

Hyper-parameter		MADDPG	IDDQN	NDDQN	Proposed method
	Q-network	-	(512,256)	(512,256)	(512,256)
	Encoder	-	-	-	(256,128)
	Decoder	-	-	-	(256,128)
	$\lambda_{\text{KL}}$	-	-	-	0.1
	$\lambda_{\text{rec}}$	-	-	-	0.5
	Actor network	(512,256)	-	-	-
	Critic network	(512,256)	-	-	-
				Shared	
	Optimizer				Adam
	Activation				ReLU
	Learning rate				$1 \times 10^{-4}$
	Discount factor				0.98
	Batch size				256
	Soft update rate				0.005
271-node	Training steps				$1 \times 10^4$
	Replay buffer				$2 \times 10^3$
899-node	Training steps				$2 \times 10^4$
	Replay buffer				$3 \times 10^3$