

Graying the black box: Understanding DQNs - Supplementary material.

000
001
002
003
004
005
006
007
008
009
010
011
012
013
014
015
016
017
018
019
020
021
022
023
024
025
026
027
028
029
030
031
032
033
034
035
036
037
038
039
040
041
042
043
044
045
046
047
048
049
050
051
052
053
054

055
056
057
058
059
060
061
062
063
064
065
066
067
068
069
070
071
072
073
074
075
076
077
078
079
080
081
082
083
084
085
086
087
088
089
090
091
092
093
094
095
096
097
098
099
100
101
102
103
104
105
106
107
108
109

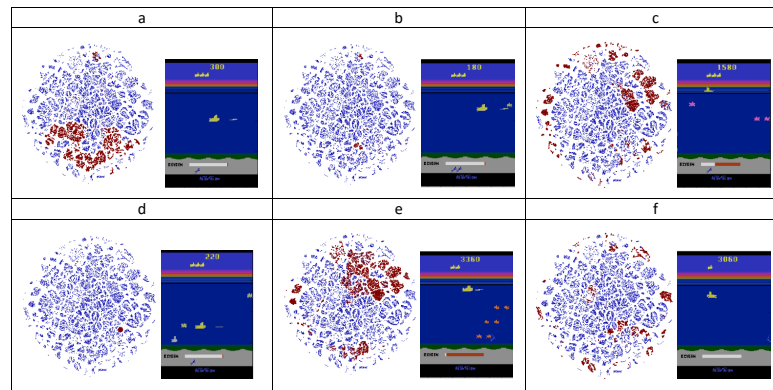


Figure 1. Seaquest states partitioned on the t-SNE map by different measure. a: Number of taken divers = 1. b: Number of taken divers = 2. c: Agent is next to sea level (filter by vertical position < 60). d: Agent is never visiting the lower 1/3 part of the frame. We mark the state of maximal diving depth (filter by vertical position > 128). e: Oxygen level < 0.05. f: Oxygen level > 0.98.

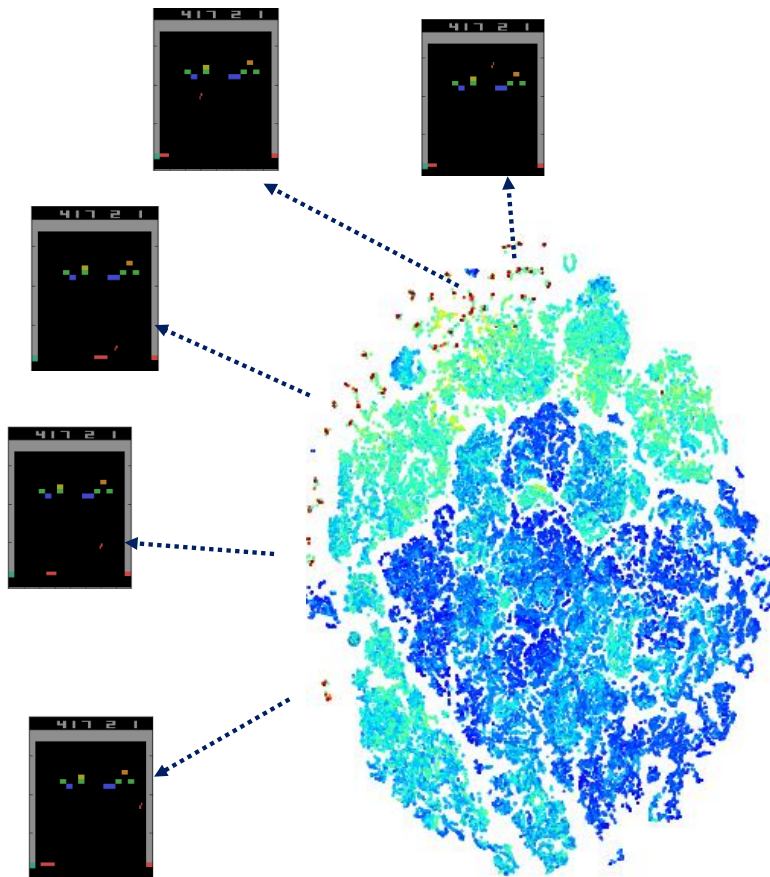


Figure 2. t-SNE for Breakout colored by time. Examples of repeated states by the agent are presented.

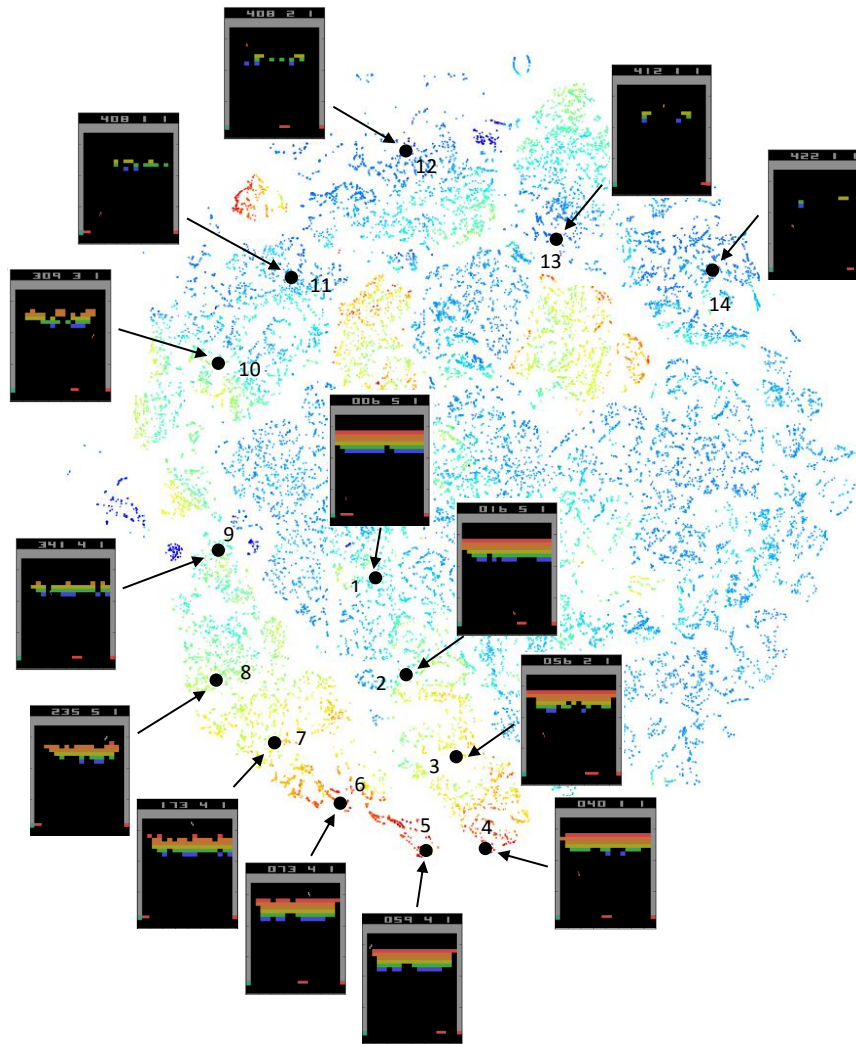


Figure 3. t-SNE for Breakout colored by time. Examples of states along the option path are presented.

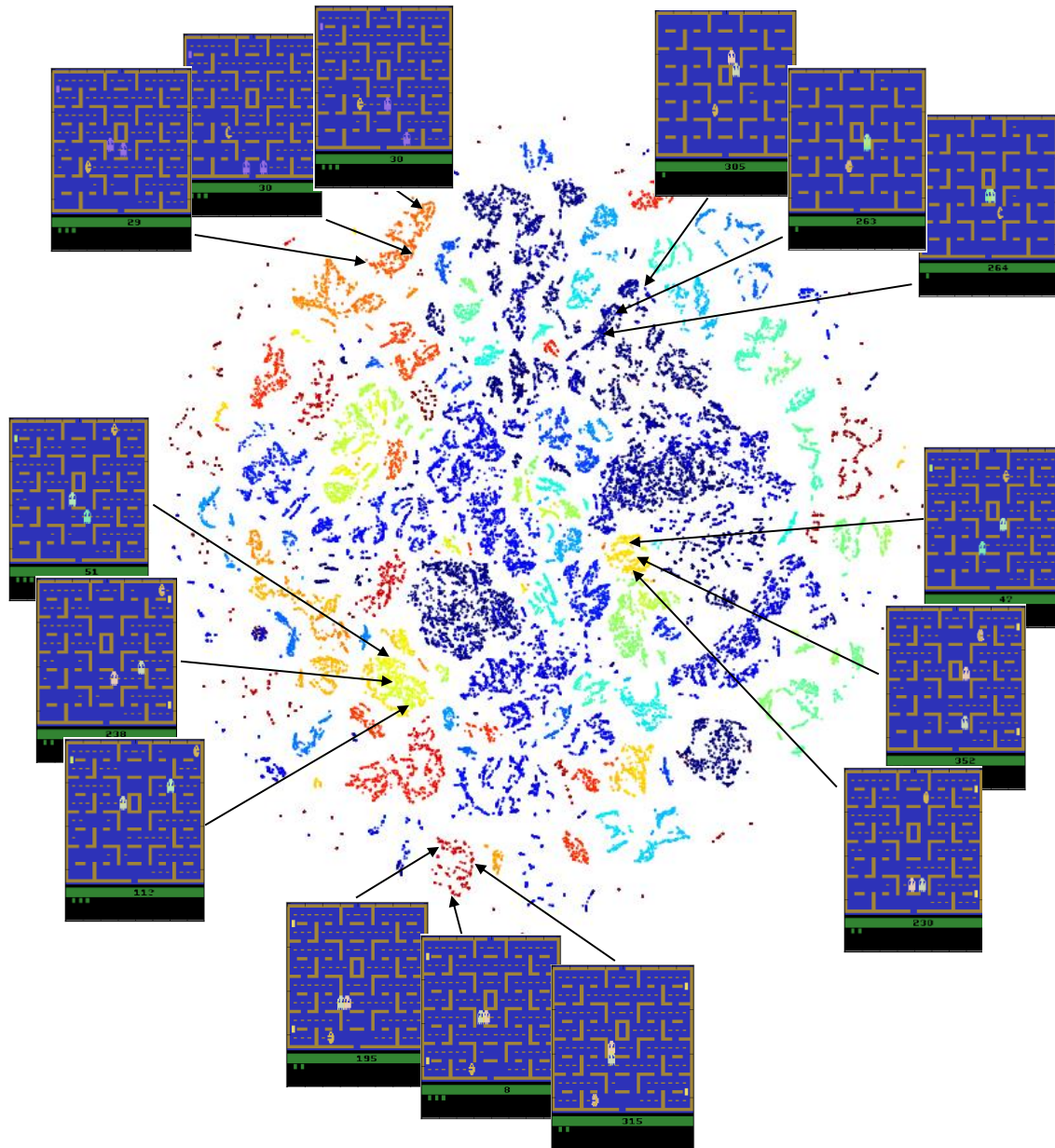


Figure 4. Pacman t-SNE with example states.