Predicting Users' Difficulty Perception in a VR Platformer Game

Erdem Murat Liuchuan Yu Siraj Sabah Haikun Huang Lap-Fai Yu

George Mason University

Abstract

Level design challenges posed by novel virtual reality (VR) gaming experiences call for new techniques. Reports of issues in VR games have pointed to drastic differences in difficulty perception between users. To address this issue, we propose a novel approach that tracks user difficulty perception on the manipulation of various game parameters. The collected user data is used to train a recurrent neural network (RNN) to predict the perceived difficulty of game levels. When predicting difficulty perception on a scale of 0-10, our model predicts on average 1.19 points away from the user's actual rating across 16 levels. Our approach presents an effective proof of concept that can open up a plethora of future research avenues for manipulating game difficulty in virtual reality.

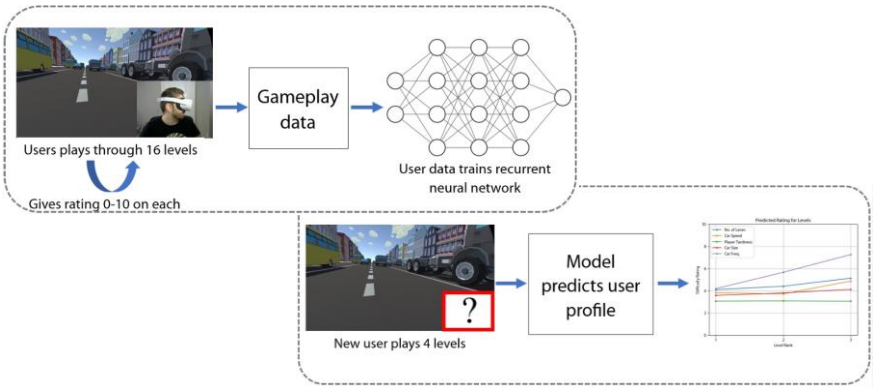

Figure 1: We use RNNs to predict user perception of difficulty through different levels, enabling a deeper understanding of how user experience is impacted by game design and adjustable parameters. We trained our model with data from 57 users to generate predictions for perceived difficulty and user profiles.

Approach

Figure 2 presents an overview of our system. The VR game developed for our experiments is called "Why did the chicken cross the road?". The goal is to cross the street without getting hit by cars. Like many virtual reality games, this was also inspired by a famous non-VR game(Crossy Roads). The game had16levels with differing parameter adjustments. The adjustable parameters for our game were the number of lanes, car speed, size of cars, and car frequency. To collected data we conducted user studies with 57 users (IRB approved). Using the data, we devised an RNN solution that takes level parameters and ratings and ratings given by each user as training data. Using this data, it is able to use the ratings of users playing 4 levels of our game and predict the overall user profile with similar predictions to the ground truth for all 16 levels.

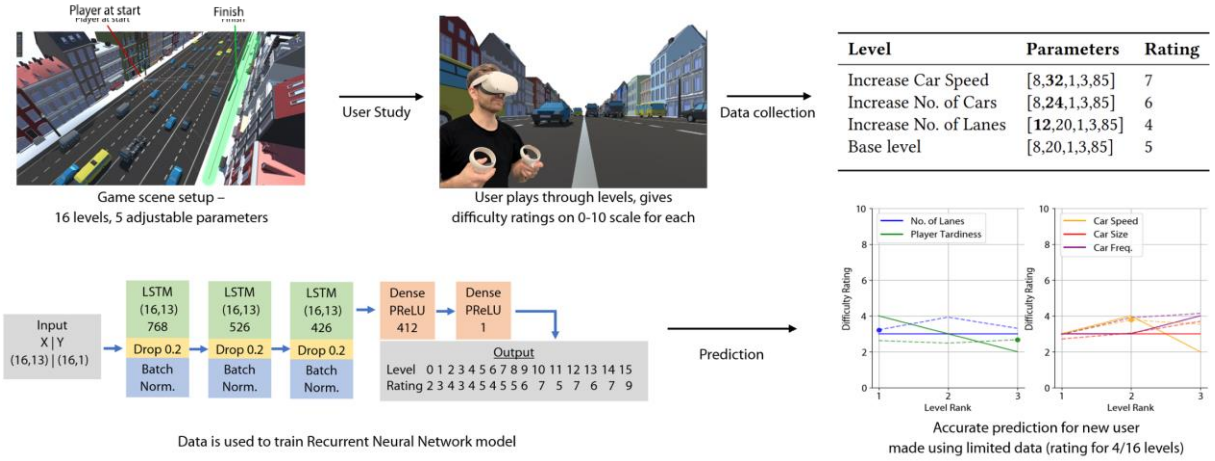

Figure 2: Visualizations of our game setup and overview of our approach

Results

Machine learning allows for a systematic approach to understanding and using data for prediction. It allows for a better understanding of how different parameters affect user experience. To a certain extent, the average rating for all levels can show which levels are more difficult. Our model produced a MAE error of 1.19 (89th per centile) and MSE of 2.61. Using one-way ANOVA test, we found with 95% confidence that our model outperforms five other popular machine learning models in accuracy (refer to Table 1).

Table 1: Mean absolute errors (MAEs), mean squared errors (MSEs) and one-way ANOVA p-values of testing each method against RNN (based on the MSEs of the predictions). All results were obtained over 5-fold cross-validation.

ML Method	MAE	MSE	p-value (vs. RNN)
MLP	3.65	22.8	≈ 0
Gaussian Naive Bayes	1.75	5.94	≈ 0
SVM Regressor	1.49	3.74	0.015
Random Forest	1.39	3.61	0.025
RNN	1.19	2.61	-