4. PENUTUP

4.1 Kesimpulan

Dengan melakukan pelatihan agen *PPO* hingga 5 juta langkah pada masing-masing *environment*, hasil menunjukkan bahwa performa agen sangat dipengaruhi oleh karakteristik environment yang digunakan. Penggunaan vectorized environment berhasil meningkatkan efisiensi pelatihan secara signifikan, menghasilkan efisiensi sebesar 52,71% atau penghematan waktu 2 jam 36 menit 43 detik dibandingkan dengan *non-vectorized*.

Secara umum, agen PPO menunjukkan hasil yang beragam tergantung pada karakteristik game. Peningkatan signifikan terlihat pada Pong (+111,36% reward) dan Flappy Bird (+98,05%), yang menandakan bahwa agen mampu memahami pola permainan dan meningkatkan performa seiring waktu. Namun, penurunan tajam terjadi pada Breakout dan Super Mario Bros, yang masing-masing mencatat penurunan reward hingga -50% dan -78,17%, mengindikasikan kesulitan adaptasi terhadap lingkungan yang lebih kompleks atau dinamis. Hal ini menegaskan bahwa efektivitas PPO sangat bergantung pada sifat dan struktur dari setiap environment.

4.2 Saran

Saran yang dapat diberikan oleh penulis dalam sistem ini adalah agen dapat dikembangkan lebih lanjut untuk memperluas cakupan eksperimen dan mengeksplorasi lebih banyak variasi lingkungan, agar pemahaman terhadap efektivitas metode semakin menyeluruh.