Analisis dan Diskusi

Metode *Q-Learning* konvensional pada dasarnya hanya sesuai digunakan pada sistem dengan ruang keadaan (state space) yang relatif kecil dan bersifat diskrit. Hal ini disebabkan karena Q-Learning menyimpan nilai untuk setiap pasangan keadaan dan aksi dalam bentuk tabel, sehingga ketika ruang keadaan menjadi sangat besar atau bahkan kontinu, pendekatan ini tidak lagi efisien. Sebagai solusi, dikembangkan *Deep Q-Network* (DQN) yang memanfaatkan jaringan saraf tiruan untuk melakukan aproksimasi fungsi nilai Q. DQN mampu melakukan generalisasi terhadap kondisi baru yang tidak secara eksplisit muncul pada data pelatihan, sehingga dapat menangani ruang keadaan yang luas dan kompleks. Penggunaan *experience* *replay* memungkinkan pemutusan korelasi antar data serta meningkatkan stabilitas pembelajaran. *Target* *network* menjadikan pembaruan nilai lebih stabil dibandingkan pendekatan konvensional. Meski demikian, DQN juga memiliki beberapa keterbatasan, seperti kebutuhan komputasi yang lebih tinggi, kesulitan dalam pengaturan hiperparameter, risiko ketidakstabilan pelatihan, serta keterbatasan saat diterapkan pada *continuous* *action* *space* yang umum ditemui pada aplikasi nyata seperti sistem kendali drone maupun robot industri.

Dalam penerapannya pada **sistem dinamis**, DQN tidak bisa langsung digunakan tanpa penyesuaian tertentu. Salah satu pendekatan adalah dengan melakukan **diskretisasi aksi**, meskipun metode ini dapat mengurangi fleksibilitas kendali. Alternatif lain adalah memanfaatkan varian pengembangan DQN, misalnya ***Double DQN*** untuk mengurangi bias estimasi, ***Dueling* DQN** untuk efisiensi representasi nilai, atau ***Rainbow* DQN** yang menggabungkan berbagai teknik peningkatan. Agar sistem tetap aman, dapat digunakan pendekatan safety-aware reinforcement learning dengan menambahkan batasan sehingga aksi berbahaya dapat dihindari.

Selain itu, pada implementasi nyata biasanya diterapkan strategi ***Sim-to-Real Transfer***, yaitu melatih agen DQN dalam simulasi sebelum diuji pada sistem fisik guna mengurangi risiko. Pendekatan **kontrol hibrida** juga kerap digunakan, di mana DQN difokuskan untuk pengambilan keputusan tingkat tinggi (seperti navigasi atau penghindaran rintangan), sementara kontrol tingkat rendah (misalnya kestabilan *drone* atau pengendalian lintasan) tetap ditangani oleh metode klasik seperti PID atau LQR.