

Parameter	Value
batch size	32
replay memory size	3000
discount factor γ	0.85
learning rate	0.0000001
gradient momentum	0.99
distance threshold l_s	0.55
negative reward t_{ter}	-100
positive reward t_{move}	1