Mahjong Master Project Proposal

Cunhan Zou (A0248334A)   *Kaisheng Deng (A0242101E)*   *Ziwei Peng (A0248341H)*

Motivation

Mahjong is a tile-based game that was developed in the 19th century in China and has spread throughout the world since the early 20th century. The game and its regional variants are widely played throughout East and Southeast Asia. Some of the interests in playing Mahjong are served for gambling, socializing, and entertaining, while the main attractiveness behind is to figure out rules out of chaos.

As all of us are Mahjong players, we would like to make the mahjong game a little bit easier by applying some AI techniques to make decisions for us. In decision-making, the understanding of the current situation is crucial, which motivates us with the idea to try to identify all Mahjong tiles on deck. It allows a full appreciation of what tiles we have and what tiles are in the pool and discarded.

Description

Our proposal is to implement a classification project to distinguish different Mahjong tiles on the deck. The project involves several stages, from image information extraction to model parsing, labeling, and output classification. We will parse the input images from the Japanese mahjong game application 雀魂麻将 (Maj Souls) as the source of the dataset. This project is aiming for Japanese mahjong games, which will be slightly different from Chinese mahjong. The tiles in Maj Souls can be divided into four different types, Wan, Tiao, Tong, and Character. The first three types each have nine different tile representations, ranging from number one to nine. The last type of Character includes seven distinct tiles, namely Dong, Xi, Nan, Bei, Zhong, Fa, and Bai. All tiles will add up to 34 different image representations in total. Our project targets to mainly identify the above 34 different Mahjong images.

Proposed Solution

We are proposing a model to first train and identify the images we collect from Maj Souls, forming a dataset of 34 different labels. As the image might be oriented differently or slightly miss certain pixels and data augmentation methods can be applied to our dataset, our training will include all possible image conditions. Accuracy of tiles is key to decision-making models, we will aim to achieve around 100% accuracy on validation and test datasets. The logic behind Mahjong is somewhat straightforward, our model would expand further to implement automated decision-making for players, nevertheless limited by the timing and resources of our project.

Milestones and Tasks for Each Member

Henry Zou: data collection and labeling, dataset 1-30, preprocessing; Kaisheng Deng: data collection and labeling, dataset 31-60, preprocessing; Ziwei Peng: data collection and labeling, dataset 61-90, preprocessing. During these processes, we are constantly brainstorming, and incorporating final thoughts into models, such as data analysis, MLP deep learning baseline, CNN/RNN/ANN, etc.