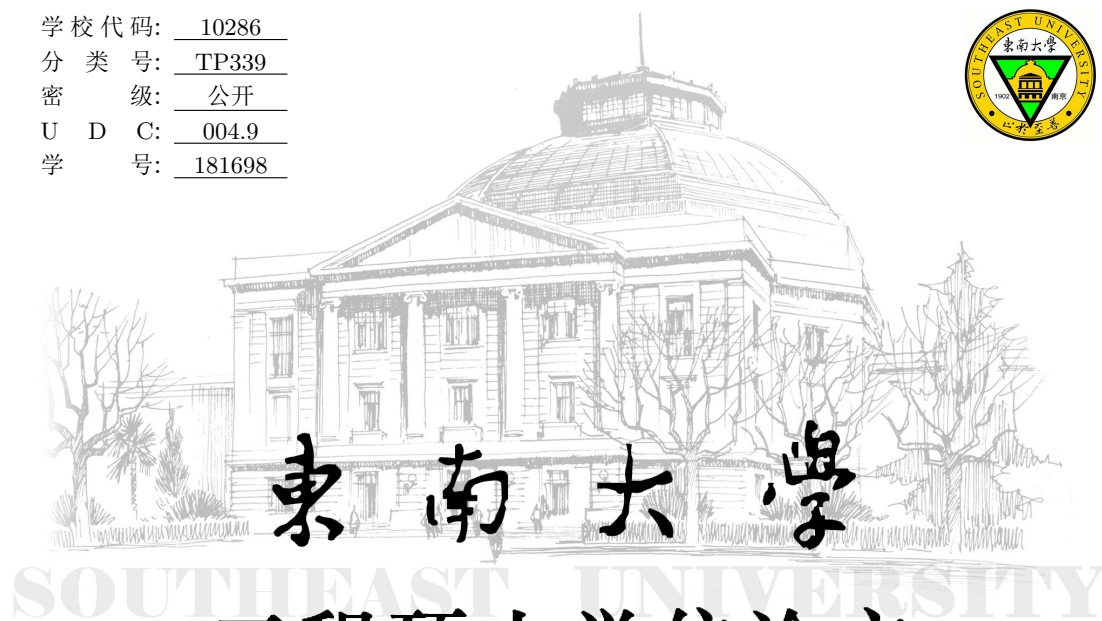


学校代码: 10286
分类号: TP339
密 级: 公开
U D C: 004.9
学 号: 181698



东南大学

工程硕士学位论文

论文题目

研究生姓名: _____ 作者姓名 _____
导师姓名: _____ 导师姓名 教授

申请学位类别 _____ 工学硕士 _____ 学位授予单位 _____ 东南大学 _____
一级学科名称 _____ 专业 _____ 论文答辩日期 _____
二级学科名称 _____ 学位授予日期 _____
答辩委员会主席 _____ 评 阅 人 _____

2021 年 5 月 21 日

東南大學

硕士学位论文

论文题目

专业名称: 专业

研究生姓名: 作者姓名

导师姓名: 导师姓名 教授

TITLE OF THE THESIS

A Thesis submitted to

Southeast University

For the Academic Degree of Master of Engineering

BY

Name

Supervised by:

Prof. Name

School of Computer Science and Engineering

Southeast University

May 21, 2021

东南大学学位论文独创性声明

本人声明所呈交的学位论文是我个人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得东南大学或其它教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

研究生签名：_____ 日期：_____

东南大学学位论文使用授权声明

东南大学、中国科学技术信息研究所、国家图书馆、《中国学术期刊（光盘版）》电子杂志社有限公司、万方数据电子出版社、北京万方数据股份有限公司有权保留本人所送交学位论文的复印件和电子文档，可以采用影印、缩印或其他复制手段保存论文。本人电子文档的内容和纸质论文的内容相一致。除在保密期内的保密论文外，允许论文被查阅和借阅，可以公布（包括以电子信息形式刊登）论文的全部内容或中、英文摘要等部分内容。论文的公布（包括以电子信息形式刊登）授权东南大学研究生院办理。

研究生签名：_____ 导师签名：_____ 日期：_____

摘 要

摘要概括全文，应该包含且仅包含核心问题、工作、方法、算法、结论以及分析结果。

关键词： 关键词 1； 关键词 2

Abstract

something

Keywords: Key word 1, key word 2

目 录

摘要	I
Abstract	II
插图目录	IV
表格目录	V
术语和符号表	VI
第一章 可能用得上的模板	1
1.1 引用	1
1.2 图片	1
1.2.1 单张图片	1
1.2.2 多张图片	1
1.3 表格	2
1.3.1 表格一	2
1.3.2 表格二	2
1.4 公式	3
1.5 算法	3
参考文献	4
致谢	5
作者简介	6

插图目录

1.1	单张图片	1
1.2	多张图片	2

表格目录

1.1 表格一 2

1.2 表格二 2

术语和符号表

缩略词	英文全称	中文全称	
DRL	Deep Reinforcement Learning	深度强化学习	
符号	含义	符号	含义
t	时间步 t	s_t	时间步 t 的状态

第一章 可能用得上的模板

1.1 引用

深度强化学习^[1, 2]中，深度学习用于复杂环境的状态表征，有效抽取图像、仿真等高维连续状态空间的特征；强化学习通过与环境交互获得经验存入经验池，采样经验更新网络，学习最大化期望回报的目标策略^[3, 4]。

引用文章使用 cite，引用文章内容使用 upcite，多个引用使用英文逗号间隔。

1.2 图片

1.2.1 单张图片



图 1.1 单张图片

单张图片如图1.1所示。

1.2.2 多张图片

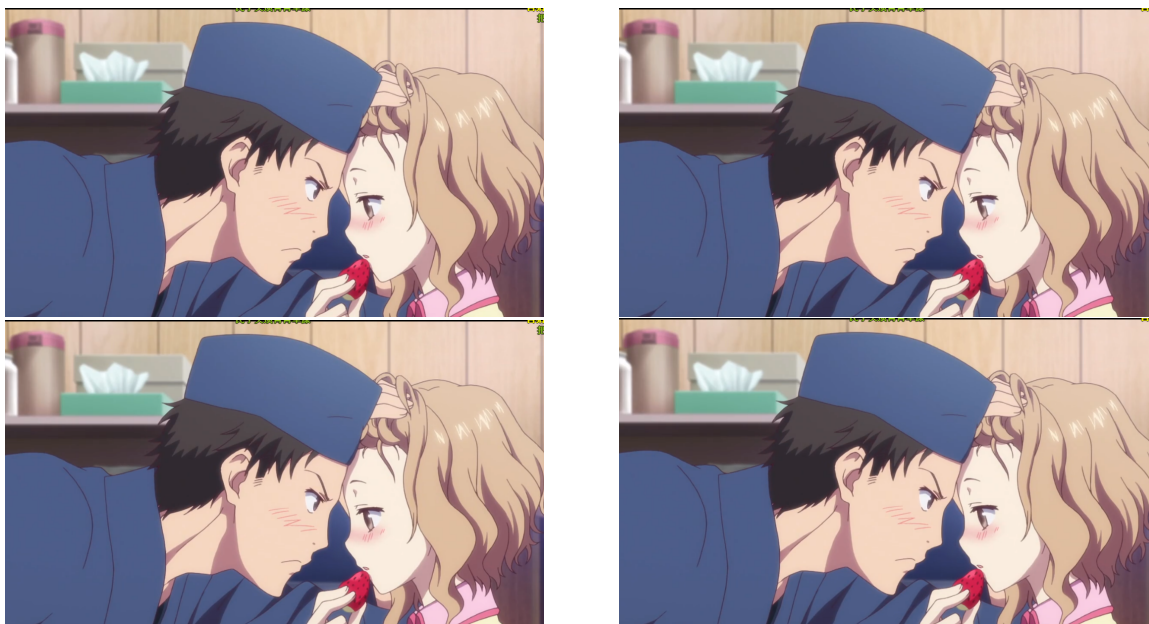


图 1.2 多张图片

1.3 表格

1.3.1 表格一

表格一如表1.1所示，表格中可以引用参考文献方便阅读。

表 1.1 表格一

方法	经验组成	经验池	保留优先级
方法一 ^[1]	转移	先进先出	时间顺序
方法二 ^[5]	转移序列	先进先出	时间顺序

1.3.2 表格二

表格二如表1.2所示。

表 1.2 表格二

数据集	方法一		方法二		方法三	
	指标 1	指标 2	指标 1	指标 2	指标 1	指标 2
数据集 1	0.975	0.53	0.983	0.56	1.000	0.32
数据集 2	0.975	0.53	0.983	0.56	1.000	0.32
平均值	0.989	0.31	0.988	0.29	1.000	0.17

1.4 公式

需要等号对齐的公式写法。

$$\begin{aligned}\nabla_{\theta} J(\pi_{\theta}) &= \int_{\mathcal{S}} \rho^{\pi}(s) \int_{\mathcal{A}} \nabla_{\theta} \pi_{\theta}(a|s) Q^{\pi}(s, a) da ds \\ &= \mathbb{E}_{s \sim \rho^{\pi}, a \sim \pi_{\theta}} [\nabla_{\theta} \log \pi_{\theta}(a|s) Q^{\pi}(s, a)]\end{aligned}\tag{1.1}$$

1.5 算法

算法如算法1.1所示。

算法 1.1 算法

输入: 输入数据

输出: 输出结果

- 1: 步骤
 - 2: for 时间步 $t = 0, \dots, T$ do
 - 3: 循环
 - 4: if 判断条件 then
 - 5: 步骤
 - 6: end if
 - 7: 步骤
 - 8: end for
-

参考文献

- 1 Mnih V, Kavukcuoglu K, Silver D, et al. Playing Atari with Deep Reinforcement Learning[J]. ArXiv, 2013, abs/1312.5602.
- 2 Ye D, Liu Z, Sun M, et al. Mastering Complex Control in MOBA Games with Deep Reinforcement Learning[J]. AAAI, 2020, 6672–6679.
- 3 Silver D, Huang A, Maddison C J, et al. Mastering the game of Go with deep neural networks and tree search[J]. Nature, 2016, 529:484–489.
- 4 Li Y. Deep Reinforcement Learning: An Overview[J]. ArXiv, 2017, abs/1701.07274.
- 5 Karimpanal T G, Bouffanais R. Experience Replay Using Transition Sequences[J]. Frontiers in Neurorobotics, 2018, 12.

致 谢

致谢

姓名

2021 年 5 月于南京

作者简介

个人介绍

硕士期间发表的学术论文

[1] 论文名称 [J]. 东南大学计算机科学与工程学院 2021 年研究生学术论坛, 2021.

硕士期间参与的研究项目

1) xxxxx

心於至善

