DSLE: A Learning Environment for Dark Souls Boss Encounters

Derin Gezgin*, Jim O'Connor*, Tanner Goodwin, Gary B. Parker

Autonomous Agent Learning Lab, Connecticut College

AAAI AIIDE 2026

Motivation

Game Environments as AI Benchmarks

Atari / ALE	Learning across many games through one interface
StarCraft II / SC2LE	Partial information and long-term strategy
Doom / ViZDoom	Visual control in a first-person 3D world
Minecraft / MineRL	Exploration, skill sequences, and demonstrations

Dark Souls: Difficult Decisions in Real Time



Timing

Read attacks and act within short openings.

Positioning

Control distance, dodge, and avoid hazards.

Resources

Manage health, stamina, and healing.

Why Boss Encounters Make a Benchmark Suite

Controlled episodes

A fixed starting save

A bounded arena

A clear goal and terminal outcome

Different demands

Timing and positioning

Hazards and multiple targets

Phase changes and objective sequences

The Full Suite: 22 Boss Encounters



DSLE-5: Five Encounters, Different Skills

A representative starting suite drawn from 22 supported encounters



Asylum Demon
Basic melee timing



Capra Demon
Tight arena and two dogs



Chaos Witch Quelaag
Persistent lava hazards



Ornstein and Smough
Two targets, changing phases



Gwyn, Lord of Cinder
Fast, sustained pressure

II Dark Souls Learning Environment

A Common Interface Across 22 Encounters

Observation	A grayscale frame from the running game
Action	One of 14 movement, attack, roll, heal, or parry actions
Control rate	250 ms per action, about four decisions per second
Episode end	Boss defeated, player death, or step limit

`reset()` and `step(action)`

Single-Episode Interaction Loop

```python
import dsle

env = dsle.make(
  boss="asylum_demon",
  instance="dsr-1",
  obs_mode="grayscale",
  max_steps=7200,
)

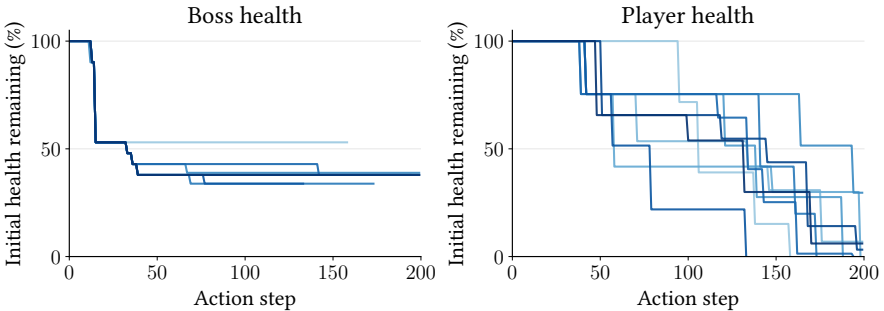
try:
  obs, info = env.reset()
  ret = 0.0
  done = False
  while not done:
    action = env.action_space.sample()
    obs, reward, term, trunc, info = env.step(action)
    ret += reward
    done = term or trunc
finally:
  env.close()
```

Repeatable Experiments in a Running Game

1. **Restore a scenario save.**
   Fix the character, equipment, inventory, and starting location.
2. **Load and enter the encounter automatically.**
   Transfer control to the agent at the start of combat.
3. **Measure outcomes from game memory.**
   Record health, damage, deaths, and boss defeat reliably.

A Fixed Start Does Not Fix the Trajectory

Asylum Demon: ten replays of the same action sequence from the same save



Results

Five Baselines on the Standard DSLE-5 Suite

Method	Controller	Budget per boss
Random	Uniform random actions	100 episodes
Scripted expert	Advance, attack, heal using health	100 episodes
PPO / DQN	Convolutional policies on 84×84 images	5 seeds, 100k steps each
SCOPE	Compressed visual policy, CMA-ES	5 seeds, 1,000 episodes each

Only the Tutorial Boss Is Beaten

Observed wins / episodes under standard saves; learning methods pooled during training

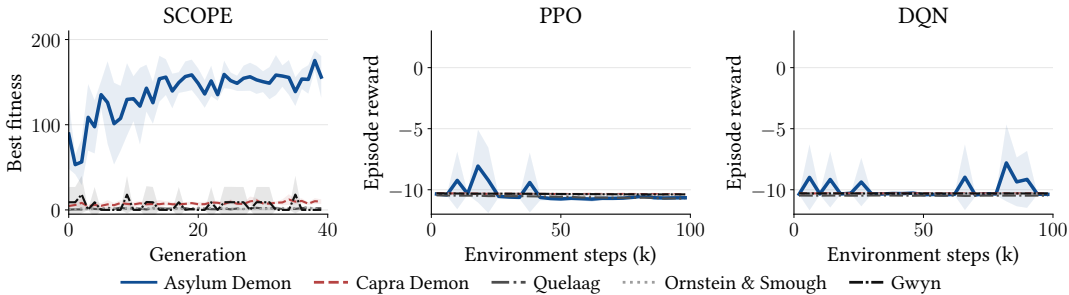
Method	Asylum	Capra	Quelaag	O&S	Gwyn
Random	0/100	0/100	0/100	0/100	0/100
Scripted expert	**63/100**	0/100	0/100	0/100	0/100
PPO	**5/1,971**	0/6,423	0/1,684	0/5,309	0/5,892
DQN	**8/2,469**	0/12,549	0/2,284	0/8,203	0/9,942
SCOPE	**1,065/5,000**	0/5,000	0/5,000	0/5,000	0/5,000

Asylum Demon: scripted expert 63%, SCOPE 21.3%, PPO / DQN below 0.4%.

Every baseline records zero wins on the four harder encounters.

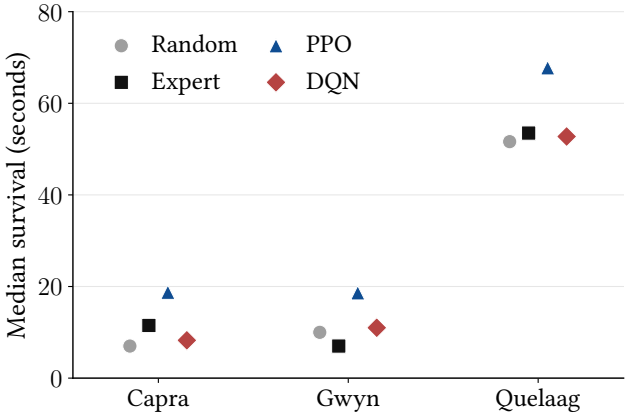
Learning Progress Is Limited Within Budget

Five seeds per boss; curves show means and shaded regions show standard deviations



SCOPE improves on Asylum Demon. PPO and DQN show no sustained gain.

Survival Alone Does Not Measure Progress



Capra and Gwyn

Median survival: 7–19 s.
Agents die before doing much damage.

Quelaag

Median survival: 52–68 s.
Longer survival, almost no damage.

Real-Time Interaction Has a Compute Cost

100,000

environment steps

About 7 hours of action time
before resets and computation

Observed wall-clock time per run

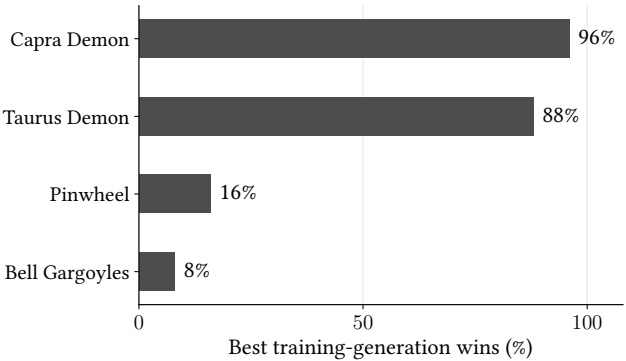
PPO: 9–21 hours

DQN: 12–41 hours

Short episodes require more resets.

Broader Exploration with Advantaged Saves

Beyond standard DSLE-5: SCOPE with every player attribute set to 50



More health, stamina, and damage.

Some early bosses become beatable.

Save states let us change the challenge while keeping the interface.

The Full Suite Also Tests Objective Sequences



Bed of Chaos

Destroy environmental targets.
Navigate a collapsing floor.



Seath the Scaleless

Find and destroy a separate crystal
before the boss becomes vulnerable.

Limitations

- Real-time interaction makes training expensive.
- Main comparisons cover only five bosses.
- Other agent families and larger budgets remain untested.

Conclusions

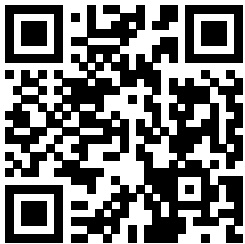
- **DSLE makes 22 distinct boss encounters repeatable.**
  Visual observations, a shared interface, and automated resets.
- **The tested baselines leave substantial room for progress.**
  Only Asylum Demon is beaten under the standard DSLE-5 protocol.
- **The suite supports richer questions about learning.**
  Memory, demonstrations, curricula, and transfer between encounters.

Questions

Thank you

Derin Gezgin

dgezgin@conncoll.edu



arXiv paper



GitHub code