Noname manuscript No.
(will be inserted by the editor)

DDPG

subtitle here

Maximilian Gehrke · Tabea Wilke · Yannik Frisch

Received: date / Accepted: date

Abstract TODO

Keywords DDPG · DQN · DPG

1 Introduction

- DQN uses deep networks to estimate the action-value function
  - it can only handle discrete and low-dim action spaces
- discretizing the action space often suffers from the course of dimensionality
- PolicyGradientTheorem from continous space to discrete space presented in DPG paper
- naive extension of DPG with nns turns out to be unstable for challenging problems
- Deep DPG (DDPG): combination of DQN and DPG, where:
  - networks are trained off-policy with samples from a replay buffer to minimize the temporal correlations between samples
  - the networks are trained with target networks to give consistent targets during temporal difference backups
  - batch normalization is used
- DDPG is able to learn from low dim observations (torques etc.), aswell as from high dim observations in pixel space

Your text comes here. Separate text sections with

F. Author
first address
Tel.: +123-45-678910
Fax: +123-45-678910
E-mail: fauthor@example.com

S. Author
second address

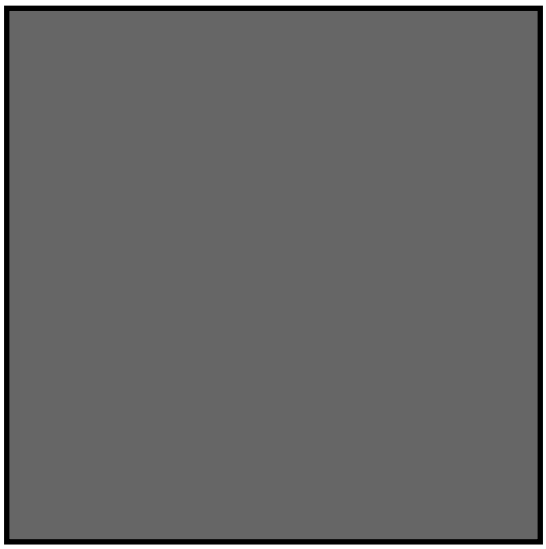

Fig. 1 Please write your figure caption here

Table 1 Please write your table caption here

first	second	third
number	number	number
number	number	number

2 Extensions to the Algorithm

We propose several possible extensions and show their performance on a task. Text with citations [2] and [1].

2.1 Subsection title

as required. Don't forget to give each section and subsection a unique label (see Sect. 2).

Paragraph headings Use paragraph headings as needed.

$$a^2 + b^2 = c^2 \tag{1}$$

References

1. Author, Article title, Journal, Volume, page numbers (year)
2. Author, Book title, page numbers. Publisher, place (year)

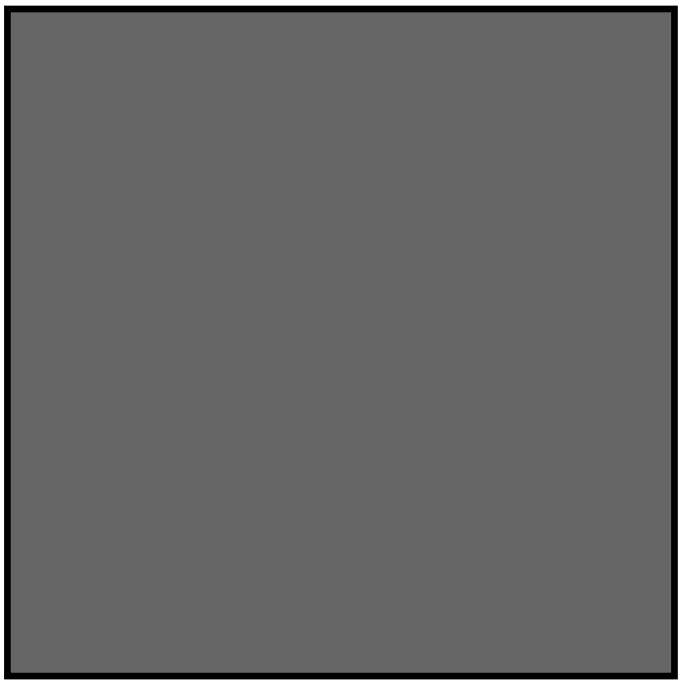

Fig. 2 Please write your figure caption here