Algorithm 1 pseudocode for Pong

1: Xavier initialization: Initialize the weight parameters
2: **while** $i = 1$ to N **do**
3:   Preprocessing
4:   Policy forward
5:   Sample action, receive reward
6:   Policy gradient
7:   **if** Finish one batch: **then**
8:     Rmsprop update
9:   **end if**
10: **end while**