



Analysis of Enthusiasm Triggers in Competitive Video Games: Overwatch Case Study

Alexis Mortelier, Boris Helleu, Wiem Belhedi, François Rioult

► To cite this version:

Alexis Mortelier, Boris Helleu, Wiem Belhedi, François Rioult. Analysis of Enthusiasm Triggers in Competitive Video Games: Overwatch Case Study. International Workshop on Intelligent Technologies for Precision Sports Science 2024 (IT4PSS), Aug 2024, Jeju Island, South Korea. <hal-04620944>

HAL Id: hal-04620944

<https://hal.science/hal-04620944v1>

Submitted on 22 Jun 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Analysis of Enthusiasm Triggers in Competitive Video Games: Overwatch Case Study

Alexis Mortelier¹, Boris Helleu², Wiem Belhedi³, François Rioult¹

¹Normandie Univ, UNICAEN, ENSICAEN, CNRS, GREYC, 14000 Caen, France

²NIMEC-UR 969, Université de Caen Basse-Normandie, 14000 Caen, France

³LrSITI, ENIT, Université Tunis El Manar, 1002 Tunis, Tunisie

alexis.mortelier@unicaen.fr

Abstract

We propose a methodology for analyzing viewer engagement during Overwatch League (OWL) broadcasts. On the one hand, we implemented image processing techniques to isolate key visual elements of the spectator interface in OWL videos, enabling an objective and granular analysis of the visual experience, qualified by measures of interest. On the other hand, we extracted audio descriptors from OWL videos of competitive events. We believe these audio descriptors are indicative of the excitement of the crowd and commentators, offering a subjective measure of the impact generated by the video broadcast of the encounter. A *wrapper* approach allowed us to characterize the relevance of the measures of interest by examining their relationship to the hype descriptors.

weather conditions, the atmosphere in the stadium, the design of the equipment, the stakes and the uncertainty of the outcome. These different variables affect spectators differently, depending on their interests and needs, so that there are many different ways of consuming a sporting spectacle.

The study of competitive matches in Overwatch is motivated by the desire to understand and improve the spectator experience. By comparing the spectator interface with the user experience, we aim to determine how what happens on screen can be better understood and appreciated by the audience. To do this, we need to assess the enthusiasm of the crowd and the commentators, and then compare it with what is presented visually.

Analyzing measures of interest in competitive Overwatch matches is then crucial not only for improving the live viewing experience, but also for creating game summaries that captivate and retain viewers, thus contributing to the growth and popularity of e-sports.

1 Introduction

Data analysis in the field of video games has become a major area of interest in recent years. The data generated by video games is extremely rich and varied, including information on player behavior, in-game performance, social interactions, preferences and more. This wealth of information offers game developers, researchers and analysts a unique opportunity to better understand game mechanics, optimize the player experience and even detect emerging trends.

One game of particular interest for data analysis is Overwatch. Overwatch is a first-person shooter developed by Blizzard Entertainment, released in 2016. It quickly gained popularity and became a flagship title in the field of competitive shooters. The game stands out for its many characters with unique skills, its various game modes and its well-established competitive aspect.

The consumption of sports entertainment is a complex phenomenon that has given rise to a variety of academic approaches. Sociologists, psychologists, economists and specialists in management and marketing have shed light on a range of seemingly simple questions: why and how do people watch sport? The literature is full of studies analyzing the importance of factors that mobilize fans in the stadium or in front of a screen: the performance of their favorite team, the talent and popularity of the players, the price of tickets,

2 State of the art

Here, we outline some of the literature on spectator engagement. We then look at Overwatch and his studies, and conclude with a review of studies in the field of crowd clamor analysis.

2.1 Spectator commitment

[Stewart *et al.*, 2003] review the usefulness and diversity of spectator typologies. They show how these models reveal complex fan behaviors, depending on whether they are sensitive to several cognitive and affective factors: emotional connection and identification with a team, search for social interaction or the excitement generated by an event. This makes it difficult to identify the archetypal sports consumer. According to these authors, these typologies enable marketers to refine their strategy, adapt product and pricing, design appropriate equipment, adjust communication campaigns and personalize a consumer experience according to each fan's needs. Essentially, this means understanding fan motivations and behaviors. Daniel C. Funk and Jeff James, at the interface of psychology and marketing, go beyond the notion of satisfying a need to produce theoretical models adapted to sport, such as the Psychological Continuum Model [Funk and James, 2001] or the Fan Attitude Network (FAN) Model [Funk and James, 2004]. So, in

the field of traditional sport, the matchday experience is the subject of wide-ranging academic investigation [Meir, 2000; Richards *et al.*, 2022].

By mobilizing the interest of numerous fans, sponsors and media around the world, esports has become a scientific object taken seriously. Indeed, according to Newzoo's 2022 Global Esports & Live Streaming Market Report, the global esports fan audience will reach 640 million by 2025, including 318 million enthusiastic fans (those who watch this type of content at least once a month). Newzoo estimates that the sector will have generated sales of \$1.38 billion by the end of 2022. So questions about how sports are consumed have found fertile extensions in the esports field. [Hamari and Sjöblom, 2017] are asking what esports is and why it is being watched. According to [Voisin and Besombes, 2023], from practice to mediated spectacle, esports covers a wide spectrum of experiences, so much so that we can speak of esports audiences. However, the first works dedicated to esports fans favor comparisons with "traditional" sports. In the article "eSport vs. Sport: A Comparison of Spectator Motives", [Pizzo *et al.*, 2018] compares spectator motivations for a traditional sport (soccer) and two esports contexts (FIFA Online 3 and StarCraft II) in South Korea. The authors listed 15 motivations for attending events (interest, excitement, enjoying players, competitive intensity, meeting people or sharing a moment with family, athletes' skill, etc.). According to the authors, 11 of the top 15 fan motivations are common to all three events denoting a certain similarity in sports and esports consumption modalities.

However, esports has a unique feature. While, like traditional sports, it manages to attract large crowds to stadiums and arenas, it also mobilizes the interest of fans on streaming platforms that enable remote interaction. Numerous studies are therefore focusing on online consumption. [Jang *et al.*, 2024] ask: "What makes esports consumers watch streamers' esports live-streaming contents?". [Qian *et al.*, 2020] have developed the Scale for Esports Spectator Demand (SESD) model, which isolates esports-specific factors: stream quality, chat room, online rewards. While demand for sports entertainment is largely explained by the uncertainty of the outcome of a live match, [Wang, 2023] focuses on asynchronous consumption of replays, and on the popularity and skills of players. Skills are particularly valued in FPS games with an assertive technical dimension, while player popularity will be more valued in a MOBA-game¹ that leaves room for entertainment and socialization. As stream viewers are themselves fans of videogame competitions, watching the best players is also a way of improving their skills [Hattingh *et al.*, 2024; Qian *et al.*, 2020].

More recent contributions attempt to synthesize the existing literature. [Rietz and Hallmann, 2023] have analyzed 25 articles dedicated to esports fan behavior. According to their survey, 22 contributions are devoted to online spectators, two to on-site fans and just one compares online/on-site. Back in 2016, Taylor noted that at the turn of the 2010s, attendance at esports events evolved from player-spectators to fans. The

¹MOBA stands for Multiplayer Online Battle Arena, such games in which two teams of heroes battle out to take on an objective.

study conducted by [Jang and Byon, 2020] highlights unique elements associated with esports events, such as cosplay but also cheering behaviors, which enrich fans' experience and motivate their intention to attend future events.

2.2 e-sport analytics

The e-sport phenomenon has emerged as a new era of sport, highlighting the opportunities it offers for sports scientists [Himmelstein and Shapiro, 2020]. The way in which traditional sports and e-sport interact implies that e-sport fits into the wider landscape of professional sport and can benefit from the experience and infrastructure of traditional sports [Castro-nova, 2020].

Data analysis in e-sport can be used to assess individual player performance, indeed understanding the specific physical requirements of games in order to optimize player [Goto and Rynkiewicz, 2019] is important. This can include measures such as shot accuracy, responsiveness, reaction times and so on. By identifying the areas in which a player excels or struggles, coaches can adapt training to reinforce the necessary skills. For example, a study conducted by [Schulte *et al.*, 2019] showed that the use of biometric data and sensors can help assess and improve the performance of e-sport players.

Data analysis can also be used to study the performance of the team as a whole, as well as that of its opponents. Teams can examine patterns of play, strategies used by opponents, and trends in play to adjust their own strategies [Patel and Lindley, 2020]. For example, research by [Wang *et al.*, 2019] used data analysis to identify the patterns of play and strategies used by top teams in the Dota 2 game, providing insights into best practice and possible counter-strategies.

Data analysis in e-sports can involve several methods, extracting and analyzing data from the games themselves [Johnson, 2020]. This can include data on player actions, match results [Mortelier and Rioult, 2022], map movements and so on. Data can be collected using the game API or third-party software [Drachen *et al.*, 2014]. The specific layout of the game map affects team strategies and match outcomes, using game data to identify patterns and trends related to the map [Reimers and Goldman, 2019].

Analyzing historical match data can provide valuable insights into game trends, player and team performance, strategies used, and so on. For example, [Schubert M., 2010] have examined user behavior in video games by analyzing historical game data to understand player preferences.

Game developers can use player data to improve game experience, game design, and player retention. It could also address the methods and tools used in game analytics, as well as the challenges and opportunities associated with using massive data from video games [Drachen *et al.*, 2013].

2.3 Overwatch (strategic FPS)

Overwatch is a game in which two teams of six players play against each other. It is halfway between an FPS (first-person shooter) and a MOBA. Each player can embody a hero who can be changed at will during the course of the game.

Heroes have their own specific abilities, as well as an ultimate ability that has a major impact on the course of the

game. Each hero belongs to a category: tank, support or damage. The composition of a team is crucial: a team with heroes with good synergy is vital to the outcome of the game. Matches are played in 4 different modes: checkpoint, escort, point attack and hybrid attack/escort. Each game mode places one team on attack and another on defense, then reverses roles when the time or objective is reached. When a hero dies, it takes 10 seconds for him to reappear.

Among the few studies on Overwatch, [Tomkinson and Van Den Ende, 2022] looks at the issue of toxicity in online multiplayer games, highlighting the strategies put in place by developers to deal with it. Similarly, [Välisalo and Ruotsalainen, 2019] look at how players and fans engage with Overwatch and its nascent esports ecosystem, highlighting the importance of personality, gender and sexual orientation in these engagement dynamics. Their study offers a comparative look at how players interact with the game’s playable characters and with professional esports players, highlighting similarities and emerging differences. Finally, [Vahlo and Karhulahti, 2022] bring a more quantitative perspective by using large-scale survey data to map the characteristics of esports players. Their cluster analysis identifies two distinct profiles of esports players, contributing to our understanding of the diversity of esports players. These studies underline the growing importance of Overwatch as an academic research field, while highlighting the ongoing need to further explore this game and its impact on video game and esports culture.

2.4 Analysis of audio clamor

[Proksch *et al.*, 2023] review advances in the acoustic detection of crowd dynamics, highlighting the use of techniques such as Fourier analysis, spectroscopy and neural network-based methods to extract meaningful features from audio data. [Dosbayev *et al.*, 2021] offer an overview of audio signal processing methods for detecting abnormal behavior in crowds, highlighting descriptors such as sound pressure level and spectral density. [Franzoni *et al.*, 2020] focus on using audio signals to detect emotions in crowds, exploring techniques such as vocal intensity analysis and prosody models. Reviews of [Başbuğ and Sert, 2019; Crocco *et al.*, 2016; Schuller, 2018; Khalil *et al.*, 2019; Lefèvre *et al.*, 2021; Rao *et al.*, 2013] provide valuable clues about the use of advanced techniques such as deep learning and prosodic analysis to automatically extract complex audio descriptors from crowd data. This research underlines the growing importance of audio analysis for detecting abnormal behavior and emotions in crowds.

3 Audio analysis of the crowd surge

The method we are proposing will enable us to highlight what arouses the interest and enthusiasm of spectators at official Overwatch games. As Overwatch does not have a native API for extracting data, we have developed a video analysis tool. This is designed to extract relevant information from the spectator overlay present in these videos. The data extracted from the video is transformed into *interest measures*, while the audio is converted into *audio descriptors* revealing the clamor of the crowd and the enthusiasm of the commentators. The

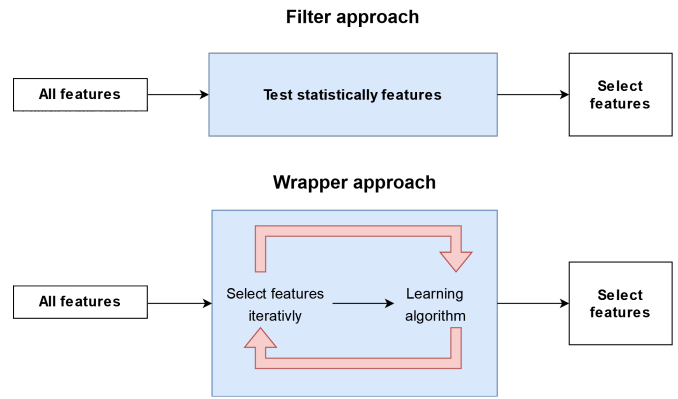


Figure 1: *filter* approach vs. *wrapper*.

inclusion of commentators in crowd noise to assess enthusiasm in a competitive game is controversial. Commentators can artificially create an enthusiastic atmosphere to captivate spectators, which may not accurately reflect the reality of the game. Their behavior may differ from that of other members of the crowd. However, they often play a crucial role in amplifying key moments and adding context, thus intensifying the general enthusiasm. Their presence is therefore considered an integral part of the atmosphere at a competitive event. Commentators strongly influence spectator perception, their enthusiasm being contagious and increasing the audience’s emotional involvement. Ignoring them could give an incomplete picture of the enthusiasm generated. Finally, including commentators helps to capture a more holistic picture of the sound environment of a competitive game. Moments of silence, sudden exclamations and crowd interaction add to the authenticity of the experience.

To achieve the highlighting of the enthusiasm, a *wrapper approach* is employed for assessing every measure’s relevancy through the learning of relation with all audio descriptors described in Section 3.3.

3.1 Wrapper approach

The wrapper approach [Kohavi and John, 1998] is originally an *attribute selection* technique that pronounces on the relevance of descriptors by examining the training performance of a predictive system. Indeed, a distinction is made between the *filter* [Germano, 1992] approach, which examines the statistical characteristics of attributes (such as correlation coefficients) to choose among them, and the *wrapper* approach, which iteratively selects attributes that induce good performance from the learning algorithm employed (see Figure 1).

The *wrapper* approach thus performs a search in the space of attribute subsets, guided by the performance of the learning algorithm. The evaluation criterion is simply the learning performance. The *wrapper* approach enables finer selection than the *filter* approach, but at the cost of a longer computation time. We use it here as a feature importance method.

3.2 Data acquisition - Analysis of transcript images

Overwatch has no native API allowing users to extract data

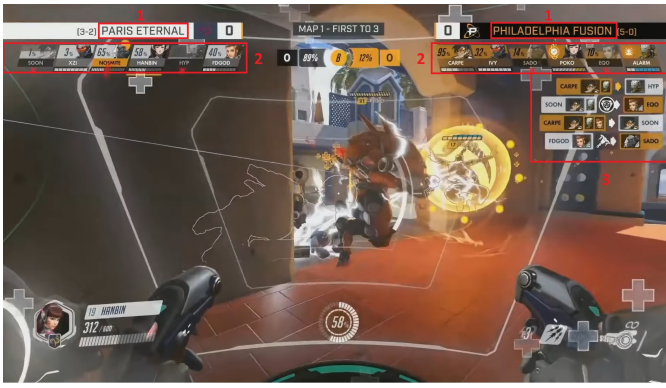


Figure 2: An image from the transcription of a match of Overwatch.



Figure 3: An image of the killfeed from Overwatch.

directly from a current match or recent game. Fortunately, many Overwatch matches are recorded and uploaded to video platforms such as YouTube. These videos contain game footage captured with the spectator interface (Figure 2), which offers an overview of the match and the various player actions: damage inflicted, healing provided, death events, ultimate ability status, ultimate ability usage. This overlay provides *objective* information on the state of the match.

To exploit this data present in videos, we have developed a specialized replay visualization tool². This tool is designed to extract relevant information from the viewer overlay. Thanks to image analysis techniques (OCR, feature matching, pattern matching, erosion and dilation), this information can be captured from the spectator interface, providing a precise status of the game.

For example, Zone 1 on Figure 2 shows the names of the two teams and the score. The team compositions, each player's hero, the hero's life, the state of his ultimate ability

²<https://github.com/Nikkunemufr/Python/tree/master/MetaWatch>

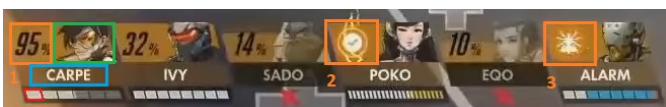


Figure 4: Team hero banner (zone 2 on the right of Figure 2).

and the player's nickname associated with this hero are displayed in Zone 2 (see Figure 4). In Zone 3, the killfeed (see Figure 3) shows the last six kills and/or destructions, with the most recent at the top. Both names are given the color appropriate to the team, and a player hero icon is placed next to them (boxed in green). The hero icons of players taking part in the kill are displayed next to them (boxed in yellow). The white arrow indicates the direction of the kill; if it is colored, it is a critical kill (Zone 1); if it is an ability (Zone 3) or an ultimate ability (Zone 2) that kills, the ability logo is shown next to it.

Each game video is then subjected to an analysis of the extracted data, beginning with the evaluation of specific measures for each category : damage suffered (in DPS), healing received (in HPS), ultimate ability status, ultimate ability use, hero death. The events of death and use of ultimate abilities are extended over 10 seconds (corresponding to the reappear-ance time) for each successive use or death. Each player's individual interest in each category is then aggregated to create a category interest.

3.3 Audio descriptor extraction

Audio analysis is useful for understanding the emotional and dynamic aspects of live events, such as Overwatch matches. For this task, a set of audio descriptors were extracted using Librosa, a Python library specialized in audio signal processing [McFee *et al.*, 2015]. These descriptors, selected for their ability to capture different aspects of the soundscape, offer a rich perspective on the variations in crowd excitement and commentator enthusiasm throughout the match.

Firstly, MFCC coefficients (*Mel-Frequency Cepstral Coefficients*) are used to represent the spectral shape of the audio signal [Davis and Mermelstein, 1980]. These coefficients, often used in speech recognition, capture frequency characteristics at different time scales. In the context of football matches, variations in MFCCs can reveal fluctuations in crowd mood and commentator excitement at key moments of the game.

STFT chroma (*Short-Time Fourier Transform*) is then used to analyze the energy distribution of harmonic components in different pitch classes, providing information on tonality and tonal variations in the audio [Fujishima, 1999]. This measurement can be linked to background music or sounds characteristic of in-game actions, enriching the analysis of the overall soundscape.

The spectral centroid, which indicates the center of gravity of frequencies in the audio signal, can reveal information about the brightness of the sound. High values of this measure can indicate an increased presence of high-pitched sounds, such as those produced by screams or applause in a stadium during moments of high intensity.

Spectral bandwidth and spectral contrast respectively provide information about the variety of frequencies present in the audio and sudden changes in the sound texture [Peeters, 2003; Baer *et al.*, 1993]. These measures can be used to detect moments of intense excitement or relative calm during the match.

Zero crossing rate and onset strength are used to assess the amount of variation and the presence of transients in the au-

dio signal [Gouyon *et al.*, 2000] [Bello *et al.*, 2005]. These aspects are essential for understanding the percussive or sonic character of audio, highlighting moments of action or strong emotion.

Tempo, power spectrogram, pitch and energy are also key descriptors used to capture the perceived rhythm, overall intensity and pitch of sounds in audio [Dixon, 2001; Klapuri, 2004; Ghaemmaghami *et al.*, 1997; Ramirez *et al.*, 2004]. These measures can be used to highlight moments of crowd excitement, commentator enthusiasm and other aspects of the game’s soundscape, providing a rich and comprehensive analysis of live events.

The use of these varied audio descriptors enables an in-depth and nuanced analysis of the soundscape of Overwatch matches, providing information on crowd excitement, commentator enthusiasm and other emotional aspects of the game in real time.

4 Experiments

4.1 Dataset

We processed data from 14 Overwatch matches, giving rise to sixty different maps (i.e. games), for a total duration of twenty hours of video. This extraction yielded 60 sequences ranging in size from 900 to 1 800 one-second time steps. For each time step of these sequences, descriptors of interest are obtained on damage (damage), healing (heal), death (death), loading of ultimate abilities (state of ulti), use of ultimate abilities (use of ulti) as well as descriptors from the audio tape of the transcripts.

4.2 Model structure

The *wrapper* approach was implemented as follows: as many models as there are descriptors combinations are trained in regression to predict the audio descriptors. The training performance (Mean Squared Error – MSE, and number of epochs for convergence) of the model was evaluated to determine the relevance of each descriptor combination for learning all the audio descriptors.

The trained model was a gated recurrent neural network (GRU) rather than a Long Short-Term Memory (LSTM) for several reasons. Firstly, there is no long-term time dependency (a strength of LSTMs) which is critical to this specific task. Unlike LSTMs, GRUs are designed to capture short- and medium-term temporal dependencies. This simplification reduces the complexity of the model, which can help it to converge faster and train more efficiently.

4.3 Results

Table 1 shows the different combinations of measures and their performance for audio descriptor prediction, measured with MSE. For each descriptor, Table 2 shows the average MSE for the models trained *without* a combination including that descriptor, the average MSE *with* that descriptor, and finally the percentage improvement (remember that the lower the MSE, the better it is). The *improvement* column of the Table 2 therefore provides an indication of the relevance of each descriptor for the MSE. A positive value suggests that the

heal	damage	use of ulti	state of ulti	death	MSE
×	×	×	×	×	0.351
	×	×	×	×	0.352
×	×		×	×	0.355
×	×		×		0.356
×	×	×	×		0.357
×		×	×	×	0.357
×			×	×	0.361
					0.363
	×	×	×		0.365
	×		×		0.365
×	×	×			0.365
×	×	×		×	0.366
×			×	×	0.367
×		×		×	0.367
	×	×			0.374
×		×			0.375
×	×			×	0.394
	×	×		×	0.439
	×		×	×	0.441
	×	×	×		0.444
×	×				0.456
	×			×	0.468
×				×	0.478
	×				0.521
×					0.534
		×		×	0.710
			×	×	0.793
		×	×	×	0.793
				×	0.819
			×		0.872
		×	×		0.880
			×		0.881

Table 1: MSE for each descriptor combination.

measure	without (MSE)	with (MSE)	improvement (%)
heal	0.55	0.38	31.4
damage	0.53	0.39	26.2
use of ulti	0.45	0.45	-1.8
state of ulti	0.44	0.46	-6.2
death	0.43	0.47	-9.5

Table 2: Contribution of each descriptor to the MSE.

423 presence of the measure improves model convergence, while
 424 a negative value indicates a deterioration in prediction.

425 Analyzing the results, we observe that for the heal and
 426 damage measures, their inclusion improves prediction accu-
 427 racy (reduced MSE). This suggests that these measures pro-
 428 vide important information for the model, making it easier to
 429 learn and improving its predictive performance.

430 On the other hand, the inclusion of ultimate ability state
 431 and death events leads to a degradation of the MSE. This may
 432 indicate that these measures introduce noise or information ir-
 433 relevant to the prediction task, disrupting the learning process
 434 and reducing prediction accuracy.

435 Analysis of this table suggests that not all measures con-
 436 tribute equally to improving model performance. Some indi-
 437 cators appear to be more relevant than others, underlining the
 438 importance of carefully selecting the features to be included
 439 in the learning process to obtain the best results.

440 5 Conclusion

441 We proposed a method for analyzing spectator engagement
 442 in Overwatch League broadcasts, examining the impact on
 443 crowd engagement of information provided by the game. The
 444 results show that heal and damage indicators significantly im-
 445 proved the performance of the crowd engagement prediction
 446 model, while features relating to ultimate abilities disrupted
 447 the learning process.

448 This study highlights the importance of the relevance of
 449 the data used to analyze spectator engagement. Identifying
 450 the most significant indicators is a step towards understanding
 451 and improving the spectator experience.

452 Acknowledgements

453 The research described here is supported by the French Na-
 454 tional Research Agency (ANR) and Region Normandie under
 455 grant HAISCoDe

456 References

457 [Baer *et al.*, 1993] Thomas Baer, Brian CJ Moore, and Stu-
 458 art Gatehouse. Spectral contrast enhancement of speech in
 459 noise for listeners with sensorineural hearing impairment:
 460 Effects on intelligibility, quality, and response times. vol-
 461 ume 30, pages 49–49. Citeseer, 1993.

462 [Başbuğ and Sert, 2019] Ahmet Melih Başbuğ and Mustafa
 463 Sert. Analysis of deep neural network models for acoustic
 464 scene classification. pages 1–4, 2019.

465 [Bello *et al.*, 2005] Juan Pablo Bello, Laurent Daudet, Samer
 466 Abdallah, Chris Duxbury, Mike Davies, and Mark B Sandler.
 467 A tutorial on onset detection in music signals. vol-
 468 ume 13, pages 1035–1047. IEEE, 2005.

469 [Castronova, 2020] Edward Castronova. The surprising im-
 470 pact of professional sports on esports. *Journal of Sports*
 471 *Economics*, 21(7):705–717, 2020.

472 [Crocco *et al.*, 2016] Marco Crocco, Marco Cristani, Andrea
 473 Trucco, and Vittorio Murino. Audio surveillance: A
 474 systematic review. *ACM Computing Surveys (CSUR)*,
 475 48(4):1–46, 2016.

[Davis and Mermelstein, 1980] Stephen Davis and Paul 476
 Mermelstein. Mel frequency cepstral coefficients (mfccs). 477
 In *Speech and Audio Signal Processing*, 1980. 478

[Dixon, 2001] Simon Dixon. Automatic tempo and beat 479
 analysis of acoustic musical signals. In *Journal of New* 480
Music Research, volume 30, pages 39–58, 2001. 481

[Dosbayev *et al.*, 2021] Zhandos Dosbayev, Rustam Ab- 482
 drakhmanov, Oxana Akhmetova, Marat Nurtas, Zhalgas- 483
 bek Iztayev, Lyazzat Zhaidakbaeva, and Lazzat Shaimer- 484
 denova. Audio surveillance: detection of audio-based 485
 emergency situations. pages 413–424, 2021. 486

[Drachen *et al.*, 2013] Anders Drachen, Magy Seif El-Nasr, 487
 and Alessandro Canossa. *Game analytics: maximizing the* 488
value of player data. Springer, 2013. 489

[Drachen *et al.*, 2014] Anders Drachen, Matthew Yancey, 490
 John Maguire, Derrek Chu, Iris Yuhui Wang, Tobias 491
 Mahlmann, Matthias Schubert, and Diego Klabajan. Skill- 492
 based differences in spatio-temporal team behaviour in de- 493
 fence of the ancients 2 (DotA 2). In *2014 IEEE Games* 494
Media Entertainment, pages 1–8. IEEE, 2014. 495

[Franzoni *et al.*, 2020] Valentina Franzoni, Giulio Biondi, 496
 and Alfredo Milani. Emotional sounds of crowds: 497
 spectrogram-based analysis using deep learning. *Multi-* 498
media tools and applications, 79(47):36063–36075, 2020. 499

[Fujishima, 1999] T. Fujishima. Chroma feature analysis and 500
 synthesis. In *Proceedings of the International Computer* 501
Music Conference, 1999. 502

[Funk and James, 2001] D.C. Funk and J.D. James. The psy- 503
 chological continuum model: A conceptual framework for 504
 understanding an individual’s psychological connection to 505
 sport. *Sport Management Review*, 4:119–150, 2001. 506

[Funk and James, 2004] Daniel C. Funk and Jeffrey D. 507
 James. The fan attitude network (FAN) model: Explor- 508
 ing attitude formation and change among sport consumers. 509
Sport Management Review, 7(1):1–26, May 2004. 510

[Germano, 1992] Massimo Germano. Turbulence: the filter- 511
 ing approach. *Journal of Fluid Mechanics*, 238:325–336, 512
 1992. 513

[Ghaemmaghani *et al.*, 1997] S Ghaemmaghani, Mohamed 514
 Deriche, and Boualem Boashash. A new approach to pitch 515
 and voicing detection through spectrum periodicity mea- 516
 surement. 2:743–746, 1997. 517

[Goto and Rynkiewicz, 2019] Masanori Goto and Joseph 518
 Rynkiewicz. Performance indicators and team styles in 519
 DotA 2. *International Journal of Performance Analysis in* 520
Sport, 19(6):1120–1131, 2019. 521

[Gouyon *et al.*, 2000] Fabien Gouyon, François Pachet, 522
 Olivier Delerue, et al. On the use of zero-crossing rate for 523
 an application of classification of percussive sounds. In 524
Proceedings of the COST G-6 conference on Digital Au- 525
dio Effects (DAFX-00), Verona, Italy, volume 5, page 16, 526
 2000. 527

[Hamari and Sjöblom, 2017] J. Hamari and M. Sjöblom. 528
 What is esports and why do people watch it? *Internet* 529
Research, 27(2):211–232, 2017. 530

- [Hattingh *et al.*, 2024] W. Hattingh, L. Van den Berg, and A. Bevan-Dye. The “why” behind generation y amateur gamers’ ongoing esports gameplay intentions. *International Journal of Sports Marketing and Sponsorship*, 25(1):67–87, 2024.
- [Himmelstein and Shapiro, 2020] Matthew Himmelstein and Joshua Shapiro. Esports: A new era of sport with opportunities for sports scientists. *International Journal of Sports Physiology and Performance*, 15(4):485–487, 2020.
- [Jang and Byon, 2020] W.W. Jang and K.K. Byon. Antecedents and consequence associated with esports gameplay. *International Journal of Sports Marketing and Sponsorship*, 21(1):1–22, 2020.
- [Jang *et al.*, 2024] W.W. Jang, W. Choi, M.J. Kim, H. Song, and K.K. Byon. What makes esports consumers watch streamers’ esports live-streaming contents? extending the theory of planned behavior. *International Journal of Sports Marketing and Sponsorship*, 25(2):288–309, 2024.
- [Johnson, 2020] Todd P Johnson. An introduction to esports data analytics. In *E-Sports Performance and Analytics*, pages 1–21. Springer, Cham, 2020.
- [Khalil *et al.*, 2019] Ruhul Amin Khalil, Edward Jones, Mohammad Inayatullah Babar, Tariqullah Jan, Mohammad Haseeb Zafar, and Thamer Alhussain. Speech emotion recognition using deep learning techniques: A review. *IEEE access*, 7:117327–117345, 2019.
- [Klapuri, 2004] Anssi Klapuri. Signal processing methods for the automatic transcription of music. Tampere University of Technology Finland, 2004.
- [Kohavi and John, 1998] Ron Kohavi and George H John. The wrapper approach. In *Feature extraction, construction and selection*, pages 33–50. Springer, 1998.
- [Lefèvre *et al.*, 2021] Bruno Lefèvre, Rachit Agarwal, Valérie Issarny, and Vivien Mallet. Mobile crowd-sensing as a resource for contextualized urban public policies: a study using three use cases on noise and soundscape monitoring. *Cities & Health*, 5(1-2):179–197, 2021.
- [McFee *et al.*, 2015] Brian McFee, Colin Raffel, Dawen Liang, Daniel PW Ellis, Matt McVicar, Eric Battenberg, and Oriol Nieto. librosa: Audio and music signal analysis in python., 2015.
- [Meir, 2000] R. Meir. Fan reaction to the match day experience: A case study in english professional rugby league football. *Sport Marketing Quarterly*, 9(1):34–42, 2000.
- [Mortelier and Rioult, 2022] Alexis Mortelier and François Rioult. Addressing the complexity of MOBA games through simple geometric clues. In *International Conference on Computational Collective Intelligence*, pages 643–649. Springer, 2022.
- [Patel and Lindley, 2020] Tirth Patel and Samuel Lindley. Analyzing physical exertion in esports: A review. *Entertainment Computing*, 34:100387, 2020.
- [Peeters, 2003] Geoffroy Peeters. Measuring the timbre of audio signals. In *Proceedings of the ISMIR*, 2003.
- [Pizzo *et al.*, 2018] A.D. Pizzo, N. Sangwon, B.J. Baker, L. Mi Ae, K. Doohan, and D.C. Funk. esports vs. sport: A comparison of spectator motives. *Sport Marketing Quarterly*, 27(2):108–123, 2018.
- [Proksch *et al.*, 2023] Shannon Proksch, Majerle Reeves, Kent Gee, Mark Transtrum, Chris Kello, and Ramesh Balasubramaniam. Recurrence quantification analysis of crowd sound dynamics. *Cognitive Science*, 47(10):e13363, 2023.
- [Qian *et al.*, 2020] T. Y. Qian, J. J. Zhang, J. J. Wang, and J. Hulland. Beyond the game: Dimensions of esports online spectator demand. *Communication & Sport*, 8(6):825–851, 2020.
- [Ramirez *et al.*, 2004] Javier Ramirez, José C Segura, Carmen Benitez, Angel De La Torre, and Antonio Rubio. Efficient voice activity detection algorithms using long-term speech information. volume 42, pages 271–287. Elsevier, 2004.
- [Rao *et al.*, 2013] K Sreenivasa Rao, Shashidhar G Koolagudi, and Ramu Reddy Vempada. Emotion recognition from speech using global and local prosodic features. *International journal of speech technology*, 16:143–160, 2013.
- [Reimers and Goldman, 2019] Victor Reimers and Josiah Goldman. League of legends: The influence of the map on performance in esports. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 2658–2666, 2019.
- [Richards *et al.*, 2022] Jessica Richards, Daniela Spanjaard, Michelle O’Shea, and Francine Garlin. The changing carnival: reimaging and recreating the match-day experience in multi-purpose stadiums. *Journal of Sport & Tourism*, 26(3):269–284, 2022.
- [Rietz and Hallmann, 2023] J. Rietz and K. Hallmann. A systematic review on spectator behavior in esports: why do people watch? *International Journal of Sports Marketing and Sponsorship*, 24(1):38–55, 2023.
- [Schubert M., 2010] Lischke L. Schubert M. User behavior in games. *Entertainment Computing-ICEC*, 2010.
- [Schuller, 2018] B. W. Schuller. Automatic emotion recognition from speech: A review of the literature and recommendations for practical realization. *Speech Communication*, 99:102–128, 2018.
- [Schulte *et al.*, 2019] Simon Schulte, Joakim Ostlund, Anders Fagergren, Mikael Thorén, and Daniel Pargman. Exploring biometric data acquisition and analysis in esports. In *2019 IEEE Conference on Games (CoG)*, pages 1–8. IEEE, 2019.
- [Stewart *et al.*, 2003] B. Stewart, A.C.T. Smith, and M. Nicholson. Sport consumer typologies: A critical review. *Sport Marketing Quarterly*, 12(4):206–216, 2003.
- [Tomkinson and Van Den Ende, 2022] Simon Tomkinson and Benjamin Van Den Ende. ‘thank you for your compliance’: Overwatch as a disciplinary system. *Games and culture*, 17(2):198–218, 2022.

- 641 [Vahlo and Karhulahti, 2022] Jarno Vahlo and Veli-Matti
642 Karhulahti. Two Overwatch player profiles. In *Modes of*
643 *esports engagement in Overwatch*, pages 11–30. Springer,
644 2022.
- 645 [Voisin and Besombes, 2023] Nicolas Voisin and Nicolas
646 Besombes. Esport. *Publictionnaire. Dictionnaire ency-*
647 *clopédique et critique des publics*, 2023.
- 648 [Välisalo and Ruotsalainen, 2019] Tero Välisalo and Mikko
649 Ruotsalainen. "I never gave up" engagement with playable
650 characters and esports players of overwatch. In *Proceed-*
651 *ings of the 14th International Conference on the Founda-*
652 *tions of Digital Games*, pages 1–6, August 2019.
- 653 [Wang *et al.*, 2019] Peng Wang, Qing Song, Xiaohua Lin,
654 and Yichao Lu. An analysis on strategy evolution in DotA
655 2. In *Proceedings of the 10th International Conference on*
656 *Internet Multimedia Computing and Service*, pages 46–50,
657 2019.
- 658 [Wang, 2023] W. Wang. Returning for skill or popularity?
659 the demand for esports match replay. *International Jour-*
660 *nal of Sports Marketing and Sponsorship*, 24(2):295–310,
661 2023.