A Implementation details

All of the baseline algorithms come from the code library $\mathcal{CORL}$: [https://github.com/tinkoff-ai/CORL]. All experiments run on a server equipped with an Intel® Xeon® Gold 6254 CPU @ 3.10GHz and NVIDIA GeForce RTX 3090 GPU.

B Additional Results

B.1 Additional Evaluations

In addition to the integration of FANS with TD3 presented in the main text, we also applied FANS to AWAC. As shown in Table 6, AWAC + FANS demonstrates a significant performance improvement over the original AWAC algorithm across a range of offline RL tasks. Notably, in challenging tasks such as *hopper-medium-expert* and *walker2d-medium-expert*, AWAC + FANS boosts the scores from 52.7 to 110.3 and from 49.4 to 109.6, respectively. These improvements are not only substantial but also come with significantly reduced variance, indicating more stable and reliable policy behavior.

Moreover, the average performance across all tasks increases from 67.7 to 83.9, further highlighting the general and consistent enhancement brought by the FANS module. These results validate the effectiveness of our approach across diverse environments and demonstrate its potential as a general-purpose enhancement to existing offline RL methods.

Table 6: Performance comparison on D4RL locomotion tasks over the final ten evaluations and five seeds (normalized scores). We **bold** the highest mean.

Tasks	AWAC	**AWAC + FANS**
halfcheetah-medium	**49.5** ± **0.6**	48.9 ± 0.5
halfcheetah-medium-replay	44.7 ± 0.7	**44.9** ± **0.2**
halfcheetah-medium-expert	93.6 ± 0.4	**94.4** ± **0.4**
hopper-medium	74.5 ± 9.1	**76.2** ± **6.1**
hopper-medium-replay	96.4 ± 5.3	**99.8** ± **2.0**
hopper-medium-expert	52.7 ± 37.5	**110.3** ± **0.7**
walker2d-medium	66.5 ± 26.0	**81.9** ± **0.7**
walker2d-medium-replay	82.2 ± 1.1	**88.9** ± **3.8**
walker2d-medium-expert	49.4 ± 38.2	**109.6** ± **1.2**
Average	67.7	**83.9**

B.2 Analyses of FASN's Structural Extensions

We conduct a systematic scaling analysis of the critic network by varying its depth (1–4) and width (64–512). For width scaling, the critic depth is fixed at 2 blocks; for depth scaling, the critic width is set to 256, following our default setup.

The results show that the best performance is achieved when the critic has depth = 2 and width = 256, yielding high returns with low standard deviations across all three tasks (*halfcheetah-medium*, *hopper-medium*, and *walker2d-medium*), indicating strong stability.

Overall, increasing the depth and width of the critic generally leads to performance improvements, suggesting that higher model capacity enhances the representational power of the value estimator.

However, we also observe significant instability under certain configurations (e.g., depth = 4 or width = 64/512), particularly on the *hopper-medium* and *walker2d-medium* tasks, where the standard deviations are notably large. This highlights the trade-off between model capacity and training stability, and the importance of balancing expressiveness with generalization.

In summary, properly scaling the critic network can significantly boost performance, but must be done with care to avoid instability.

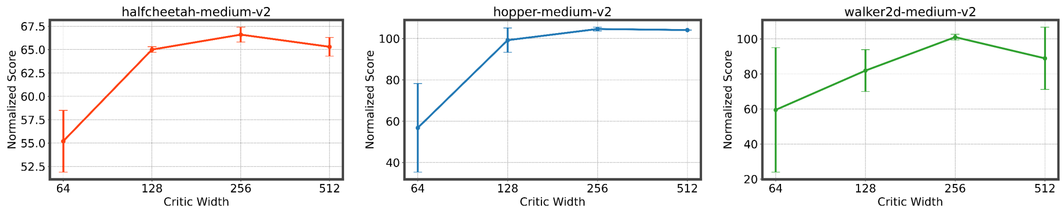

Figure 4: Performance of TD3 with FANS by varying width for the critic network.

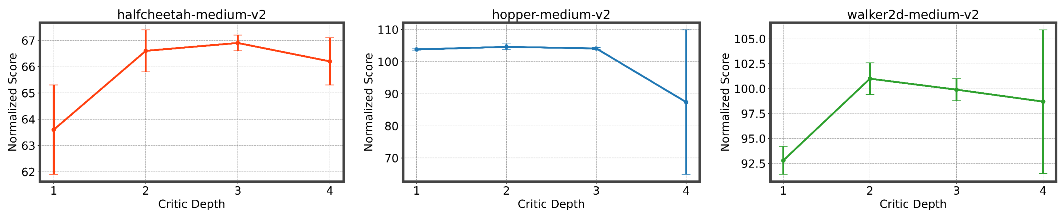

Figure 5: Performance of TD3 with FANS by varying depth for the critic network.

B.3 Learning Curves

The learning curves of TD3 + FANS for all tasks corresponding to Table 2 in the main text are shown in the figure below.

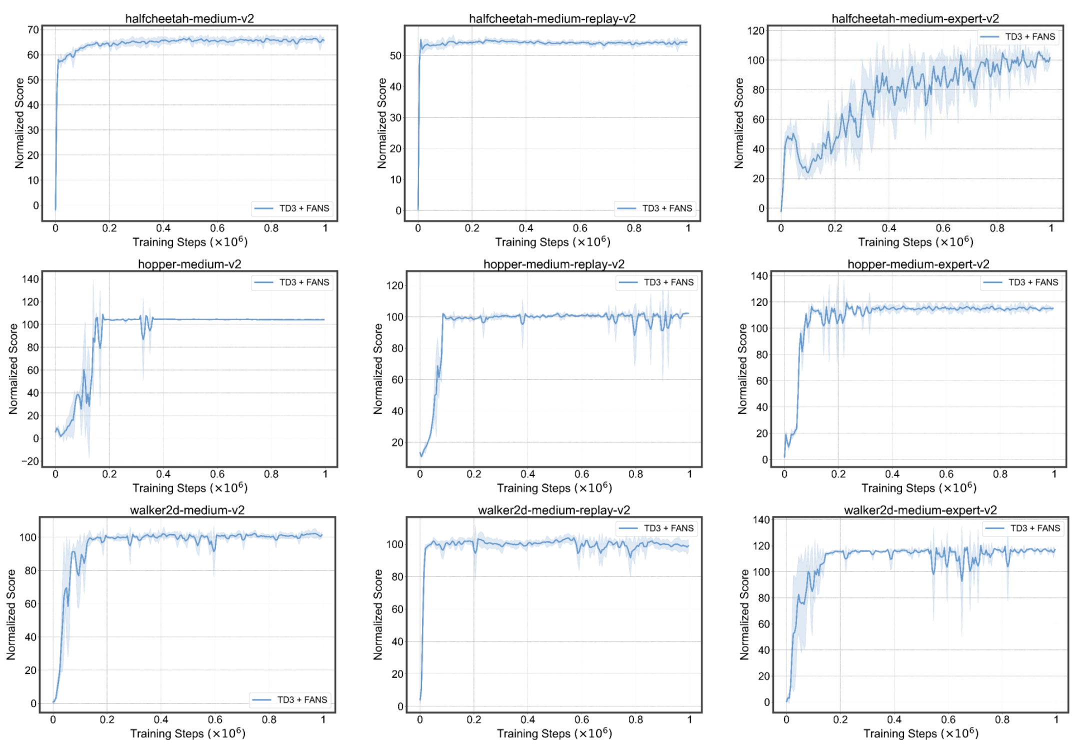

Figure 6: Performance of TD3 with FANS on different mujoco tasks.