Meeting summaries DRL project

Meeting 1: 27.07.2022, 12 pm

About the Idea:

- Why not use the fMRI data? Because it is not publicly available
- Three approaches, start with one and then do the other ones later if we need more

Approach 1 (start)

Train agent on games and compare salience maps
-> should be relatively easy (not so much risk)
-> Leon S. send material from Tensorflow about salience maps, we might need to optimise it
-> then everything in place for approach 2

Approach 2:

Give agent information about human gaze and compare learning to the normal one from before
-> also relatively easy

Approach 3:

Create more human like vision for the agent
E.g. restricting the view or blurring the image
LeonS: more experimental & risk
Ideas: having one model deciding what to do and where to go next, vs. Two models
-> blurring difficult because the mask would have to be differentiable
-> discuss that with Leon S. if we are there in the project

Decision on which game

Start with game with good performance in comparisons in other papers (e.g. rainbow DQN)

DQN implementation:

1. Using existing Rainbow DQN implementation
    1. Stable baselines has good ones
    2. Need Approach 1 and 2 for a good project
    3. Choose carefully a good working implementation that we can understand

2. Implementing it ourselves
    1. Approach 1 might be enough for a good grade
    2. Rainbow what is reasonable
    3. Something with advantages (because they are already checking where the network is looking?)
    4. Noisy networks not necessary

- Give feedback on that decision via Email

Workload:

Maybe it is easier to do more different approaches and then not so deeply.

Meeting 2: 08.08.22, 12 pm

Leon S. you sadly did not attend the meeting, so this features only our decisions.

DQN implementation:

- We will further improve and test it
- Implementing prioritised experience replay
- Changing some hyperparameters

Saliency:

- It is okay to use external code for calculations, because there is still enough to implement
- DIfferent methods are going to be compared
- Scale the big q-values maybe with softmax to be able to compute the calculations (otherwise np.exp overflow)

Gaze Prediction:

- Scale heatmaps down to 0-1?
- How to stack the frames most efficiently?

How to use Tensorboard:

- Also works during training
- Command: `tensorboard --logdir path`

Meeting 3: 11.08.22, 12 pm

Saliency maps suggestions:

- If performs bad, maybe the bias in the middle is not good in information gain
- Compare scores of our dqn with the performance of old basic dqn papers
- MNIST classifier, to check if the saliency methods are working

Risks approach 3,

if it does not work, these might be the causes :

- Deepq stuff -> misuse part of the q-surface -> need a large delay to solve that
- DQN always trained on old data
- Idea Leon L: with another target network to environment for blurring

Writing together:

- Todonotes -> write in parallel, review what the others did by writing a comment, or telling others what you are doing in a comment
- Do a general structure at the beginning so you know what goes where

Meeting 4: 13.09.22, 11 pm

Deadline extension:

Because we tried to have a meeting for a week already and the training of our model takes a lot of time (after the first there are 2 more to train) we asked for a deadline extension.

Buffer improvements:

- Do not use normal system memory as runtime memory, because it will be super slow
- Size down the images to uint8
- Buffer should be 4x the batch size already loaded (we already have 8 x)
- Lossless compression algorithm?
- Better use sparse arrays, who only save nonzero values, because the whole background is black

Parameters:

- Try to go in a direction and see how good at beginning of training, because we do not have time for a grid search
- 6 times more training than sampling (as we have) sounds reasonable

Prioritised replay:

What to do with new samples?

- Directly use them once or give them a high value
- TD error of 0-100 (as we have) is to big, maybe normalise the rewards for better training

Saliency:

- Try binary comparison
- Try correlation
- Use 4 different frames
- Use scaling instead of blurring after creating saliency not for every pixel
- Human might look in the future, DQN in the present
- Write down precisely what we did at the end
- Also try with integrated gradients