ABSTRACT
This paper is about a small eye tracking study for scan path classification. Seven participants played Mario Kart while wearing a head mounted eye tracker. In total, we had 64 recordings, but one had to be removed (Only 79 gaze samples were recorded). We compared different scan path classification features to estimate the performance of the participants based on the ranking they achieved. The best performing feature was ENCODJI which incorporates saccades and the heatmap in one feature. HOV, which uses saccade angles, performed well for all tasks but was outperformed by the heatmap (HEAT) for the last two groups.

CCS CONCEPTS

KEYWORDS
Eye Tracking, Gaze, Mario Kart, Classification, Score, Scanpath

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).


1 INTRODUCTION AND STUDY DESCRIPTION
The study was conducted at a university with seven people. Two participants were post doc, four were PhD students, and one participant was a student. Four of the participants are male and three are female. In each recording session, four participants played against each other together with bots and each of the participants has worn a Look! [Kübler 2021] head mounted eye tracker. In total 16 sessions in groups of four were conducted. After each session, the ranking position of each participant was stored. Therefore, we had 64 recordings in total and a ranking position for each recording. One recording (Number 35) had only 79 gaze points, which is why we removed this recording from our evaluation. There were no time restrictions nor any special instructions for the participants. They could look freely.

2 EVALUATION
For the classification task, we grouped the twelve possible ranking positions into three groups. The first group consists of the rankings 1 to 3, the second group of the rankings 4 to 7, and the third group consists of the rankings 8 to 12. This was done to simplify the classification task, since we expect that there will be not much difference between the first and the second ranking, for example. In contrast to this, there should be differences between very good performing players and less good performing players. In total, we had 64 recordings but had to remove one since there were only 79 gaze points saved. Since 63 recordings is a small amount of data, we increased the dataset by selecting 100 subsets of each recording. Each subset of a recording was computed by randomly selecting 60% up to 80% of the recording. We evaluated three different features
namely histogram of oriented velocities (HOV) [Fuhl et al. 2018], heatmap (HEAT) [Fuhl et al. 2021], and ENCODJI [Fuhl et al. 2019] in a tenfold cross validation. In each fold for the cross validation only the subsets of recordings were allowed so that there is no data of the same recording in two folds. With 63 recordings, this means that each of the first 7 folds contained the subsets of 6 recordings and each of the last three contained the subsets of 7 recordings. For classification, we used an ensemble of bagged decision trees with a fixed amount of trees set to 50.

Figure 1 shows the confusion matrices for the different features. As can be seen, the HEAT feature seems to be good for the classification of the two less good performing groups (Group 2 and 3) while the HOV feature outperforms the HEAT feature for the top group (Group 1). Since the ENCODJI feature contains the saccades as well as the heatmap, it performs best for each group. In total, it is still challenging to classify the first group but already highlights, that expertise information lies in the gaze signal [Castner et al. 2020; Hosp et al. 2021]. Further work will go into the extension of the dataset for skill level classification. It is necessary to increase the amount of data and especially the amount of tasks for further research in scan path classification and understanding.

REFERENCES