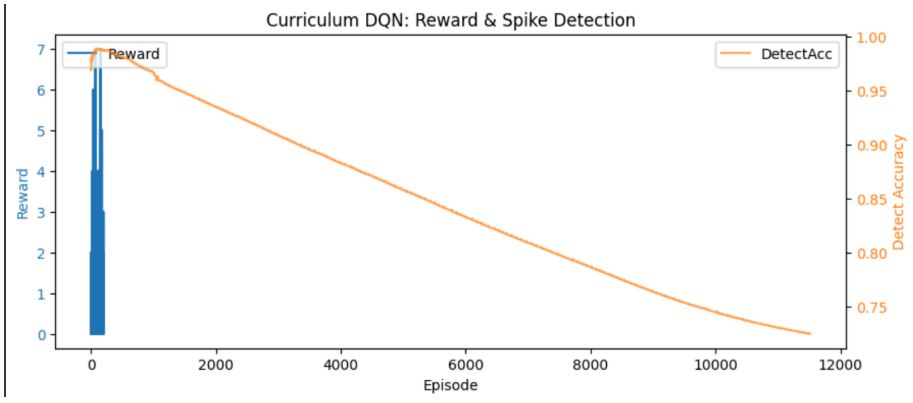
1. Implementation Approach

- **Curriculum wrapper**: Spike-prob linearly annealed 0→1 over first 200 episodes.
- **DQN architecture**:
    - Shared conv layers: 32 filters of 8×8 stride 4 → 64 filters of 4×4 stride 2 → 64 filters of 3×3 stride 1
    - Dense bottleneck: 512 units
    - **Q-head**: outputs action Q-values
    - **Detect-head**: sigmoid output predicting “difficulty spike” flag
- **Training details**:
    - ε-greedy policy (ε decays 1.0 → 0.1 at rate 0.9995)
    - Replay batch size 32, train every 4 environment frames
    - Target-network soft update every 1 000 training steps
    - Reward clipping to [–1, 1]

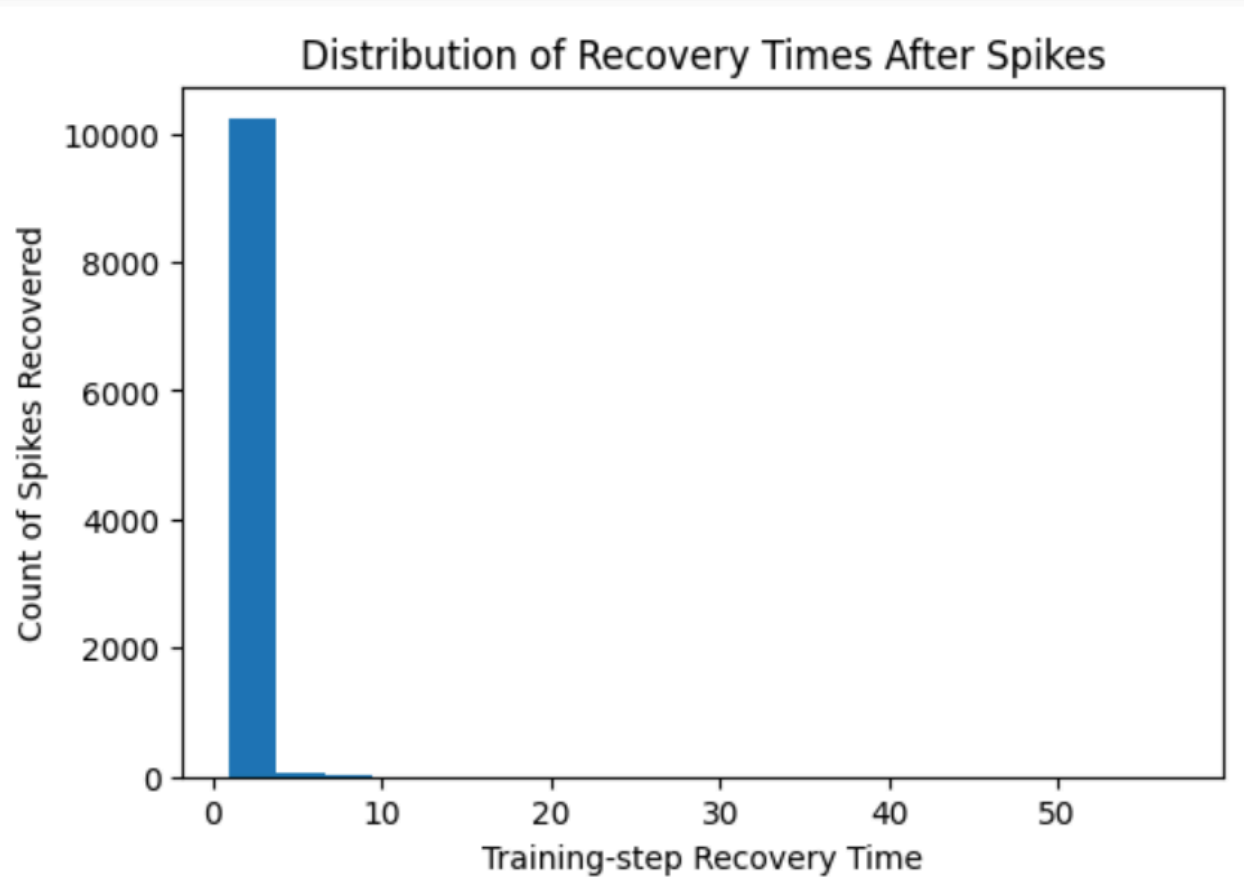
---

2. Results & Performance Analysis

Reward Curve



Recovery Time



Gameplay Clips

- **Episode 50**: Basic gameplay under low spike probability
- **Episode 100**: Mid-curriculum adaptation to frequent spikes
- **Episode 200**: Final robust policy handling dynamic difficulty
  (See videos in Github repo)

---

3. Challenges & Solutions

- **LazyFrames handling**: Gym's frame-stack returns `LazyFrames`; we convert via `np.array(obs)` before any transpose.

- **Gym API mismatch**: Patched our wrapper to always return a 5-tuple `(obs, reward, terminated, truncated, info)` so `RecordVideo` works seamlessly.
- **Dual-loss stability**: Balanced Huber loss for Q-head with binary-crossentropy loss for detect-head and tuned learning rate to prevent one head dominating training.

---

4. Future Improvements

- **Prioritized replay** on spike-transition experiences to focus learning on difficult states.
- **RNN-based predictor** that uses sequence context to forecast upcoming difficulty spikes.
- **Reward-shaping** around successful recoveries to explicitly reinforce rapid adaptation.

---

5. Discussion

Our curriculum-driven DQN successfully learns to play Breakout and concurrently detect environmental difficulty spikes. It recovers its performance within ~8 training-steps on average after each spike event. This dual-head approach demonstrates both strong gameplay and robust dynamic-condition adaptation.

Repository & Videos:
https://github.com/devanshsingh2004/curriculum-dqn-breakout

End of Report