

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