Hierarchical VAE Based Semantic Communications for POMDP Tasks

Paper ID: 2730

Dezhao Chen[1] and Wenhui Hua[2]

[1] School of Informatics, Xiamen University, Xiamen, China

[2] School of Electronic Science and Engineering(National Model Microelectronics College), Xiamen University, Xiamen, China

INTRODUCTION

Background

Semantic communication, a potential 6G technology, focuses on efficiently transmitting compact information for tasks without compromising performance, even with limited bandwidth. This is especially beneficial for dynamic tasks like remote control.

Semantic communication for POMDP tasks

- Rich semantics in observation and action spaces.
- Offering diverse control scenarios.
- Delivering a well-defined semantic objective that is closer to the fundamental challenge of semantic emergence.

Main challenges

- The extraction of decision semantic information.
- The trade-off between sufficiency and compactness.
- Low sample efficiency and poor generalization.

METHODOLOGY

System model

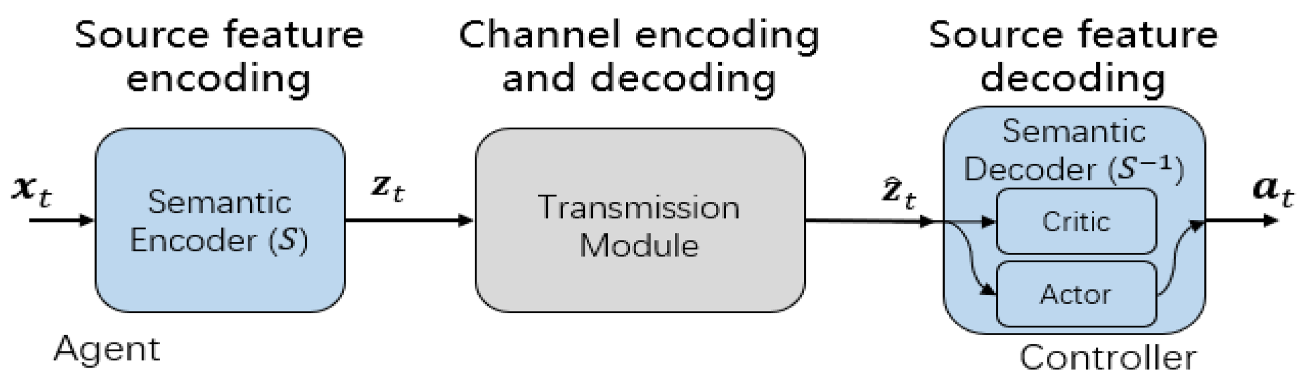

Fig. 1 Semantic communication system for remote control tasks.

Hierarchical Feature Encoding Based on HVAE

- Probability model
  - Encoder: $q(z_1, z_2, \ldots, z_n|o) = q(z_1 \mid o)q(z_2 \mid o)\ldots q(z_n \mid o)$
  - Decoder: $p(o, z_1, \ldots, z_n) = p(z_n) \prod_{i=1}^{n-1} p(z_i \mid z_{i+1})$
- Optimization objective:
  - $$\mathcal{L}_{\text{HVAE'}} = \mathbb{E}_{q_\theta(z_1|o)}[\log p_\theta(o \mid z_1)] - \mathcal{D}_{\text{KL}}[q_\phi(z_1, \ldots, z_n \mid o) \parallel p_\theta(z_1, \ldots, z_n)]$$
  - $$\mathcal{D}_{\text{KL}}[p\|q] = \mathcal{D}_{\text{KL}}(q_\phi(z_n|o)\|p(z_n)) + \mathbb{E}_{q_\phi(z_n|o)}[\mathcal{D}_{\text{KL}}(q_\phi(z_{n-1}|o) \parallel p_\theta(z_{n-1}|z_n))] + \ldots + \mathbb{E}_{q_\phi(z_2|o)}[\mathcal{D}_{\text{KL}}(q_\phi(z_1|o) \parallel p_\theta(z_1|z_2))]$$

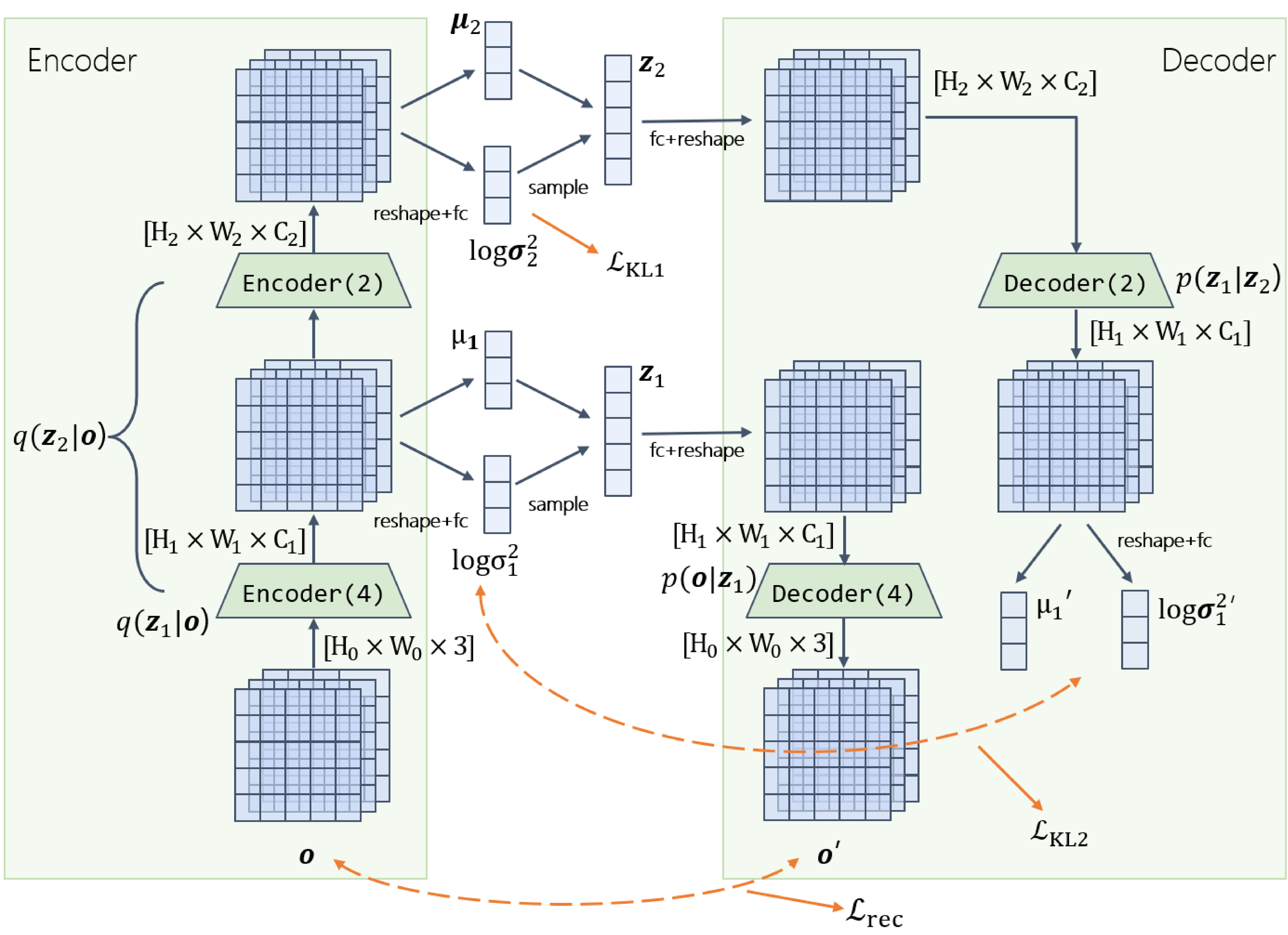

Fig. 2 The implementation of HVAE-2.

Hierarchical VAE for RL

- Hierarchical coding
  - z_1: sufficient feature, encode enough information of obs.
  - z_2: compact feature, encode task-relevant information.
- Concatenate z_1 and z_2 as the transmission feature.
- Using reinforcement learning training to approximately optimize the first term of the KL objective.

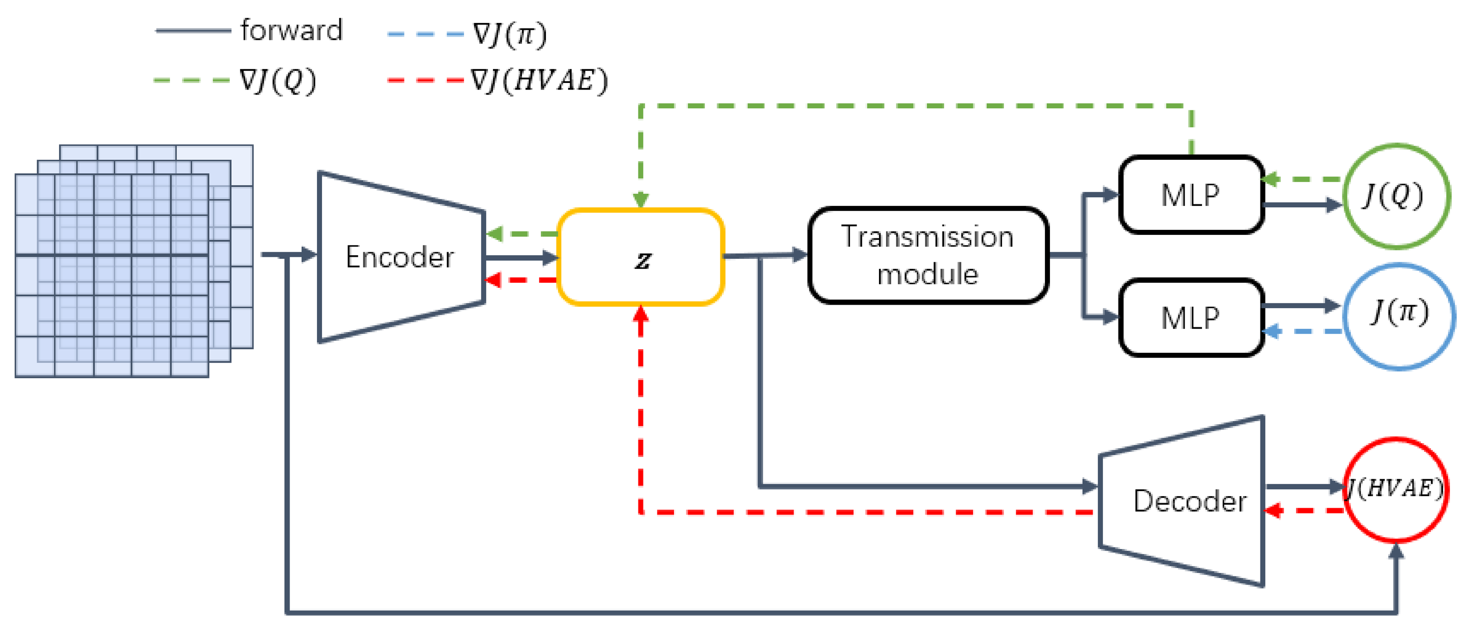

Fig. 3 The implementation of HVRL.

EXPERIMENT

Settings

- Environment: Cartpole swingup of DeepMind Control Suite.
- Metrics: cumulative reward and MSE of state prediction.

Results and discussion

Predictor	VRL	HVRL	z1	z2
Linear	0.8426(4)	0.0102(1)	0.0718(3)	0.0249(2)
Non-linear	0.8124(4)	0.0016(1)	0.0017(2)	0.0043(3)

Table 2. The predictor performance of different methods.

- The proposed method achieves a good task performance and improves the sampling efficiency.
- z_1 has low sampling efficiency; and z_2 is a good representation of the state, but it cannot complete the task well.

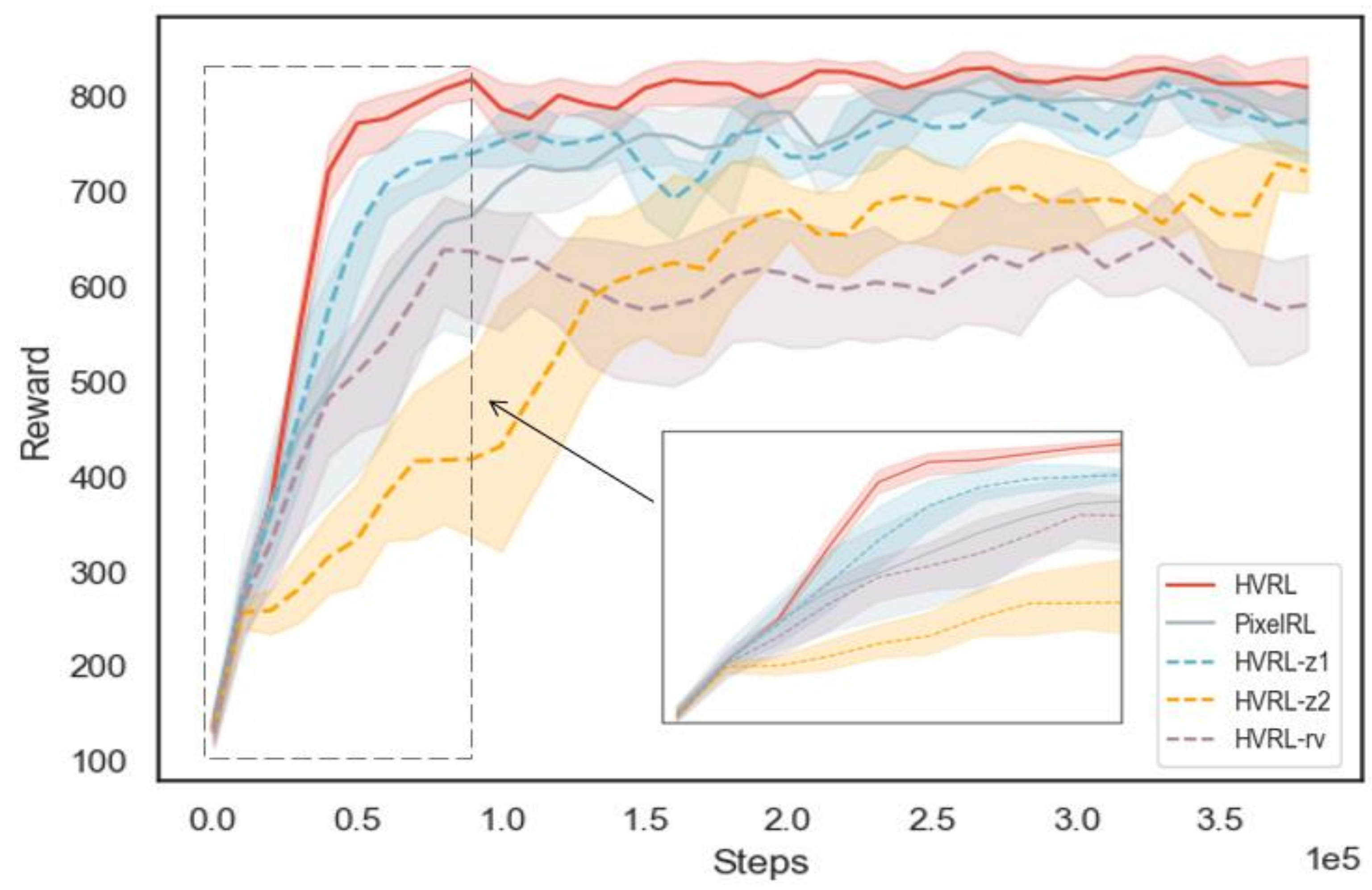

Fig. 4 Performance of different methods on cartpole-swingup environment.

CONCLUSION

- **HVRL**, a VAE-based hierarchical encoding technique for RL.
- Balancing the **trade-off** between **compactness** and **sufficiency**.
- Reducing the **representation gap** between features and ground-truth endogenous states.
- Achieving desirable **sample efficiency** and **control performance**.