Bicycle crash motion using instance segmentation

Introduction

Bicycle usage carries several benefits, from less traffic congestion in cities, less polution, healthier lifestyle, among others. However, cyclist are vulnerable road users, and crash rates are increasing [European Comission 2024][1]. For this reason, research in the field of bicycle crashes during the last 15 years has gained popularity. Within the types of bicycle crashes, one that stands out is the single-cyclist crash, where no other road user is involved, and its share ranges from 50 to 85% of bicycle-related hospital admissions [Utriainen et al. 2023][3].

From these situations, we find in the literature three known bicycle crash motions, which are 'pitch-over', 'roll-over' and 'skid-out'. Pitch-over is when, due to excessive longitudinal load transfer, the rear wheel lifts from the ground and the bicycle rotates around the contact point of the front wheel. Similarly, roll-over refers to the crash due to an excesive roll rate and the cyclist falling to the side. Lastly, skid-out occurs when the grip in one of the tyres suddenly decrease and the bicycle falls beneath the cyclist [Gildea et al. 2021][7].

The lack of data of these events is commonly discussed in the literature [Utriainen 2020][2]. However, no solutions have been provided, and recreating crashes lead to dangerous situations for test subjects even in laboratory controlled setups. For this reason, we propose a solution to gather single-cyclist crash data based on monocular video analysis. In order to automatise this analysis, we delve into computer vision techniques. Among the common computer vision tasks, we find object detection, semantic segmentation and instance segmentation. Object detection consists in identifying an object of certain class in a digital image [Zou et al. 2023][6]. Semantic segmentation deals with demarcating different objects and parts of an unknown image [Guo et al. 2018][5]. Then, instance segmentation is the task that combine both previously mentioned goals simultaneously, identifying the class of the object and segmentating it in the image [Hafiz, A.M., Bhat, G.M.][4].

Although in the literature we find bicycle crash analysis using computer vision methods [Gildea et al. 2024][8], it was focused on the rider fall outcome, not in the motion of the bicycle and its particular dynamics. Therefore, the aim of the present study is to gather and identify bicycle crash motion data from monocular videos. To this end, we created a dataset of single-cyclist crashes from web sources. From these videos, we track the motion of the bicycle by identifying and tracking the position of the wheeels.

Methods

To create a database of single-cyclist crash videos, we manually selected videos from public domain sources, such as YouTube. Selected videos contain bicycle crashes where no other road user is involved, or if there is another user, its participation in the event is assumed as a passive perturbation, i.e. the excerted force is uncontrolled and it does not depend on the source. The dataset consists of 100 videos with an average duration of 3 seconds and their time stamp on the original source.

The dataset was annotated using **C**omputer **V**ision **A**nnotation **T**ool (CVAT) for instance segmentation with ellipse shapes. Thus, using this approach, we label the wheels as ellipses using 'front' and 'rear' as features. These anotations give the boundary points of the ellipses, from which we obtain the centres of the wheels and their orientation by estimating the rotation angle with respect to the vertical axis. In addition, we calculate the ratio between the minor and major axes of the ellipses, which give us the rotation angle around the vertical axis, also understood as the angle with respect to the camera view. At last, we calculate the distance in x and y axes between the centres of both ellipses. The measurement unit for this process are the pixels of the image.

Results

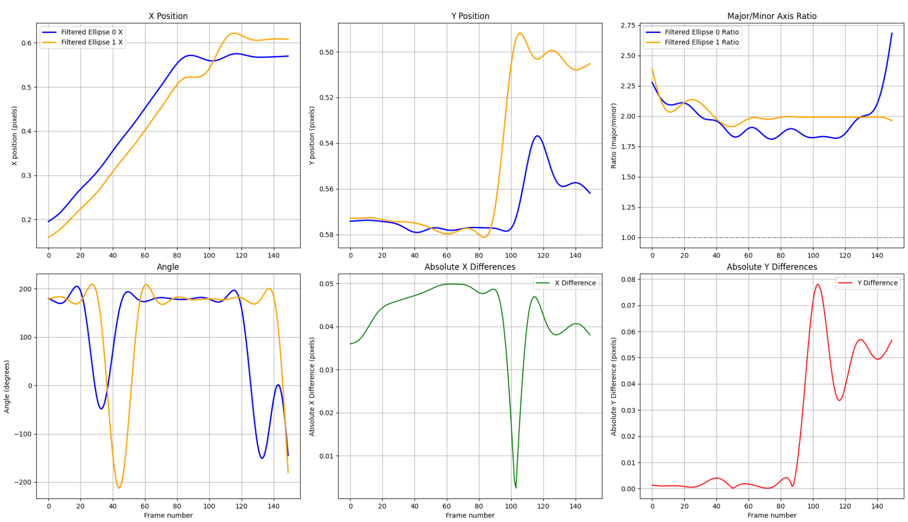

Figure 1: a) x position of the centre of the ellipses in time. b) y position of the centre of the ellipses in time. c) Ratio between the major and minor axes. d) Angle of the ellipses with respect to the vertical axis. e) Distance in pixels between both wheel centres in x axis. f) Distance in pixels between both wheel centres in y axis. Note that position plots are normalised.

Figure 1 shows the data of a 'pitch-over' crash.

We observe that after frame 100, the rear wheel moves forward over the front wheel. Similarly, between frame 85 and 105, there is a major increase in the vertical position of the rear wheel. Taking into account the average distance in the y axis between both wheel centres, measured in pixels of the image, a peak

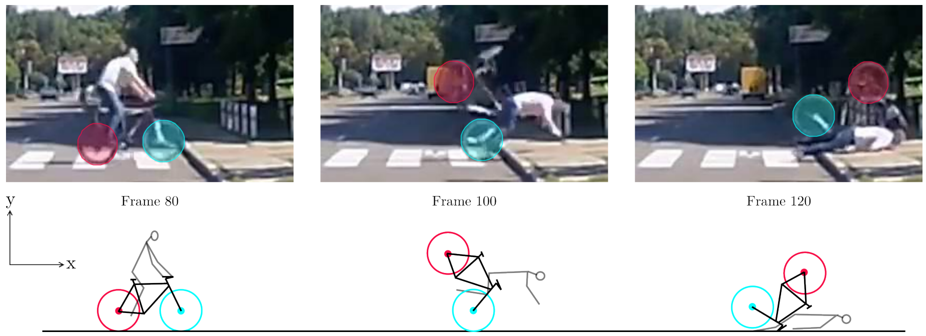

Figure 2: Annotated frames of the crash video.

is observed during the crash motion, which is equivalent to 283% of the average difference.

Following the same methodology, we extract data from more videos of the dataset. The main findings are summarised in the table below.

Table 1: Distance between wheels from crash videos.

Number	Crash Type	Avg x diff.	Avg y diff.	Max x %	Max y %
1	Pitch-over	-0.08	0.08	-93.2	178.5
3	Start riding	-0.07	0.01	-10.3	226.1
5	Pitch-over	0.02	0.02	167.7	282.7
6	Pitch-over	-0.1	0.03	-93.7	577.7
7	Pitch-over	0	0.01	-844	1203.2
8	Lost control	0.01	0.06	1142	196
9	Pitch-over	-0.09	0.04	-84	141.6
10	Disengagement	-0.04	-0.01	-57.7	-309.9
11	Pitch-over	0.01	0.01	412.9	1286.6

Table 2: Data of no-crash videos.

Number	Avg x diff.	Avg y diff.	Max x %	Max y %
1	0	-0.01	559.4	-88.5
3	-0.3	0	-15.4	1832.4
4	-0.38	0	-79.5	-1078.6

Discussion

Problems addressed in this research are bicycle crash detection and the lack of single-cyclist real-world crash data. Using instance segmentation, we were able to track the position of the wheels of the bicycle in the image. From this data, we can analyse crash motions…

We hypothesize that sudden variations in the distance between wheel centres will allow to identify crashes with large vertical rotations of the bicycle, such as pitch-over crashes.

We observe that crash scenarios present higher average variations in the motion of the wheels. Furthermore, pitch-over crash configurations always show variations over 150% with respect to the average distance. Similarly, the average difference for both directions (x and y) is greater than 0.01 pixels. Contrarily, no-crash scenarios tend to show average differences below 0.01 pixels for one of the directions.

Ongoing work is focused on applying regression methods to the gathered data to predict the crash type.

Limitations

This work faces limitations that are not fully covered in the preliminary results.

On one hand, present results are based on selected videos from the dataset with a clear view of the wheels. On the other hand, this approach neglects the rotation of the wheels, which leads to not detecting the longitudinal slip if it exists.

Since our dataset is made mainly from dashcam videos, we assume a frame rate of 30 FPS. Additionally the motion of the camera could lead to outliers where no-crash videos generate data with variations similar to crash videos (see Table 1).

To summarise, the presented methodology allows to gather data from real-world scenarios. Results show that the obtained data is useful for analysis and could be enhanced if challenges are tackled up. Integrating different techniques into the pipeline used in this work, several of the mentioned challenges can be solved. This research will continue in development to enhance results.

Conclusion

In this work, we have created a video dataset of single-cyclist crashes. Using computer vision techniques on these videos, we extracted 2-dimensional motion data of the wheels in different crash scenarios. We conclude that crash configurations where large vertical motions occur are easier to identify. Further work will be to assess near-miss situations.

References

[1]: European Commission (2024) Facts and Figures Cyclists. European Road Safety Observatory. Brussels, European Commission, Directorate General for Transport.

[2]: Utrianen, R. (2020) Characteristics of Commuters' Single-Bicycle Crashes in Insurance Data.

[3]: Roni Utriainen, Steve O'Hern & Markus Pöllänen (2023) Review on single-bicycle crashes in the recent scientific literature, Transport Reviews, 43:2, 159-177, DOI: 10.1080/01441647.2022.2055674

[4]: Hafiz, A.M., Bhat, G.M. A survey on instance segmentation: state of the art. Int J Multimed Info Retr 9, 171–189 (2020). https://doi.org/10.1007/s13735-020-00195-x

[5]: Guo, Y., Liu, Y., Georgiou, T. et al. A review of semantic segmentation using deep neural networks. Int J Multimed Info Retr 7, 87–93 (2018). https://doi-org.tudelft.idm.oclc.org/10.1007/s13735-017-0141-z

[6]: Z. Zou, K. Chen, Z. Shi, Y. Guo and J. Ye, "Object Detection in 20 Years: A Survey," in Proceedings of the IEEE, vol. 111, no. 3, pp. 257-276, March 2023, doi: 10.1109/JPROC.2023.3238524

[7]: Gildea, K., Hall, D., & Simms, C. (2021). Configurations of under-reported cyclist-motorised vehicle and single cyclist collisions: Analysis of a self-reported survey. Accident Analysis &Amp; Prevention, 159, 106264. https://doi.org/10.1016/j.aap.2021.106264

[8]: Gildea, K., Hall, D., Cherry, C. R., & Simms, C. (2024). Forward dynamics computational modelling of a cyclist fall with the inclusion of protective response using deep learning-based human pose estimation. Journal of Biomechanics, 163, 111959. https://doi.org/10.1016/j.jbiomech.2024.111959