

Computer Vision for MOBA Analytics: A Dataset and Baseline for Visibility Analysis in Dota 2

Ricardo da Rocha Carvalho¹, Eloísa Oliveira¹, Luiz Bernardo Martins Kummer¹,
Emerson Cabrera Paraiso¹, Rayson Laroca¹

¹Pontifícia Universidade Católica do Paraná (PUCPR), Curitiba, Brasil

{ricardo.rcarvalho,oliveira.eloisa,luiz.kummer,paraiso,rayson}@ppgia.pucpr.br

Abstract. Introduction: Most Multiplayer Online Battle Arena (MOBA) analytics studies rely on structured data, which does not directly capture what each team could actually see during a match. **Objective:** This work introduces Dota2-Vis, a video-based dataset, and a baseline pipeline for visibility analysis in professional Dota 2 matches. **Methodology:** The dataset comprises all 144 matches from The International 2025, recorded from both team perspectives, totaling 288 Full HD videos, together with 2,477 manually annotated minimap images. We evaluate multiple variants of a modern object detector for player-icon detection and use the best-performing model to estimate opponent-visible player presence over time. **Results:** YOLO11l (large) achieved the best overall performance, reliably identifying player icons even in dense and visually cluttered minimap scenes. The resulting visibility curves reveal player, hero, role, and team-level patterns that complement conventional MOBA analytics, highlighting behavioral differences that are difficult to obtain from structured data alone. The Dota2-Vis dataset is publicly available at <https://github.com/RicardoRCarvalho/dota2-vis/>.

Keywords Computer Vision, Dota 2, Esports, Game Analytics.

1. Introduction

Game analytics has become an important tool for understanding player behavior, supporting decision-making, and extracting actionable knowledge from digital games [Seif El-Nasr et al. 2013]. This is particularly relevant in esports, where matches generate large volumes of behavioral, spatial, and temporal information that can be used to characterize strategies, compare players, and support predictive models. Among esports genres, Multiplayer Online Battle Arenas (MOBAs) are especially relevant for artificial intelligence research because they combine long-term planning, team coordination, real-time decision-making, and imperfect information [Katona et al. 2019, Costa et al. 2024].

Most prior work on MOBA analytics has relied primarily on structured data from drafts, event logs, replay files, and public Application Programming Interfaces (APIs), supporting tasks such as draft recommendation, match outcome prediction, team composition analysis, event prediction, and live win prediction [Summerville et al. 2016, Semenov et al. 2017, Hodge et al. 2021, Yang et al. 2022, Yang et al. 2023]. These sources remain central, but capture only part of match dynamics. Visual information can complement them by revealing what players can actually observe during a match. In Dota 2, for instance, player coordinates indicate where a hero is located, but not necessarily what the opposing team can see. Opponent visibility depends on fog of war,

terrain elevation, trees, wards, unit vision ranges, temporary abilities, and other map-dependent factors, making it difficult to infer accurately from tabular data alone.

Computer vision offers a complementary way to study these dynamics directly from gameplay footage, using the visual information presented to players during the match. This perspective is especially useful for analyzing visibility cues that depend on spatial context, transient game elements, and patch-dependent map changes. Recent studies have begun to explore visual information in MOBAs, including player camera analysis in Dota 2 [Tot et al. 2021] and player trajectory extraction in *League of Legends* videos [Kim et al. 2025]. Nevertheless, to the best of our knowledge, no publicly available dataset currently provides the necessary annotations or structure to systematically study visibility-related information in professional MOBA matches.

In this work, we present Dota2-Vis, a video-based dataset designed to support visibility analysis in professional Dota 2 matches. It brings together dual-perspective recordings of all matches from *The International 2025* and manually annotated minimap images collected from professional matches. The annotations include player icons, clones, and other relevant minimap elements, providing fine-grained data for training object detectors under visually dense MOBA conditions. We then explore this dataset to build a baseline pipeline that detects player icons on the minimap and estimates when each player is visible from the opposing team’s perspective throughout a match.

By aggregating detections over time, we obtain temporal descriptors of opponent-visible map presence for players, heroes, roles, and teams. These descriptors can reveal stylistic differences among players using the same hero and role, hero and role-dependent activation windows, and team-level visibility patterns associated with match outcomes.

The main contributions of this paper are threefold:

- We introduce Dota2-Vis, a dataset composed of professional Dota 2 match videos from both teams’ perspectives and manually annotated minimap images for player/icon detection;
- We provide baseline object detection experiments for detecting player icons and related minimap elements in visually dense MOBA scenes;
- We provide initial visibility-based analyses that characterize opponent-visible map presence across players, roles, heroes, teams, match stages, and match outcomes. Although not exhaustive, these analyses illustrate how visibility descriptors extracted from video can be leveraged in practice for post-match review, scouting, opponent preparation, and data-driven strategic analysis.

The remainder of this paper is organized as follows. Section 2 presents the main Dota 2 characteristics relevant to this study. Section 3 reviews related work on artificial intelligence, game analytics, and computer vision in MOBAs. Section 4 describes the dataset, annotations, and detection baseline. Section 5 presents the visibility-based analyses. Finally, Section 6 concludes the paper and outlines future work.

2. Dota 2

Defense of the Ancients 2 (Dota 2) is a MOBA in which two teams, Dire and Radiant, compete on a shared map. Each team has five players, and each player controls a unique hero with specific abilities, attributes, and strategic roles. Competitive lineups are

commonly organized into five roles: the hard carry, who receives resources to become a late-game damage threat; the mid laner, who gains early levels and helps control the pace of the early game; the offlaner, who contests space and initiates fights; and two supports, often referred to as position 4 and position 5, who provide utility, map control, and vision. The main objective is to destroy the opposing team's Ancient, a central structure located inside its base. Professional matches usually last between 30 and 40 minutes [Katona et al. 2019] and involve different phases, including laning, resource acquisition, objective control, team fights, and base sieges.

A central characteristic of Dota 2 is imperfect information. Teams only observe regions revealed by allied heroes, units, buildings, wards, and other temporary sources of vision, while the remaining areas are hidden by fog of war. Thus, visibility is a strategic resource: knowing where opponents are, inferring where they might be, and denying information to the opposing team directly affects movement, objective control, and team-fight preparation [Chitayat et al. 2020, Tot et al. 2021].

These mechanics make visibility a state-dependent signal rather than a simple function of hero location. As illustrated in Figure 1, nearby heroes are not necessarily mutually aware, since line of sight may be blocked or enabled by local map elements such as trees, terrain transitions such as stairs, and temporary vision sources.

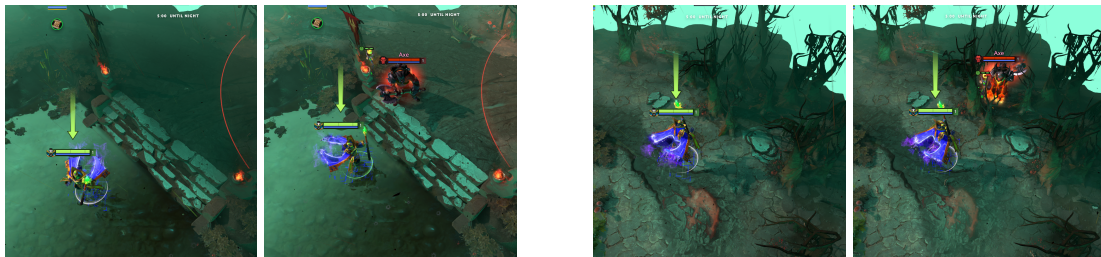


Figure 1. Representative examples illustrating why visibility in Dota 2 cannot be inferred from player coordinates alone. Within each pair, the left image shows the information available from a player or team perspective, while the right image shows the complete game state, highlighting the effect of terrain elevation, trees, and local map geometry.

3. Related Work

Prior work on MOBA analytics has largely relied on structured match data, such as drafts, event logs, replay files, and public APIs. In DotA 2, this line of research includes draft analysis and hero recommendation [Summerville et al. 2016], outcome prediction from team composition [Semenov et al. 2017], and live win prediction in professional matches [Hodge et al. 2021]. Similar predictive tasks have also been explored in other MOBA, including interpretable win prediction for *Honor of Kings* [Yang et al. 2022] and event prediction in MOBA matches [Yang et al. 2023]. Together, these studies show the relevance of machine learning for esports analysis, while also illustrating the common dependence on information already available through logs, replays, or telemetry.

Other studies have investigated more specific aspects of match dynamics. [Katona et al. 2019] and [Ringer et al. 2023], for example, addressed in-game death prediction in DotA, while [Chitayat et al. 2020] modeled the value of vision through

warding information in *League of Legends* and Dota 2. The latter is particularly related to our work, as it highlights the strategic importance of visibility in MOBAs. However, it estimates the value of vision from available game data, whereas we aim to recover visibility information directly from gameplay footage.

Computer vision remains less explored than structured-data approaches in MOBA research. [Tot et al. 2021] modeled player intent in Dota 2 by combining hero and player-camera positions to predict team fights shortly before they occur, while [Kim et al. 2025] proposed an object-detection-based framework (adapting YOLOv5) to extract player trajectories from *League of Legends* videos. Beyond MOBAs, [Joo et al. 2025] studied hierarchical scene detection and tracking in *StarCraft* esports videos, aiming to improve automatic observing and spectator experience. These studies suggest that visual information can complement traditional game data, especially when relevant events or states are unavailable or difficult to recover from structured form.

Our work follows this direction, but focuses on a different aspect of Dota 2: player visibility. By detecting player positions on the minimap from videos recorded from both teams' perspectives, we estimate when players are visible to opponents and analyze how this information evolves throughout professional matches. To our knowledge, no publicly available MOBA dataset combines video data, dual-team perspectives, and minimap annotations in a way that supports this type of opponent-visibility analysis. Thus, the proposed dataset and baseline experiments provide an initial step toward using computer vision to study visibility patterns that are difficult to obtain from tabular data alone.

4. The Dota2-Vis Dataset: Data Collection and Player Detection

We introduce *Dota2-Vis*, a dataset designed to support visibility analysis in professional Dota 2 matches. The dataset has two complementary components: (i) full match videos recorded from both teams' perspectives, used for visibility analysis; and (ii) manually annotated minimap images, used to train and evaluate the player detector. Both components, along with their annotations and predefined data splits, are publicly available¹.

To build the video component, we replayed professional matches using the official game engine and recorded them with OBS Studio, keeping the screen, game, and recording resolutions fixed at 1920×1080 pixels. As our goal is to analyze visibility at the highest level of play, we recorded all matches from *The International 2025*, the main annual tournament in the professional Dota 2 scene [Liquipedia 2026]. The tournament comprised 144 matches, each recorded twice, once from the Radiant perspective and once from the Dire perspective, resulting in 288 Full HD videos.

To train the player detector, we collected 2,477 minimap images from professional matches that were not part of *The International 2025*, preventing overlap between the training data and the matches analyzed later in this work [Barz e Denzler 2020]. We focus on the minimap rather than the full gameplay screen because it provides a compact, standardized representation of what each team can observe at any given moment. This choice reduces the influence of camera movement, zoom, combat effects, and hero animations, while preserving the information required for our target task: estimating whether players are visible to the opposing team.

¹<https://github.com/RicardoRCarvalho/dota2-vis/>

The annotation protocol was designed to account for both regular players and game-specific visual elements that may interfere with visibility estimation. We initially considered 10 player classes, one for each player in the match. However, some heroes, items, and runes can create controllable *clones*, i.e., additional units that may appear on the minimap similarly to the original hero. As clones may affect visibility analysis, we added one clone class for each player, resulting in 20 player-related classes. In addition, players can communicate with their teammates through *pings*, which are temporary minimap markers used, for example, to indicate danger, movement, or attack intentions. As pings share the same color as the player who created them and may be visually confused with player icons, we added an *Other* class to represent non-player elements that could affect detection. Thus, the final annotation protocol contains 21 classes. To avoid making the detector dependent on specific heroes, we configured the game to display players on the minimap using color and symbol icons instead of hero portraits. Notably, this makes the detector independent of the current hero pool and avoids the need for retraining whenever new heroes are added to the game.

All minimap images have resolution 240×240 pixels and were manually annotated with bounding boxes (x, y, w, h) . This process is time-consuming because minimaps often contain many small, overlapping objects amid visual clutter and similar colors. As shown in Figure 2, the samples cover different match moments, ranging from sparse situations to dense scenes with several visible players and non-player elements, including challenging cases with clones, pings, and cluttered regions. The rightmost column shows the corresponding bounding boxes, illustrating the level of detail required during annotation. To improve reliability, the annotations were reviewed in multiple rounds using different data splits, with the remaining inconsistencies observed in preliminary experiments mostly restricted to difficult cases involving severe occlusion or visual ambiguity.

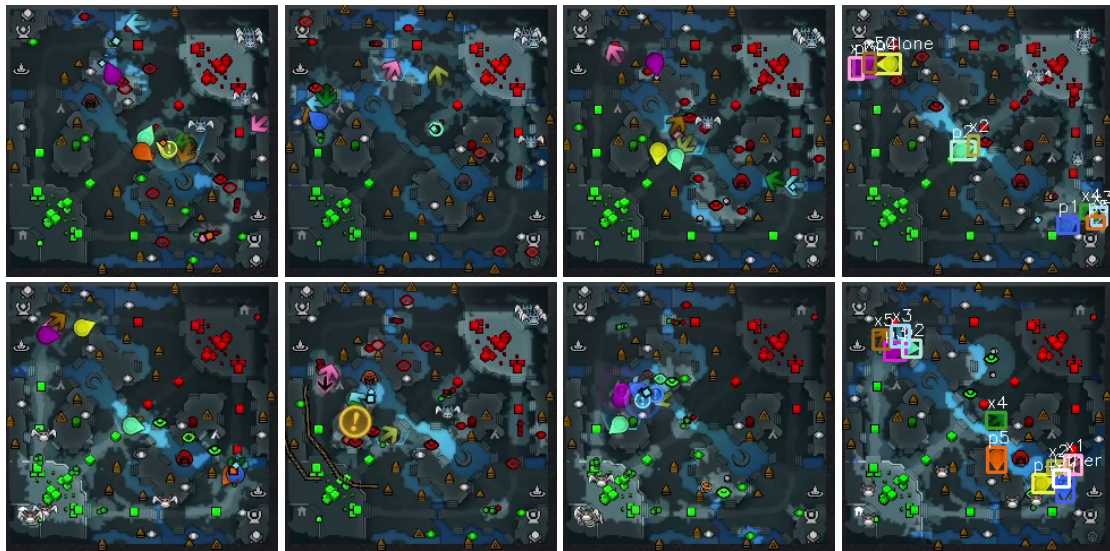


Figure 2. Representative examples from the proposed Dota2-Vis dataset, illustrating the variability of minimap samples in terms of team side, game moment, object density, and visual clutter. The rightmost column shows manually labeled bounding boxes for reference.

The 2,477 images were split into training, validation, and test subsets, with the test

set containing only frames from matches not used for training or validation. This reduces visual overlap between training and evaluation data. The final split contains 1,670 training, 407 validation, and 400 test images, following an approximately 70/15/15 distribution.

We adopted YOLOv11 [Ultralytics 2024] in the detection experiments, motivated by the strong performance of the YOLO family across diverse object detection tasks [Terven et al. 2023, Kim et al. 2025, Laroca et al. 2025]. YOLOv11 provides five variants, from YOLO11n (nano) to YOLO11x (extra-large), which differ in size, complexity, and speed-accuracy trade-off. Considering that the most suitable variant depends on the target application, we evaluate all five YOLO11 variants in this work.

The hyperparameters were selected based on preliminary experiments conducted on the validation set. The final training configuration used 200 epochs with early stopping patience of 15, image size of 480×480 pixels, a maximum of 40 detections per image, and dropout of 0.1. We also applied data augmentation, including small changes in saturation and brightness, vertical flips, mosaic augmentation, scale variation, erasing, and small geometric transformations, such as rotation, shear, and perspective changes.

For evaluation, we report precision, recall, F-score, and $\text{mAP}_{50:95}$. A detection is considered correct when its predicted class matches the ground truth and the overlap between the predicted and annotated bounding boxes satisfies Intersection over Union ($\text{IoU} > 0.5$), following the standard protocol adopted in object detection benchmarks [Everingham et al. 2010]. We place particular emphasis on F-score and $\text{mAP}_{50:95}$. The F-score is the harmonic mean of precision and recall, capturing the balance between false detections and missed players, both of which directly affect visibility estimation. In turn, $\text{mAP}_{50:95}$ evaluates detection quality across multiple IoU thresholds, providing a stricter assessment of localization performance [Lin et al. 2014].

As shown in Table 1, the YOLO11l (large) variant achieved the best overall results, obtaining the highest F-score and $\text{mAP}_{50:95}$ among the evaluated variants. Although YOLO11x is substantially larger than YOLO11l (194.9 BFLOPs vs. 86.9 BFLOPs), it did not improve performance in this setting. One possible explanation is that the task involves small and visually simple objects, while the available training data may not be sufficient for the additional model capacity to translate into better generalization. Thus, we selected YOLO11l as the model used in the visibility analyses reported in this work.

Table 1. Results obtained by the different YOLOv11 variants. The *large* variant provided the best trade-off between computational cost and accuracy.

Model	Precision	Recall	F-score	$\text{mAP}_{50:95}$
YOLO11n (nano)	0.964	0.881	0.921	0.661
YOLO11s (small)	0.965	0.896	0.929	0.713
YOLO11m (medium)	0.963	0.903	0.932	0.721
YOLO11l (large)	0.974	0.906	0.939	0.729
YOLO11x (extra-large)	0.965	0.911	0.937	0.720

Figure 3 shows representative detections from the selected model. Overall, the detector identifies players reliably, even in dense minimap scenes and visually cluttered situations. Most errors occur in rare cases of severe occlusion or strong overlap between small objects, as illustrated by the missed detection highlighted in red in the figure.



Figure 3. Representative detections produced by YOLO11l (large). Green bounding boxes denote correct detections, whereas the red bounding box in the leftmost image highlights a missed detection of a heavily occluded player icon. Overall, the model detects players reliably, with failures occurring mainly under severe occlusion.

Importantly, as an additional target-domain sanity check, we manually annotated 400 held-out minimap frames from *The International 2025* videos, sampling two to three random frames per game. The detector achieved performance consistent with Table 1, with an F-score of 0.940 and an $\text{mAP}_{50:95}$ of 0.729.

All models were implemented using the Ultralytics framework [Ultralytics 2024]. The experiments were run on a computer equipped with an NVIDIA RTX 4090 GPU with 24GB of memory.

5. Visibility-Based Game Analytics

The previous section established the feasibility of detecting player icons on the Dota 2 minimap. We now shift the focus from detector performance to the type of game analytics enabled by the proposed pipeline. Rather than using visual data only as an auxiliary representation, our goal is to recover a match signal that is difficult to obtain from structured data alone: which players were visible to the opposing team over time. This information complements traditional MOBA analytics based on drafts, events, public APIs, or match statistics [Seif El-Nasr et al. 2013, Costa et al. 2024], and is especially relevant because visibility is directly related to map control, pressure, and uncertainty in decision-making [Chitayat et al. 2020].

For each match, we process the videos recorded from both team perspectives. A player is considered visible when their minimap icon is detected in the video corresponding to the opponent’s perspective. We aggregate these detections in five-minute intervals and compute the visibility percentage as:

$$\text{Vis}(\mathcal{E}, I) = 100 \cdot \frac{1}{|\mathcal{F}_I| |\mathcal{E}|} \sum_{f \in \mathcal{F}_I} \sum_{e \in \mathcal{E}} \mathbf{1}(e \in \mathcal{D}_f), \quad (1)$$

where $\mathcal{F}_I$ is the set of frames in interval I , $\mathcal{E}$ is the entity or group of entities being analyzed, and $\mathcal{D}_f$ is the set of detections in frame f . Depending on the choice of $\mathcal{E}$, this metric can represent the visibility of a player, a hero, a role, or an entire team. In the following analyses, we report curves up to the 45-minute mark to reduce the influence of late-game intervals represented by fewer matches.

Figure 4a illustrates a player-level analysis in which the hero and role are fixed. We selected *Earthshaker*, the most picked mid-lane hero in the tournament, and compared the

five players with the largest number of matches using this hero, with each player appearing in at least three matches. This setting reduces part of the variability caused by hero choice and allows the visibility curves to highlight differences in player behavior. For example, *4nalog*, one of the analyzed players, presents higher early-game visibility than the others, suggesting a more active or pressure-oriented use of the hero during the first minutes. In contrast, player *Malr1ne* remains less visible in the same period, which may indicate a more conservative approach, safer resource acquisition, or a strategy based on keeping the opponent uncertain about his position. These curves should not be interpreted as direct measures of player quality. Instead, they provide a quantitative way to describe stylistic differences usually discussed qualitatively by analysts, coaches, and players.

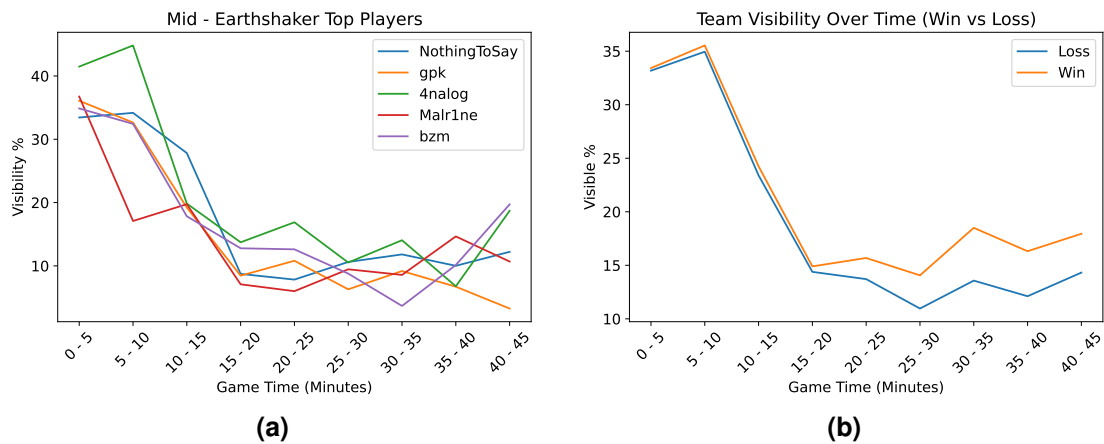


Figure 4. Examples of visibility-based analyses enabled by the proposed pipeline. (a) Player-level visibility curves for the mid-lane *Earthshaker* players with the largest number of matches in the tournament. (b) Team-level visibility curves grouped by match outcome. Curves are computed in five-minute intervals from videos of *The International 2025*.

At the team level, Figure 4b shows that winning and losing teams have similar visibility during the early game, especially during the laning stage. After approximately 20 minutes, however, the curves start to separate, with winning teams becoming more visible on the map. A plausible interpretation is that, once a team gains resources and map control, it can occupy more contested areas, pressure objectives, and force the opposing team to play from safer or hidden positions. This behavior is consistent with the snowballing dynamics commonly observed in MOBAs, in which an initial advantage is progressively converted into more control over space and resources [Li et al. 2017]. From a modeling perspective, this suggests that opponent-visible map presence could be used as an additional feature for live win prediction or event prediction, complementing structured indicators such as gold, experience, kills, objectives, and draft information [Semenov et al. 2017, Hodge et al. 2021, Yang et al. 2022, Yang et al. 2023].

The same representation can also be used to compare heroes and roles. Figure 5 shows visibility curves for the six most picked heroes in the mid lane and offlane roles. In the mid lane (Figure 5a), heroes exhibit different activation patterns: some are more visible in the early game, while others become more present only after acquiring levels, items, or mobility. For instance, the curve of *Storm Spirit* suggests a later increase in map activity, which is consistent with a hero that often requires resources before

becoming a constant threat across the map. In the offlane (Figure 5b), visibility tends to reflect the role’s responsibility of contesting space, initiating fights, and disrupting the enemy carry. Heroes such as *Mars* and *Bristleback* illustrate different timings of this pressure, with visibility changing as their relative strength evolves throughout the match. These observations should not be interpreted as patch-invariant conclusions about specific heroes. Rather, they show that the proposed visual signal can transform strategic concepts such as activity, pressure, and hiddenness into measurable temporal patterns.

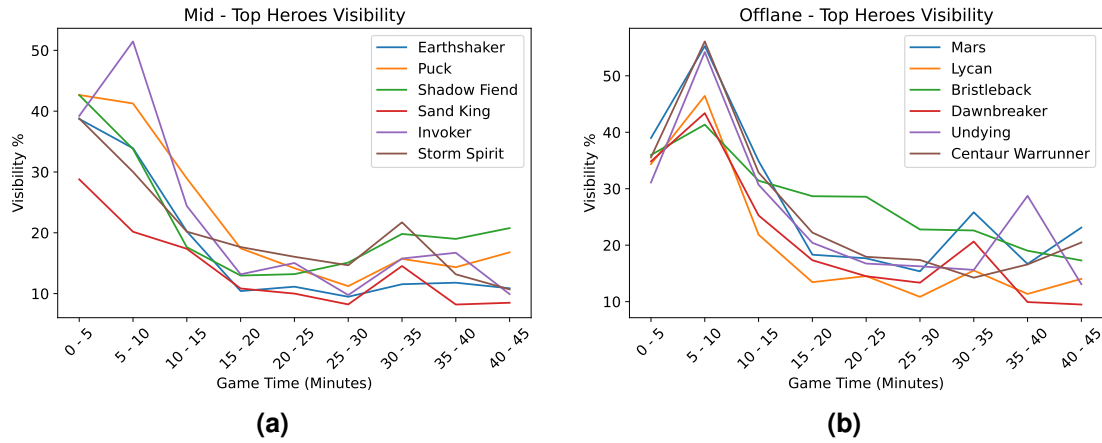


Figure 5. Hero-level visibility curves for the six most picked heroes in two roles at *The International 2025*. (a) Mid lane. (b) Offlane. The curves illustrate how the proposed visual signal can reveal hero-specific activation windows and role-dependent map presence patterns.

Beyond aggregate analysis, the recovered visibility signal can support practical forms of performance review and training. For players, it can be used to compare map presence against reference profiles from high-level players using the same hero and role, helping identify moments of unusual exposure, excessive passivity, or different activity timings. For teams, it can guide post-match review by highlighting intervals in which visible map presence did not translate into objectives, pressure, or resource denial, as well as moments in which low visibility may indicate missed opportunities to contest space. In opponent preparation, player-hero visibility profiles may also reveal recurring timing patterns, such as when a mid player tends to rotate, when an offlaner starts occupying dangerous areas, or when a team begins showing itself before major objectives. These uses should be understood as diagnostic support rather than automatic coaching prescriptions: visibility alone does not explain intent, but it can guide replay review and provide objective evidence for discussions about map control, risk, and timing [Novak et al. 2019, Bialecki et al. 2024].

These examples highlight why the proposed Dota2-Vis dataset goes beyond a minimap detection benchmark. By extracting visibility directly from video, it provides behavioral descriptors that are difficult to recover from structured data alone. Player coordinates, for instance, do not necessarily indicate what the opposing team could actually see. As discussed in Section 2 and illustrated in Figure 1, visibility depends on occlusions and map geometry, including trees that may block vision and elevation changes that affect line of sight. However, some trees can be destroyed during a match, and map layouts may change across game patches. Hence, even a hard-coded

visibility model would require constant updates and could still diverge from what was actually displayed to players during the match. Our video-based approach therefore provides a complementary path for MOBA analytics, connecting recent efforts on visual understanding of esports footage [Tot et al. 2021, Kim et al. 2025, Joo et al. 2025] with strategic analysis in professional Dota 2.

6. Conclusions

This paper presented Dota2-Vis, a dataset for studying visibility in professional Dota 2 matches through computer vision. The dataset comprises dual-perspective gameplay videos from *The International 2025* and manually annotated minimap images, enabling the training and evaluation of methods that extract visibility-related information from gameplay footage. To our knowledge, Dota2-Vis is the first dataset to support this type of opponent-visibility analysis in professional Dota 2 matches from video.

We also introduced a baseline pipeline for player-icon detection on the minimap and for estimating when each player is visible to the opposing team. This pipeline converts a visual, context-dependent game mechanic into quantitative temporal descriptors. The results show that this formulation is feasible and provides a useful starting point for visibility-centered analysis, with YOLO11l (large) achieving the best overall trade-off among the evaluated detectors. Initial analyses of the extracted visibility curves reveal patterns that are difficult to obtain from structured data alone, including differences among players using the same hero and role, hero- and role-dependent activation windows, and changes in visibility trends between winning and losing teams after the early game. These preliminary findings suggest that opponent-visible map presence may provide a useful signal for studying activity, pressure, risk, and map control in Dota 2.

The analyses reported in this work should be interpreted as an initial demonstration of what Dota2-Vis enables, rather than as a complete characterization of professional Dota 2 visibility patterns. Given the scope of this work, we focus on representative cases that illustrate how the proposed visual signal can support analyses of players, heroes, roles, teams, match stages, and outcomes. Broader analyses across different matchups and contexts remain an important direction for future work. Moreover, as the dataset is centered on a single tournament, some findings may reflect the specific teams, patch, metagame, and broadcast conditions of *The International 2025*.

Future work will expand the dataset to additional tournaments and patches, combine visibility with structured data from public APIs, and incorporate temporal tracking and uncertainty estimation. These extensions would support more detailed analyses of team-level visibility profiles, objective-contest scenarios, smoke plays, ganks, warding patterns, and match-stage differences. More broadly, the same video-based formulation could be adapted to other MOBAs, such as *League of Legends* and *Honor of Kings*, further extending the use of computer vision for visibility-centered game analytics.

Acknowledgments

This study was financed in part by the *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil* (CAPES), Finance Code 001.

References

- Barz, B. e Denzler, J. (2020). Do we train on test data? Purging CIFAR of near-duplicates. *Journal of Imaging*, 6(6):41.
- Białecki, A., Michalak, B., e Gajewski, J. (2024). Esports training, periodization, and software—a scoping review. *Applied Sciences*, 14(22):10354.
- Chitayat, A. P., Kokkinakis, A., Patra, S., Demediuk, S., Robertson, J., Olarewaju, O., Ursu, M., Kirmann, B., Hook, J., Block, F., e Drachen, A. (2020). WARDS: Modelling the worth of vision in MOBA's. In *Science and Information Conference*, pages 63–81.
- Costa, L. M., Drachen, A., Souza, F. C. M., e Xexéo, G. (2024). Artificial intelligence in MOBA games: A multivocal literature mapping. *IEEE Transactions on Games*, 16(2):250–269.
- Everingham, M., Van Gool, L., Williams, C. K. I., Winn, J., e Zisserman, A. (2010). The pascal visual object classes (VOC) challenge. *International Journal of Computer Vision*, 88(2):303–338.
- Hodge, V. J., Devlin, S., Sephton, N., Block, F., Cowling, P. I., e Drachen, A. (2021). Win prediction in multiplayer esports: Live professional match prediction. *IEEE Transactions on Games*, 13(4):368–379.
- Joo, H.-T., Lee, S.-H., Chung, I., e Kim, K.-J. (2025). Improving esports viewing experience through hierarchical scene detection and tracking. *Scientific Reports*, 15(1):9803.
- Katona, A., Spick, R., Hodge, V. J., Demediuk, S., Block, F., Drachen, A., e Walker, J. A. (2019). Time to die: Death prediction in Dota 2 using deep learning. In *IEEE Conference on Games (CoG)*, pages 1–8.
- Kim, D.-W., Park, S.-Y., Yang, S.-I., e Lee, S.-K. (2025). Real-time player tracking framework on MOBA game video through object detection. *IEEE Transactions on Games*, 17(2):498–509.
- Laroca, R., dos Santos, M., e Menotti, D. (2025). Improving small drone detection through multi-scale processing and data augmentation. In *International Joint Conference on Neural Networks (IJCNN)*, pages 1–8.
- Li, Q., Xu, P., Chan, Y. Y., Wang, Y., Wang, Z., Qu, H., e Ma, X. (2017). A visual analytics approach for understanding reasons behind snowballing and comeback in MOBA games. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):211–220.
- Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., e Zitnick, C. L. (2014). Microsoft COCO: Common objects in context. In *European Conference on Computer Vision (ECCV)*, pages 740–755.
- Liquipedia (2026). The International. https://liquipedia.net/dota2/The_International. Accessed: 2026-04-18.
- Novak, A. R., Bennett, K. J. M., Pluss, M. A., e Fransen, J. (2019). Performance analysis in esports: Part 1 – the validity and reliability of match statistics and notational analysis in League of Legends. *SportRxiv*.

- Ringer, C., Missaoui, S., Hodge, V. J., Chitayat, A. P., Kokkinakis, A., Patra, S., Demediuk, S., Munoz, A. C., Olarewaju, O., Ursu, M., Kirman, B., Hook, J., Block, F., Drachen, A., e Walker, J. A. (2023). Time to die 2: Improved in-game death prediction in Dota 2. *Machine Learning with Applications*, 12:100466.
- Seif El-Nasr, M., Drachen, A., e Canossa, A., editors (2013). *Game Analytics: Maximizing the Value of Player Data*. Springer, London.
- Semenov, A., Romov, P., Korolev, S., Yashkov, D., e Neklyudov, K. (2017). Performance of machine learning algorithms in predicting game outcome from drafts in Dota 2. In *International Conf. on Analysis of Images, Social Networks and Texts*, pages 26–37.
- Summerville, A., Cook, M., e Steenhuisen, B. (2016). Draft-analysis of the ancients: predicting draft picks in Dota 2 using machine learning. In *AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, pages 100–106.
- Terven, J., Córdova-Esparza, D.-M., e Romero-González, J.-A. (2023). A comprehensive review of YOLO architectures in computer vision: From YOLOv1 to YOLOv8 and YOLO-NAS. *Machine Learning and Knowledge Extraction*, 5(4):1680–1716.
- Tot, M., Conserva, M., Chitayat, A. P., Kokkinakis, A., Patra, S., Demediuk, S., Munoz, A. C., Olarewaju, O., Ursu, M., Kirmann, B., Hook, J., Block, F., Drachen, A., e Perez-Liebana, D. (2021). What are you looking at? Team fight prediction through player camera. In *IEEE Conference on Games (CoG)*, pages 1–8.
- Ultralytics (2024). YOLOv11. <https://docs.ultralytics.com/models/yolo11/>. Accessed: 2026-04-26.
- Yang, Z., Pan, Z., Wang, Y., Cai, D., Shi, S., Huang, S.-L., Bi, W., e Liu, X. (2022). Interpretable real-time win prediction for Honor of Kings—a popular mobile MOBA esports. *IEEE Transactions on Games*, 14(4):589–597.
- Yang, Z., Wang, Y., Li, P., Lin, S., Shi, S., Huang, S.-L., e Bi, W. (2023). Predicting events in MOBA games: Prediction, attribution, and evaluation. *IEEE Transactions on Games*, 15(2):193–201.