Atari-HEAD

Game images and decision-making prediction

Team Apollo
Nicole Basinski, Danny Wan, Jason Xing, Nydia Chang

Overview

Problem:

It is hard to get quantifiable information on players' decision-making process in video games. Information on decision making is limited to player input into a controller. This gives us no information from players before they make an input.

Solution:

Team Apollo created a model that utilizes a player's gaze positions and video game frames to predict their actions. This gives us measurable data on players' decision-making process.

Stakeholders:

Game developers that want to create more engaging video games.

What this is?

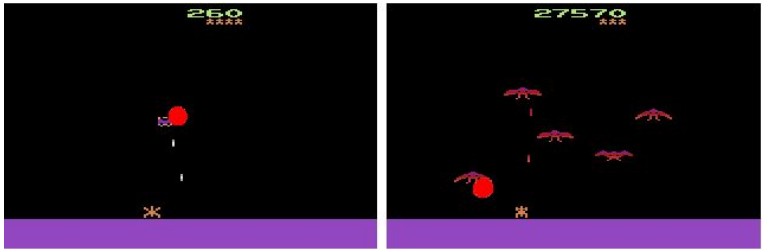
Atari Human **E**ye-Tracking **A**nd **D**emonstration Dataset (Atari-HEAD)

- Provides information about human decision-making strategies
- Gaze information - human motivation behind actions and the anticipated rewards

http://ai.stanford.edu/~zharu/publications/

Semi-frame-by-frame game playing

- Game pauses until action
- Allows multiple eye movements per frame
  - Reduces inattentional blindness
  - Allows sophisticated planning
  - Allows enough decision time in order to obtain near-optimal decisions



(a) 250ms (b) 1200ms

(a) 217ms (b) 1158ms

Data selection for training

Selection:

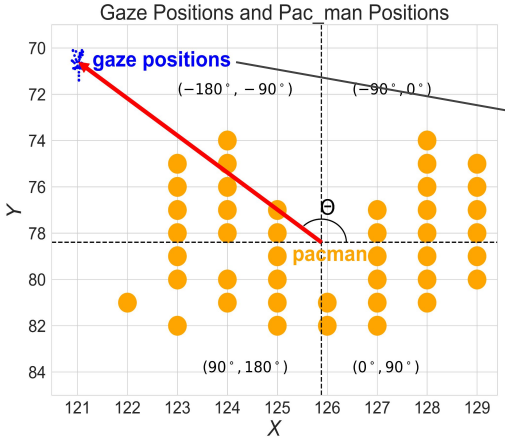
- Ms Pac-Man games (over a set of 20 games)
- 2-hour, ‘highscore’ gameplay trials for training
- ~15,000 actions
- 1.7 million gaze locations

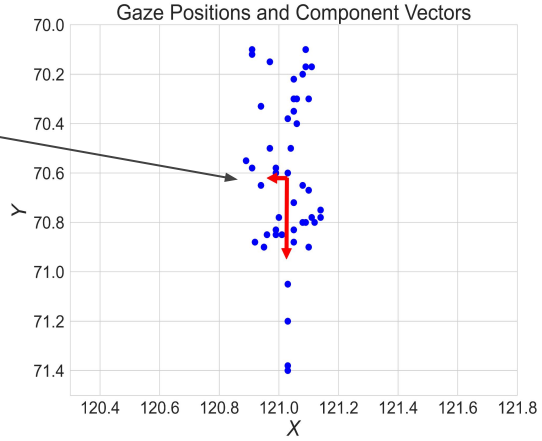
Source:

- Large-scale dataset was obtained from
  https://zenodo.org/record/3451402
  https://arxiv.org/pdf/1903.06754.pdf

Modeling approach: features extraction

*Find **possible features** that impact the player's action (parameter of interest).*





gaze

image

1. Mean gaze position, start position, end position relative to the Pac-Man.

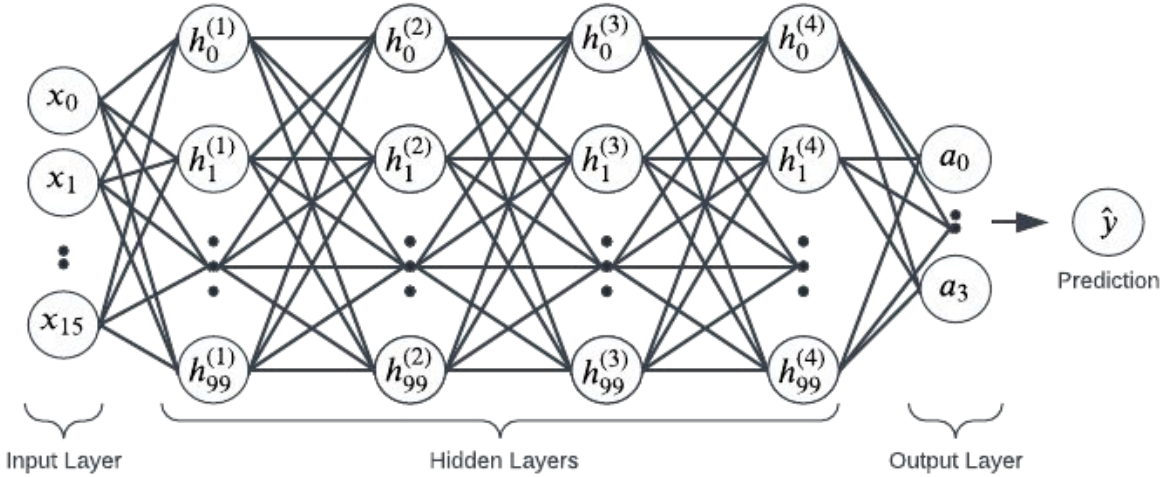
2. Variance, component of gaze positions

3. Position of Pac-Man and possible action in each frame (up, down, left, right).

4. Nearest ghost positions relative to the Pac-Man

Multi-layer Perceptron (MLP) Architecture

- Input Layers
- 4 Hidden Layers containing 100 Nodes each
- Output Layers



Prediction

80% training, 20% test run and prediction

1. Normal logistic regression

Testing Accuracy:
49% (not good)

	Predicted 0	Predicted 1	Predicted 2	Predicted 3
Actual 0	426	6	69	159
Actual 1	165	62	425	91
Actual 2	140	60	562	175
Actual 3	128	2	103	395

2. Sklearn multilayer neural network

Testing Accuracy:
73%

	Predicted 0	Predicted 1	Predicted 2	Predicted 3
Actual 0	458	56	80	66
Actual 1	49	553	87	54
Actual 2	70	87	705	75
Actual 3	57	48	64	459

Atari-HEAD: Game images and decision-making prediction

Nicole Basinski, Nydia Chang, Danny Wan, & Jason Xing

Our model uses a single frame of a Ms Pacman game, as well as eye gaze data, to predict the next move a human might choose in an optimal setting.

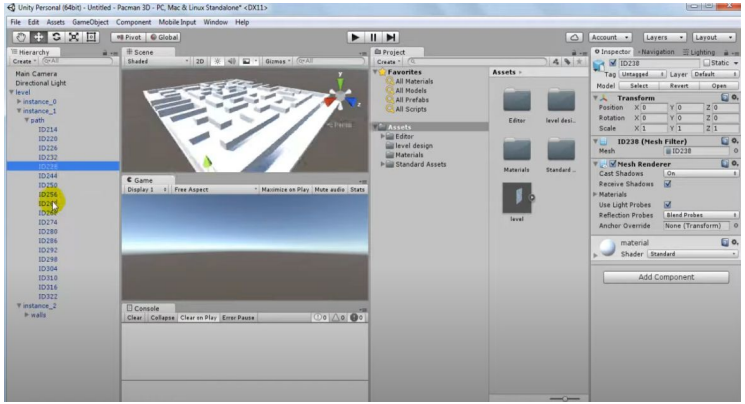
To simulate eye gaze data, please click on the image where you would like the 'eye gaze' target to be. *Hint: In gameplay, the human eye gaze tends to be where the player is planning to go next.*

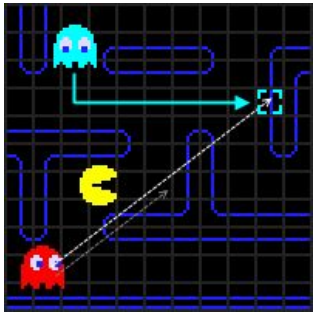
Use case & future work:

Use case:

Video Game Development

- Level Design
- Game A.I.





Future work:

- Optimize hyperparameters to increase accuracy
- Include time elapsed until the human player made a decision
- Incorporate modeling technique to different games

Citations

Ruohan Zhang, Calen Walshe, Zhuode Liu, Lin Guan, Karl S. Muller, Jake A. Whritner, Luxin Zhang, Mary Hayhoe, & Dana Ballard. (2019). Atari-HEAD: Atari Human Eye-Tracking and Demonstration Dataset (Version 4) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.3451402

https://github.com/nicole-sb/erdos-project-2022--atari-HEAD/tree/main/Data_prediction