RL Learnings

1. Was not working at all

State represented as

* Player attributes
  + Player position
  + Player step size
  + Player step size
* 4 closest enemies and 4 closest rewards
  + Velocity
  + Position
  + Size

Each env step we would replay and train

Each step was one sequence of actions

**Results:**

Would not learn the states at all

State values would be almost the same for each action no matter what

Would just run into wall and do nothing

1. Improvement

Simplify the problem:

* Instead of dodging many enemies and collecting many rewards… just remove enemies and have agent collect rewards (validate that the RL algorithm is valid)

Restructuring the state space:

* Now we look at the last 4 states as “one” state and pass an action that will be executed over the next 4 steps. Helps significantly with learning as 700x700 board made it hard to learn
* Add penalty for crashing into walls to force movement
* **Add distance to walls as state attribute**
  + **The agent learned to not hit the edges WAY before it learned to pick up rewards… this suggests it is better to pass distances to the player rather than just positions in the state space as arguments**

Tweak parameters

* Set minimum exploration quite high at 20%
* Only replay every 10 steps to let us get more experience
* Ensure rewards are constant despite having 4 frames

Best results so far:

* 1 object detected
* Learning rate 0.005
* Gamma 0.95
* Structure, 48, 24, 12, 4
* Rewards 100, -200, -1
* Plateau after about 8 or 9 trials
* Pass player pos, player step, player size, distance to walls,
  + Enemy and good velocities and distances and if they are above below, left or right

Next Test (state shape now 148 for sum of 4 frames) **Best so far**

* Remove player position and size
* Detect 2 objects rather than just 1
* Add a larger layer to network
* Reduce Learning rate to 0.001
* Rewards 100, -150, -1