

## INTISARI

*Autonomous Vehicle* (AV) merupakan teknologi dengan kendaraan mampu beroperasi secara otomatis tanpa intervensi dari manusia. Salah satu fokus pengembangan dari AV adalah *Autonomous Valet Parking* (AVP) yang menghadapi tantangan manuver parkir di berbagai kondisi lingkungan. Selain itu, kendaraan harus dapat melakukan manuver parkir secara presisi pada setiap posisi yang ada yang menjadi tantangan tersendiri dalam sistem AVP. *Deep Reinforcement Learning* (DRL) menjadi solusi keterbatasan tersebut dengan menghasilkan strategi yang optimal dalam melakukan manuver parkir dengan memanfaatkan *neural network* untuk pembaruan. Beberapa algoritma DRL sudah diterapkan sebelumnya pada sistem AVP seperti TD3 dan PPO. Namun, terdapat algoritma SAC yang belum diterapkan pada sistem AVP yang merupakan algoritma yang dapat menyeimbangkan tingkat eksplorasi dan eksploitasi sehingga akan cocok untuk sistem AVP. Selain itu, belum dilakukan perbandingan pada ketiga algoritma tersebut di skenario manuver parkir pada *Autonomous Valet Parking* menjadi tantangan untuk pemilihan algoritma terbaik. Penelitian ini bertujuan untuk mengimplementasikan algoritma SAC serta membandingkannya dengan algoritma TD3 dan PPO dengan menggunakan improvisasi pada environment berupa penambahan *noise*.

Penelitian ini membandingkan 3 algoritma DRL berbasis *actor-critic* yaitu PPO, TD3, dan SAC serta menggunakan 4 skenario perbedaan jenis *noise* yang diberikan pada LiDAR yakni Normal(tanpa noise), *Internal Noise*, *External Noise*, dan *Combined Noise*. Performa pelatihan dievaluasi berdasarkan 10 parameter termasuk stabilitas, kecepatan pembelajaran, pengoptimalan strategi, dan keberhasilan dalam manuver parkir. Selain itu, penelitian ini juga menguji generalisasi pada model yang dihasilkan dari proses pembelajaran untuk manuver parkir di seluruh lokasi parkir.

Hasil penelitian menunjukkan bahwa algoritma SAC unggul hampir di seluruh parameter di keempat skenario dengan menghasilkan *average reward* sebesar 144.6209 (tanpa *noise*), 138.0145 (*internal noise*), 135.9286(*external noise*), dan 137.2005(*Combined noise*). SAC juga unggul dalam pengujian generalisasi dengan rata-rata keberhasilan sebesar 98.047% di seluruh lokasi parkir. Hasil tersebut membuktikan bahwa SAC merupakan algoritma RL paling optimal untuk AVP dalam menghadapi berbagai gangguan di lingkungan serta variasi lokasi parkir.

Kata kunci: Pendalaman Pembelajaran Penguatan, Parkir Otomatis, *Proximal Policy Optimization*, *Twin Delayed Deep Deterministic*, *Soft Actor-Critic*

## ABSTRACT

Autonomous Vehicles (AVs) are a technology in which vehicles are capable of operating automatically without human intervention. One of the key areas of AV development is Autonomous Valet Parking (AVP), which faces challenges related to maneuvering in various environmental conditions. In particular, vehicles must be able to perform precise parking maneuvers at different positions, presenting a significant challenge for AVP systems. Deep Reinforcement Learning (DRL) offers a promising solution to these limitations by generating optimal parking strategies using neural networks for policy updates. Several DRL algorithms, such as Twin Delayed Deep Deterministic Policy Gradient (TD3) and Proximal Policy Optimization (PPO), have previously been implemented in AVP systems. However, the Soft Actor-Critic (SAC) algorithm, which is capable of balancing exploration and exploitation, has not yet been applied to AVP systems. This characteristic makes SAC a suitable candidate for AVP applications. Furthermore, no comparative analysis has been conducted among the three algorithms in parking maneuver scenarios for AVP, presenting a challenge in identifying the most suitable algorithm. This study aims to implement the SAC algorithm and compare its performance with TD3 and PPO through an enhanced simulation environment by introducing noise into the system.

The research compares three actor-critic-based DRL algorithms: PPO, TD3, and SAC. It utilizes four noise scenarios applied to the LiDAR sensor: Normal (no noise), Internal Noise, External Noise, and Combined Noise. Training performance is evaluated using ten parameters, including stability, learning speed, strategy optimization, and success rate in executing parking maneuvers. In addition, this study assesses the generalization capability of the trained models in performing parking maneuvers across all parking locations.

The results demonstrate that SAC outperforms the other algorithms across most parameters in all four scenarios, achieving average rewards of 144.6209 (no noise), 138.0145 (internal noise), 135.9286 (external noise), and 137.2005 (combined noise). SAC also excels in generalization testing, with an average success rate of 98.047% across all parking locations. These findings indicate that SAC is the most optimal RL algorithm for AVP, effectively handling environmental disturbances and variations in parking locations.

Key words: Deep Reinforcement Learning, Parking, Proximal Policy Optimization, Twin Delayed Deep Deterministic, Soft Actor-Critic