

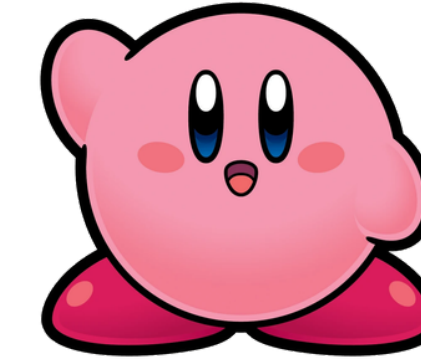


Hier geht's zum Projekt!

Autor

Maurice Füsser

# Schafft es Kirby?



IM AUFTRAG FÜR DAS MODUL:  
ADVANCES IN AI  
BEI PROF. DR. DENNIS MÜLLER

Hochschule Düsseldorf  
University of Applied Sciences  
**HSD**

## Reinforcement Learning für Kirby's Dreamland: Optimierung durch DDQN und Reward-System



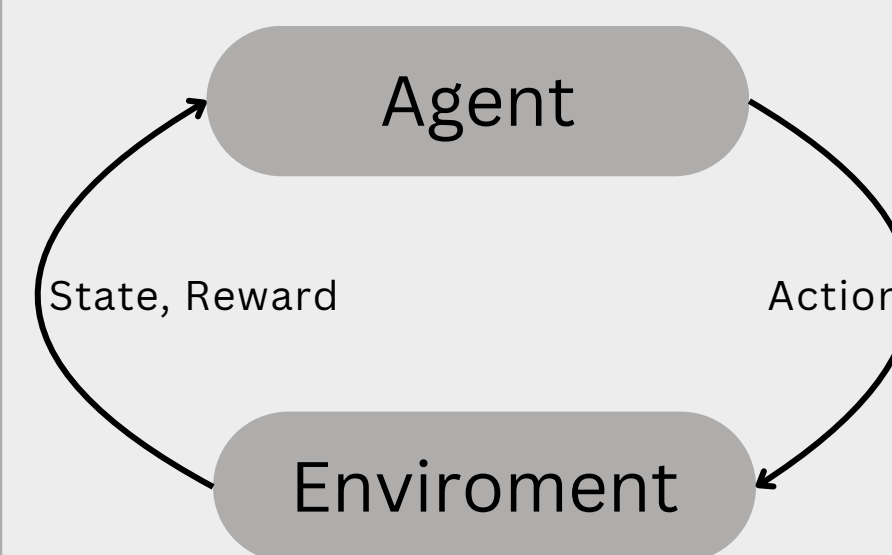
### 01. Einführung

Die Anwendung von Reinforcement Learning in Retro-Spielen bietet eine einzigartige Plattform, um maschinelles Lernen in dynamischen, unvorhersehbaren Umgebungen zu testen.

**Das Ziel:** Eine KI zu entwickeln, die das Spiel erfolgreich abschließt.

### 03. Methodik

Das Training basiert auf einem Double Deep Q-Network (DDQN). Der Agent analysiert Spielframes, wählt basierend auf einem neuronalen Netzwerk eine Aktion und erhält dafür eine Belohnung. Belohnungen und Bestrafungen wurden so gewählt, um ein abschließen des Spiels zu erreichen. Das Ganze verläuft in einer Schleife. Wodurch der Agent am ende lernt und sein Trainingsfortschritt verbessert.



### 04. Aufbau

- Architektur:**
  - CNN zur Verarbeitung der Spielframes (20x16x4)
  - Fully Connected Layer zur Vorhersage der Q-Werte für Aktionen.
- Reward-System:**
  - Kirby erhält für die Bewegung nach rechts eine Belohnung.
  - Kirby erhält für die Bewegung nach links eine Bestrafung
  - Kirby erhält für das Besiegen von Gegnern eine Belohnung
  - Kirby erhält eine Bestrafung wenn er Leben verliert
  - Kirby erhält eine Belohnung wenn er das Level abschließt
- Trainingsdetails:**
  - 25000 Epochen mit jeweils 2500 Schritten
  - Trainingszeit: 5 Tage
- Eine Epoche:**
  - Kirby absolviert das Level
  - Kirby verliert alle Leben
  - Kirby macht 2500 Aktionen
  - Ein Speicherung der Erfahrung verläuft über den Replay Buffer



### 06. Schlussfolgerung

#### Erkenntnisse:

- Arrays statt Bilder
- Training ressourcenschonender
- Belohnungssystem: Entscheidender Faktor für den Lernerfolg.
- Epsilon-Decay: Erfolgreicher Übergang zu strategischem Verhalten.

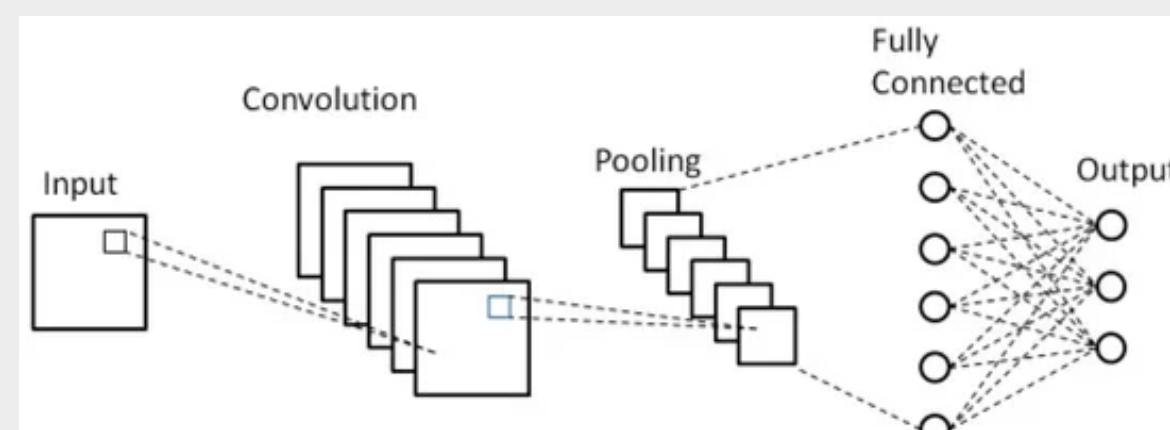
#### Zukünftige Maßnahmen:

- Training auf weitere Level ausweiten.
- Vergleich mit anderen RL-Methoden (z. B. A3C, PPO).
- Optimierung des Belohnungssystems für komplexere Szenarien.

### 02. Hintergrund

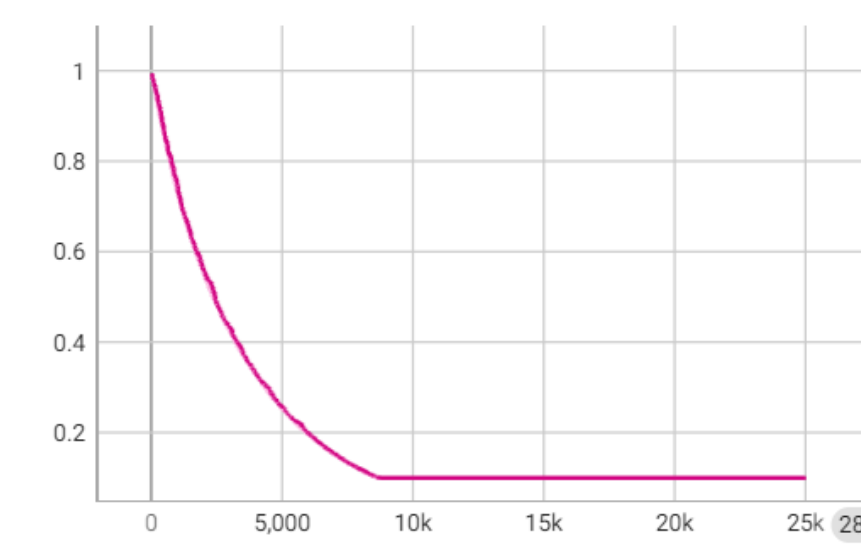
Theoretische Grundlagen und technische Anforderungen

- Reinforcement Learning (RL):**
  - Der Agent lernt durch Interaktion mit der Umgebung, Belohnungen zu maximieren.
- Double Deep Q-Network (DDQN):**
  - Zwei Netzwerke sorgen für stabile Aktionsbewertung.
- Belohnungssystem:**
  - Fortschritt wird belohnt, ineffiziente Aktionen bestraft.
- Effizientes Training:**
  - Spielframes werden als Arrays verarbeitet, nicht als Bilder.
  - Training mit einem CNN
- Strategisches Verhalten:**
  - Anfangs zufällige Aktionen, später gezielte Strategien.
- Replay Buffer:**
  - Erfahrungen werden gespeichert, um stabil zu trainieren.

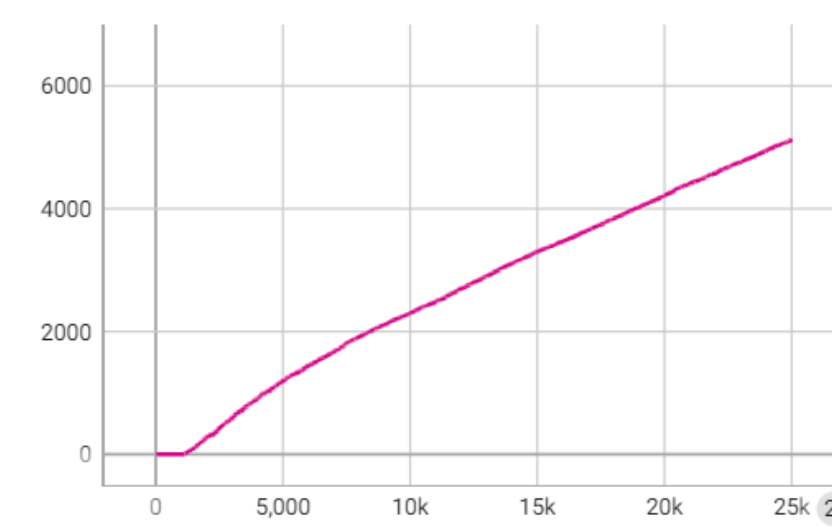


### 05. Ergebnisse

Epsilon



Level Abschlüsse



Episoden Länge



Belohnung / Level Abschlüsse

