverage rights markets and the growth of media sport portals," *Media, Culture & Society*, vol. 41:7, pp. 975–994, 2019.

[9] F. Block, V. Hodge, S. Hobson, N. Sephton, S. Devlin, M. F. Ursu, A. Drachen, and P. I. Cowling, "Narrative bytes: Data-Driven content production in esports," in *Proceedings of the 2018 ACM International Conference on Interactive Experiences for TV and Online Video*, ser. TVX '18. SEOUL, Republic of Korea: Association for Computing Machinery, Jun. 2018, p. 29–41. [Online]. Available: https://doi.org/10.1145/3210825.3210833

[10] A. Katona, R. Spick, V. J. Hodge, S. Demediuk, F. Block, A. Drachen, and J. A. Walker, "Time to die: Death prediction in dota 2 using deep learning," in *2019 IEEE Conference on Games (CoG)*. Queen Mary University of London: ieeexplore.ieee.org, Aug. 2019, p. 1–8. [Online]. Available: http://dx.doi.org/10.1109/CIG.2019.8847997

[11] A. P. Chitayat, A. Kokkinakis, S. Patra, S. Demediuk, J. Robertson, O. Olarewaju, M. Ursu, B. Kirmann, J. Hook, F. Block *et al.*, "Wards: Modelling the worth of vision in moba's," in *Science and Information Conference*. Springer, 2020, pp. 63–81.

[12] N. Kinkade, L. Jolla, and K. Lim, "Dota 2 win prediction," 2015. [Online]. Available: http://jmcauley.ucsd.edu/cse255/projects/fa15/018.pdf

[13] V. J. Hodge, S. M. Devlin, N. J. Sephton, F. O. Block, P. I. Cowling, and A. Drachen, "Win prediction in Multi-Player esports: Live professional match prediction," *IEEE Trans. Comput. Intell. AI Games*, p. 1–12, Oct. 2019. [Online]. Available: http://eprints.whiterose.ac.uk/152931/

[14] S. Demediuk, P. York, A. Drachen, J. A. Walker, and F. Block, "Role identification for accurate analysis in dota 2," in *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, vol. 15, no. 1, 2019, p. 130–138.

[15] M. Schubert, A. Drachen, and T. Mahlmann, "Esports analytics through encounter detection," in *Proceedings of the MIT Sloan Sports Analytics Conference*, vol. 1. Boston, Massachusetts: sloansportsconference.com, 2016, p. 1–18.

[16] Valve Corporation. Open source dota 2 data platform. [Online]. Available: https://www.opendota.com/

[17] spheenik, "Clarity analyzer," https://github.com/spheenik/clarity-analyzer, 2015.

[18] A. L. Maas, A. Y. Hannun, and A. Y. Ng, "Rectifier nonlinearities improve neural network acoustic models," in *Proc. icml*, vol. 30, no. 1. Citeseer, 2013, p. 3.

[19] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

APPENDIX A
COMPLETE MATCH ANALYSIS

Full analysis of the positive predictions is presented in Table III below. All of the predictions are calculated with a confidence threshold of 0.77 on the match presented in Figure 6. The criteria for the engagement was taken from [15], and is as follows:

1) At least three heroes in range of each other.
2) Heroes of both teams present.
3) Transaction of damage from one team to the other.

TABLE III
ANALYSIS OF THE POSITIVE PREDICTIONS

Event Number	Game Time	Event Description	Network Prediction	OpenDota Label
1	0:17	A small engagement on the top lane, Enchantress fires a couple of shots at 2 Dire characters going through the lane.	1	0
2	0:41	Radiant heroes engage in 3v2 on the top lane, Morphling gets caught out of position. First kill of the game.	1	0
3	1:16	Top lane, 3v2 fight, lasts for 8 seconds. Nobody dies.	1	0
4	1:33	Top lane, 1v2, Enchantress doing some damage to the enemy heroes. One character very close to dying.	1	0
5	3:09	Engagement starts on top, 4 seconds later another one starts on bottom. Top engagement ends with one death on the Dire side, no casualties on the bottom.	1	0
6	5:23	Phoenix and Morphling start trading shots at 5:16. Additional characters come to aid both sides at 5:23, Morphling starts running away but gets chased down.	1	0
7	5:51	A small engagement on top, dire characters teleports in and gets engaged on. Engagement finishes in a couple of seconds. No deaths.	1	0
8	6:20	Two engagements at the same time at last for 15 seconds. 6 characters present in the engagement in the top lane, 4 characters on the bottom one. Two heroes die, one in each engagement.	1	0
9	11:39	A big team fight starts. At first, only 3 characters present, others joined during the duration of the fight. The fight lasts for 30 seconds with 3 characters dying.	1	1
10	13:57	No engagement at this moment. There are several heroes close by, but nothing happens.	1	0
11	14:04	A small engagement on top, 2v1. One character dies, short engagement, only lasts for 4 seconds.	1	0
12	15:56	A team fight happening near the top lane. Starts with a small engagement, heroes from both sides join the fight. The fight ends in 20 seconds with 4 deaths.	1	1
13	17:30	Engagement near the bottom lane on the Dire side of the map, Morphling (D) was caught out of position but runs away quickly.	1	0
14	19:24	The biggest team fight of the game. All characters from both sides are present, the fight lasts for 35 seconds, three characters die.	1	1
15	20:35	A small engagement that starts a few seconds before (20:28) turns into a bigger one after several characters join the fight. 5 heroes die.	1	1
16	21:37	A bit mistimed prediction. Just as an engagement finished.	1	0
17	23:19	The radiant team goes into the Dire area near their base, Dire engages, 4 Radiant characters and 1 Dire character dies. The fight lasts for 40 seconds.	1	1
18	26:01	A team fight in the middle of the map. Engagement starts slowly but ends with 7 characters losing their lives. Engagement lasts for 50 seconds.	1	1
19	29:42	Engagement starts at 29:32 and ends at 29:40. The prediction was two seconds late.	1	0
20	32:36	Another small engagement happens before the prediction, the prediction happens two seconds after the engagement ends.	1	0
21	33:28	Engagement near the Dire base. 9 out of 10 characters are present, lasts for 25 seconds.	1	0
22	36:18	All 10 characters present. Relatively quick engagement on the Radiant side of the map, two Dire characters die.	1	0
23	37:51	Probably the game-deciding team fight. Both teams are ready, every character is in the position, and prepared for the fight. The engagement starts at 37:58 and lasts until 38:38. All of the Radiant characters die.	1	1
24	40:33	Team fight in front of the Radiant base. Long engagement, all of the characters present, multiple characters die, revive and come back to the fight. Radiant successfully defend the base.	1	1
25	46:29	Last engagement in the game. Starts with a chase, and ends with the Dire team winning the game. The duration of the fight was 58 seconds, all of the characters on the Radiant side die.	1	0