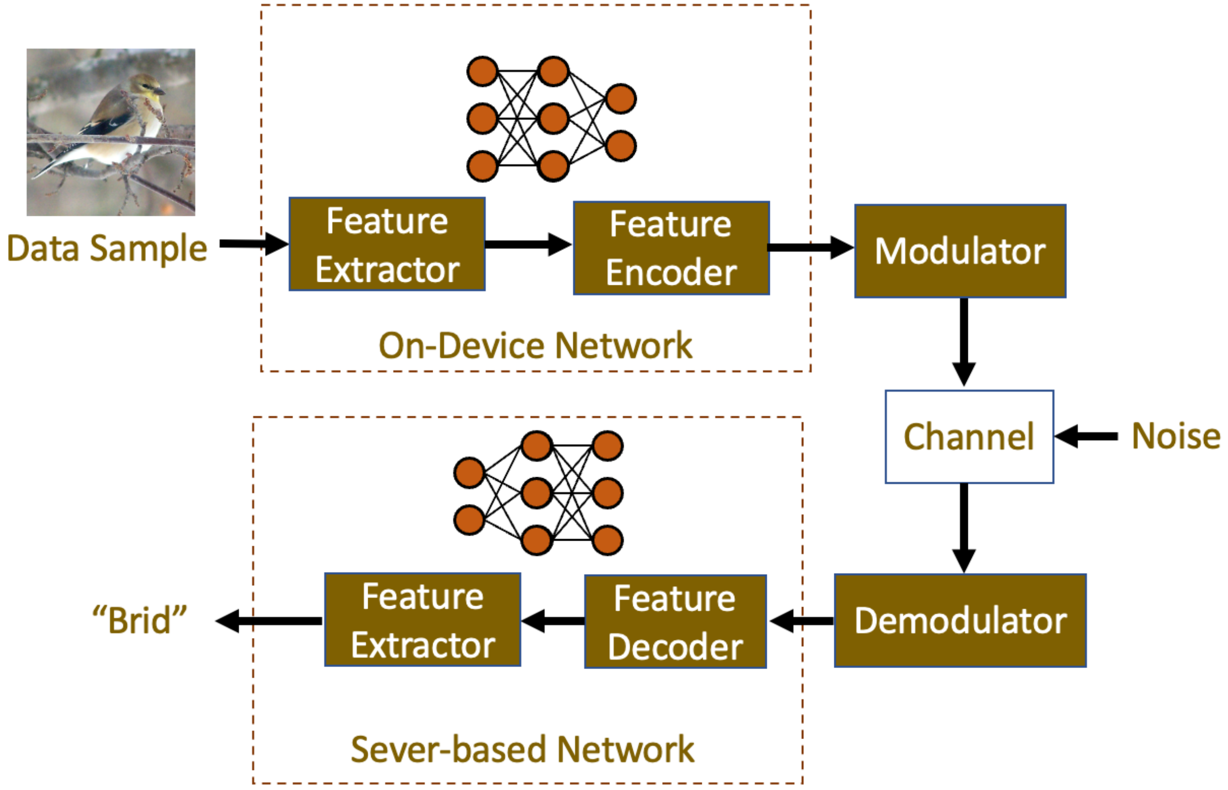
Robust Information Bottleneck for Task-Oriented Communication with Digital Modulation

Songjie Xie, Youlong Wu, Shuai Ma, Ming Ding, Yuanming Shi, and Mingjian Tang

Jan 17, 2023

Task-Oriented Communication

- To transmit informative features for downstream inference



- Current key techniques:

Feature extraction via Information Bottleneck[1]

End-to-end optimization via deep learning[2]

1: Shao, J., Mao, Y., & Zhang, J. (2021). Learning task-oriented communication for edge inference: An information bottleneck approach. IEEE Journal on Selected Areas in Communications, 40(1), 197-211.
2: Bourtsoulatze, E., Kurka, D. B., & Gündüz, D. (2019). Deep joint source-channel coding for wireless image transmission. IEEE Transactions on Cognitive Communications and Networking, 5(3), 567-579.

Issues and Our Solutions

Feature extraction via Information Bottleneck

Redundancy reduction

Degrading the robustness of communication

New principle: Robust IB

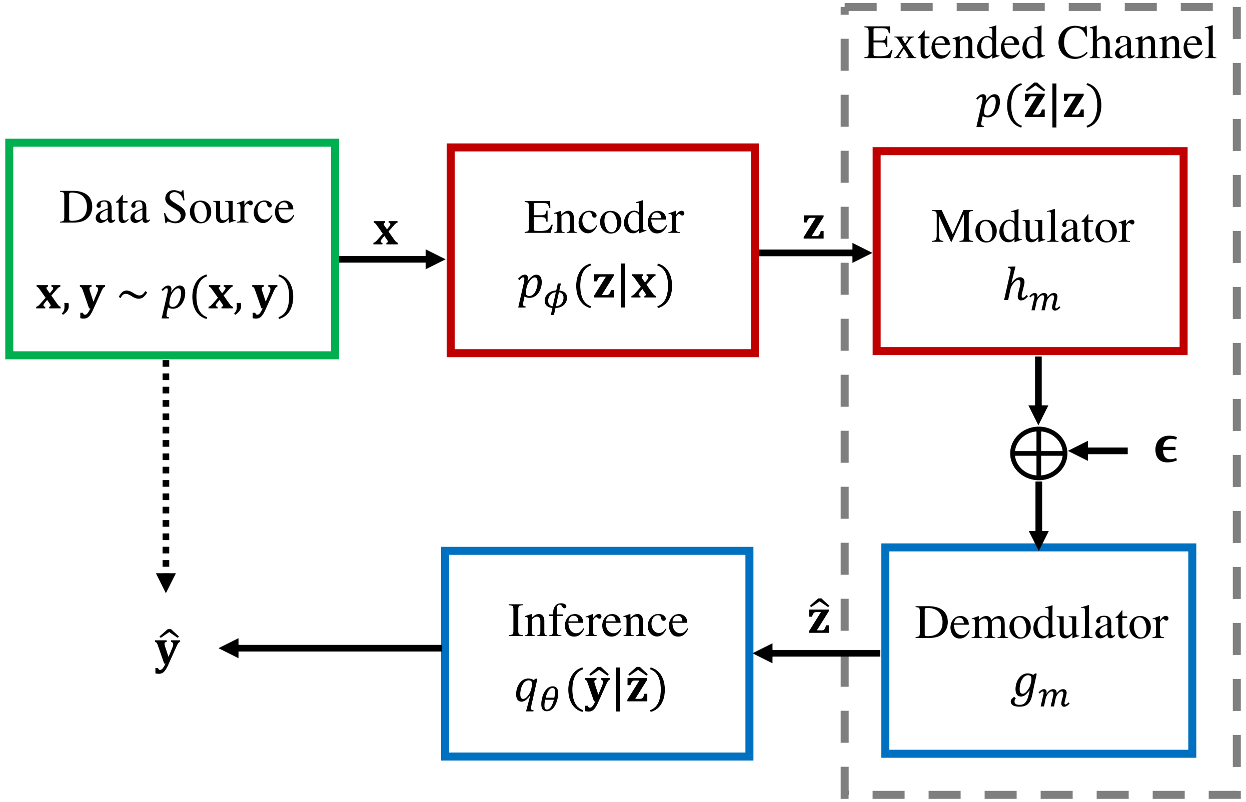
End-to-end optimization via deep learning

Floating point output value

Incompatibility with digital communiaction

New model: Discrete representation learning

System Model



Encoding:

$$Y \rightarrow X \xrightarrow{\phi} Z \rightarrow \hat{Z}$$

Optimal inference:

$$p_\phi(\mathbf{y}|\hat{\mathbf{z}}) = \frac{\int p(\mathbf{x}, \mathbf{y}) p_\phi(\hat{\mathbf{z}}|\mathbf{y}) d\mathbf{x}}{p_\phi(\hat{\mathbf{z}})}$$

Variational approximation for inference:

$$p_\phi(\mathbf{y}|\hat{\mathbf{z}}) \longrightarrow q_{\boldsymbol{\theta}}(\mathbf{y}|\hat{\mathbf{z}})$$

Informativeness-Robustness Tradeoff

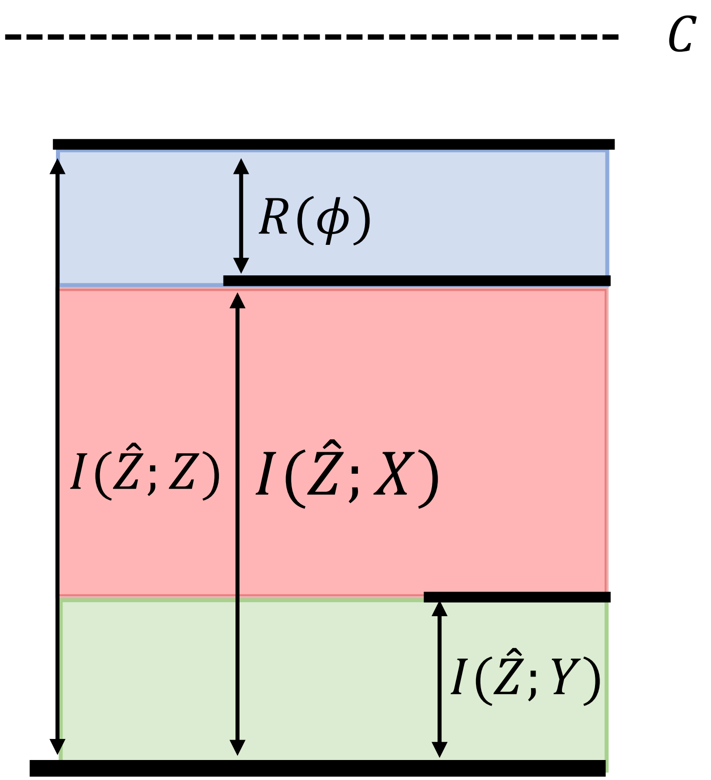
$$Y \to X \xrightarrow{\phi} Z \to \hat{Z}$$

- Task-relevant information: $I(Y;\hat{Z})$
- Task-irrelevant information: $I(X;\hat{Z}) - I(Y;\hat{Z})$
- Coded redundancy:

$$R(\phi) \triangleq I(Z;\hat{Z}) - I(X;\hat{Z})$$

- Channel capacity:

$$C \triangleq \max_{p_\phi(z)} I(Z;\hat{Z})$$

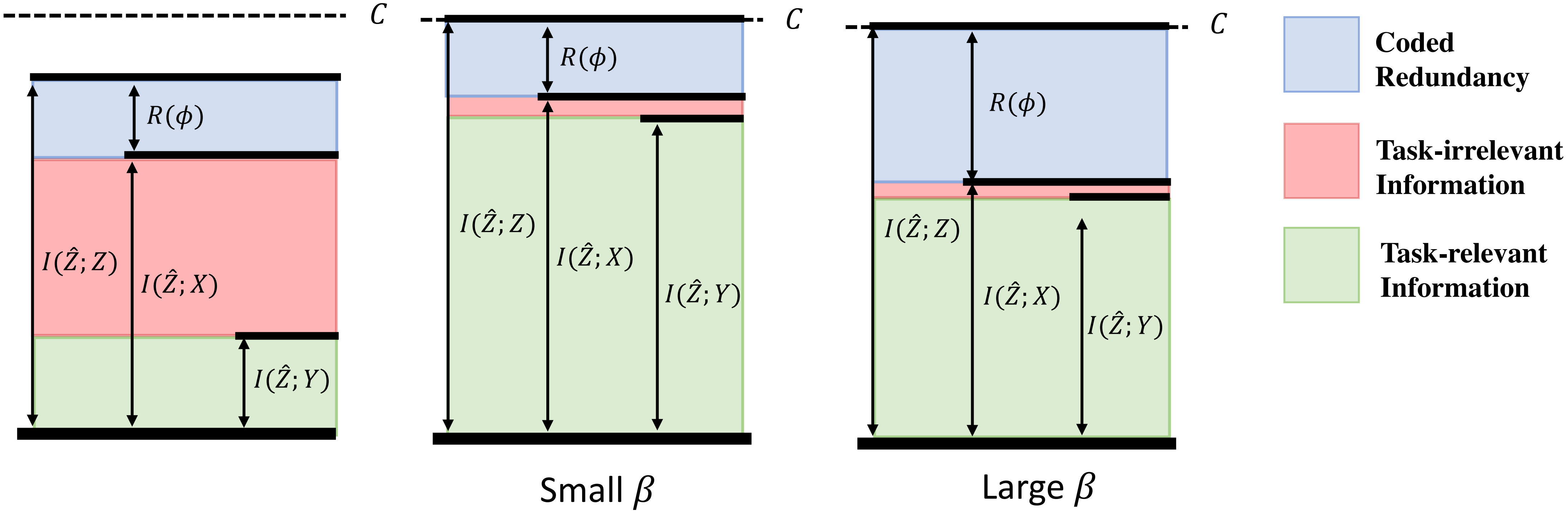


Data processing inequality:

$$I(Y;\hat{Z}) \leq I(X;\hat{Z}) \leq I(Z;\hat{Z}) \leq C$$

Robust Information Bottleneck (RIB)

$$\mathcal{L}_{\text{RIB}}(\phi) = -I(Y;\hat{Z}) - \beta[\underbrace{I(Z;\hat{Z}) - I(X;\hat{Z})}_{R(\phi)}]$$



Key Idea: *Keep minimal but sufficient task-relevant information and leave the rest redundancy utilized for robust encoding.*

Variational Encoding

Variational upper bound:

$$\begin{aligned}\mathcal{L}_{\mathrm{RIB}}(\boldsymbol{\phi}) \leq \mathcal{L}_{\mathrm{VRIB}}(\boldsymbol{\phi}, \boldsymbol{\theta}) &= \mathbb{E}_{p(\mathbf{x},\mathbf{y})}[\mathbb{E}_{p_{\boldsymbol{\phi}}(\hat{\mathbf{z}}|\mathbf{x})}[-\log q_{\boldsymbol{\theta}}(\mathbf{y}|\hat{\mathbf{z}})] \\ &\quad + \beta\mathbb{E}_{p_{\boldsymbol{\phi}}(\mathbf{z}|\mathbf{x})}[H_{\boldsymbol{\phi}}(\hat{Z}|\mathbf{z})] - \beta H_{\boldsymbol{\phi}}(\hat{Z}|\mathbf{x})]\end{aligned}$$

$q_\theta(y|\hat{z})$ is a variational distribution to approximate $p_\phi(y|\hat{z})$

Empirical estimation:

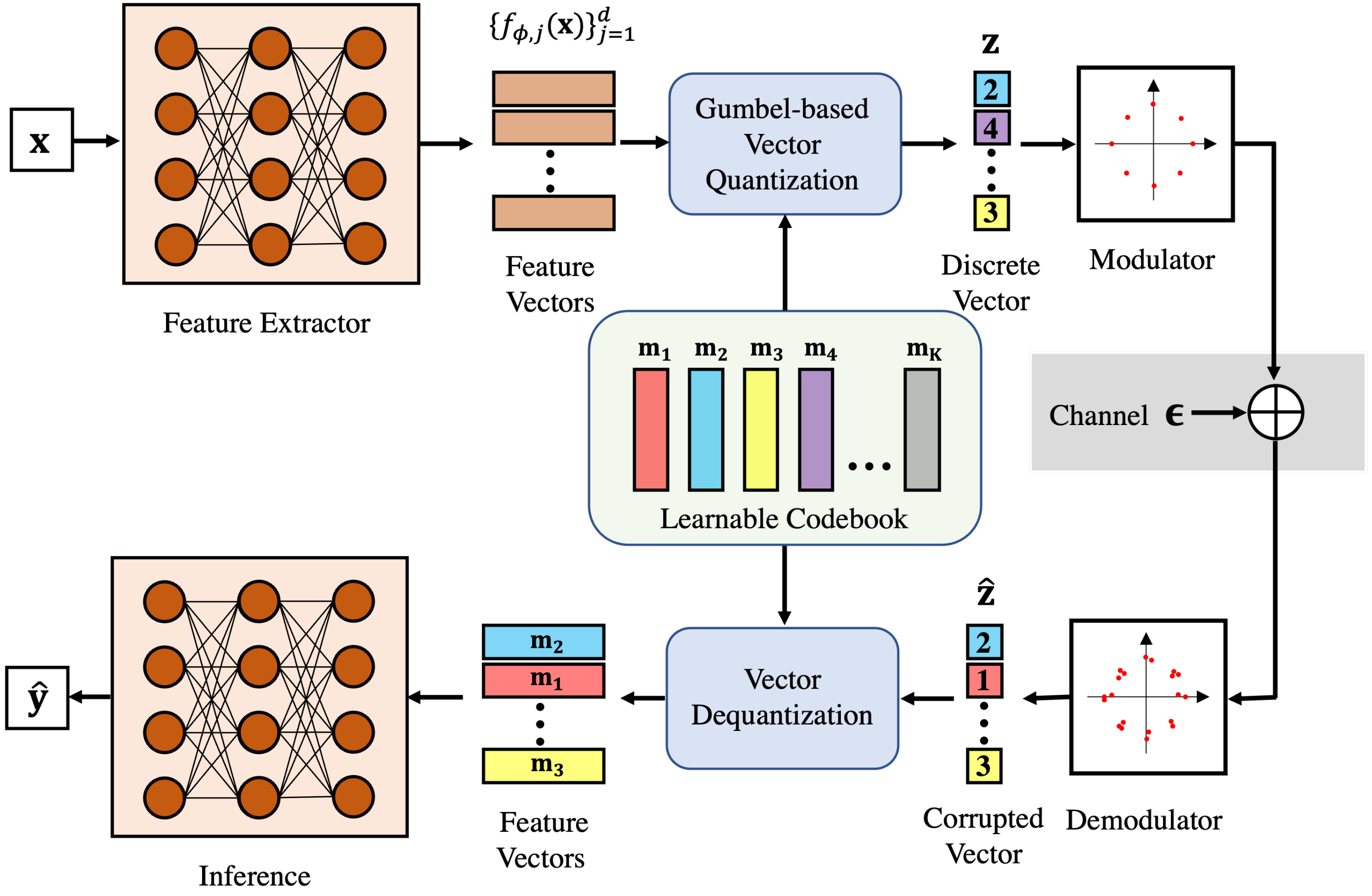
$$\begin{aligned}\tilde{\mathcal{L}}_{\mathrm{VRIB}}(\boldsymbol{\phi}, \boldsymbol{\theta}) = \frac{1}{N}\sum_{i=1}^{N}\{\frac{1}{L}\sum_{l=1}^{L}[-\log q_{\boldsymbol{\theta}}(\mathbf{y}^{(i)}|\hat{\mathbf{z}}^{(i,l)}) \\ + \beta\sum_{j=1}^{d} H(\hat{Z}_j|z_j^{(i,l)})] - \beta\sum_{j=1}^{d} H_{\boldsymbol{\phi}}(\hat{Z}_j|\mathbf{x}^{(i)})\},\end{aligned}$$

where $\hat{\mathbf{z}}^{(i,l)} = (\hat{z}_j^{(i,l)})_{j=1}^{d}$, $\hat{z}_j^{(i,l)} = g_{\mathrm{m}}(h_{\mathrm{m}}(z_j^{(i,l)}) + \epsilon_j^{(i,l)})$, $z_j^{(i,l)} \sim p_{\boldsymbol{\phi}}(z_j|\mathbf{x}^{(i)})$, and $\epsilon_j^{(i,l)} \sim \mathcal{CN}(0, \sigma^2)$.

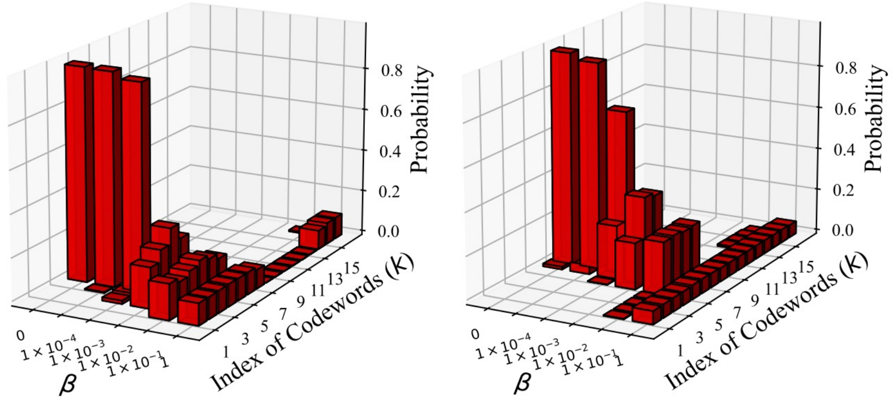
Advantages:

- No need to propose a variational prior $q(z)$ that conforms well to the aggregated posterior $p_\phi(z)$.
- z can be continuous or discrete.

Discrete Task-Oriented JSCC (DT-JSCC)



$$\mathcal{L}_{\text{RIB}}(\phi) = -I(Y;\hat{Z}) - \beta\underbrace{[I(Z;\hat{Z}) - I(X;\hat{Z})]}_{R(\phi)}$$



(a) (b)

Experimental Results

- **State-of-the-art baselines**: DeepJSCC, VFE
- Inference Performance on MNIST and CIFAR-10 classification tasks:

TABLE IV
THE INFERENCE ACCURACY OF EVALUATED METHODS FOR THE MNIST CLASSIFICATION TASK.

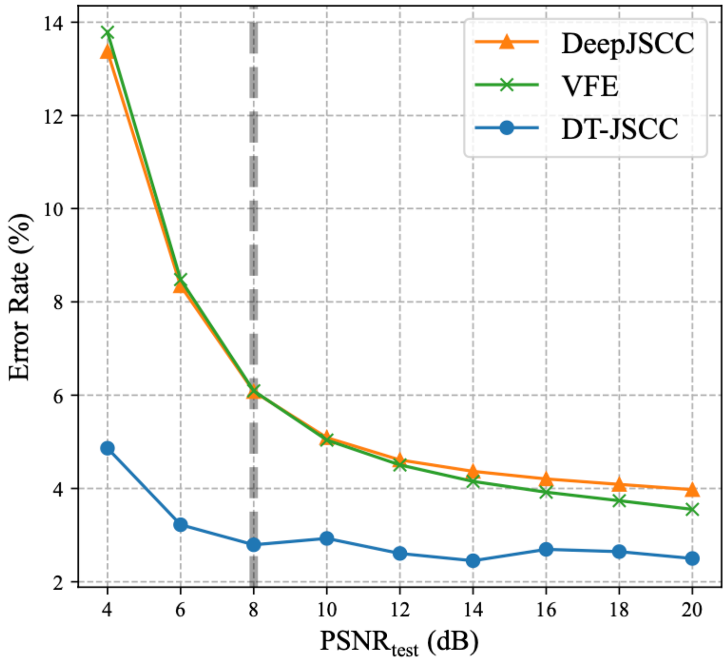
PSNR	4 dB	8 dB	12 dB	16 dB	20 dB
DeepJSCC	86.63	93.92	95.39	95.63	95.91
VFE	86.69	93.95	95.41	95.79	96.03
DT-JSCC	**96.66**	**97.21**	**97.25**	**97.72**	**97.93**

TABLE V
THE INFERENCE ACCURACY OF EVALUATED METHODS FOR THE CIFAR-10 CLASSIFICATION TASK.

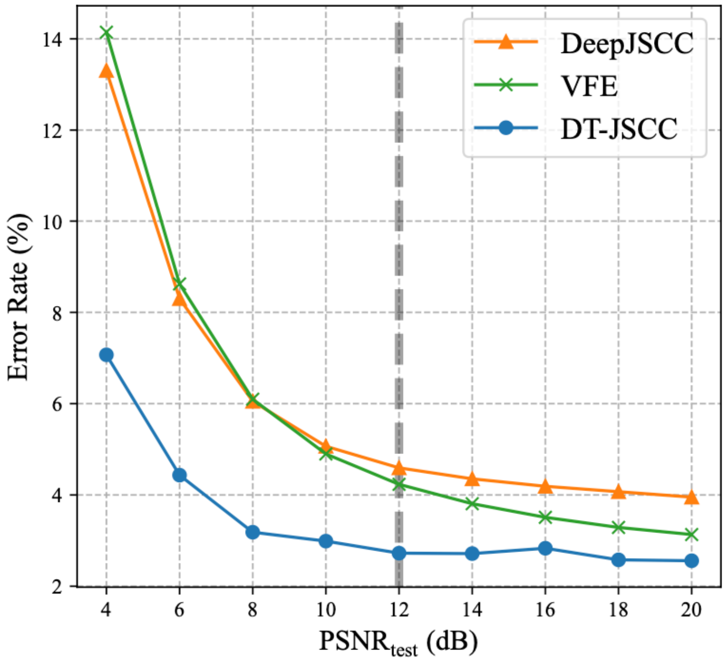
PSNR	4 dB	8 dB	12 dB	16 dB	20 dB
DeepJSCC	91.22	91.66	91.80	91.90	91.93
VFE	91.33	91.67	91.84	91.94	91.98
DT-JSCC	**91.46**	**91.93**	**91.91**	**92.26**	**92.14**

Experimental Results

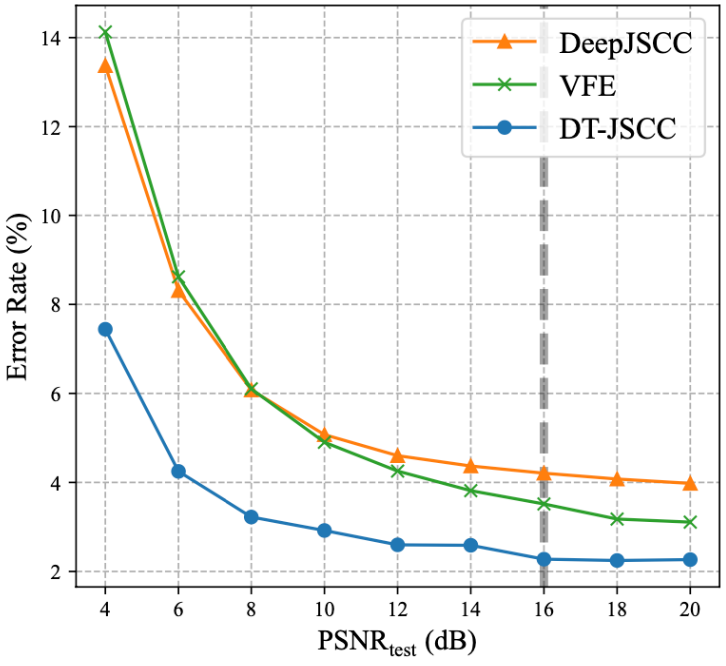
- Robustness of evaluated methods (MNIST):



(a) PSNR$_{train}$ = 8 dB



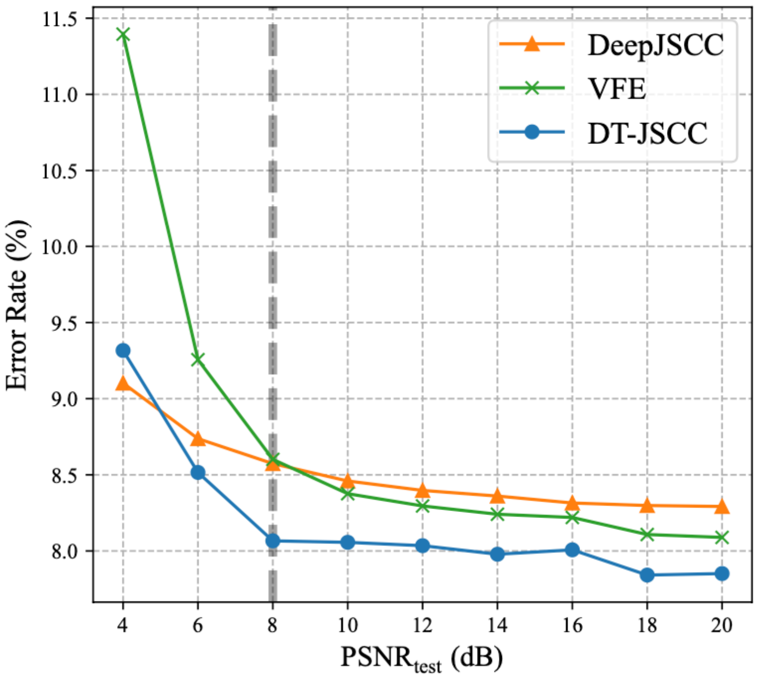
(b) PSNR$_{train}$ = 12 dB



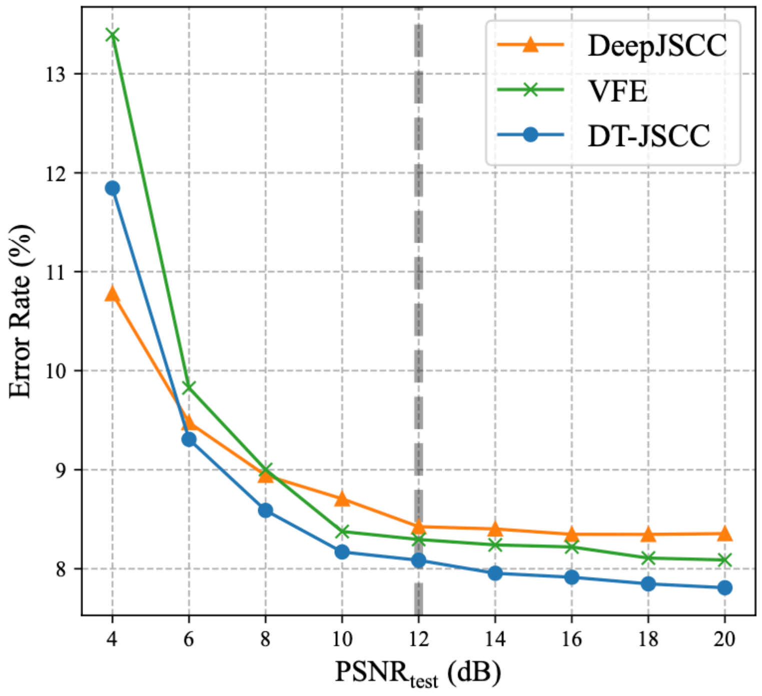
(c) PSNR$_{train}$ = 16 dB

Experimental Results

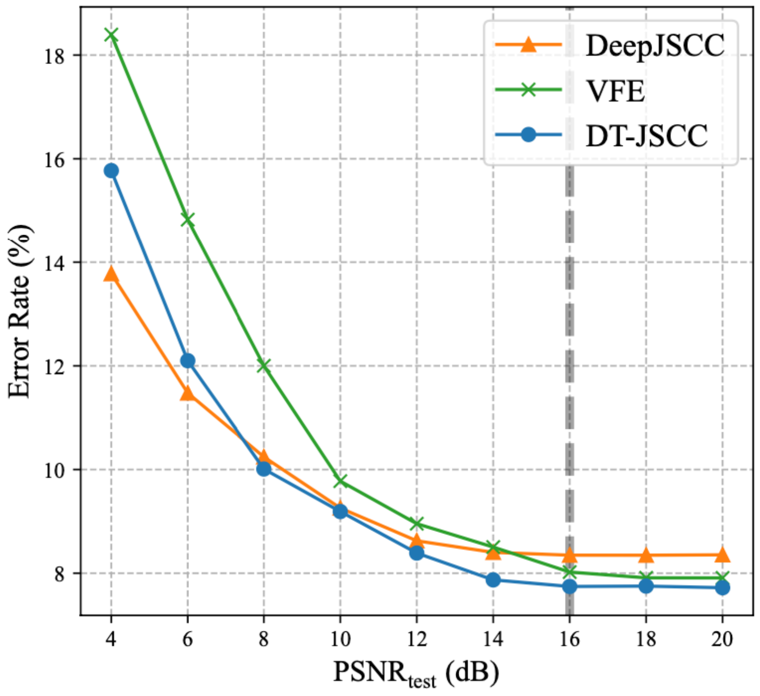
- Robustness of evaluated methods (CIFAR-10):



(a) $\mathrm{PSNR}_{\mathrm{train}} = 8$ dB



(b) $\mathrm{PSNR}_{\mathrm{train}} = 12$ dB

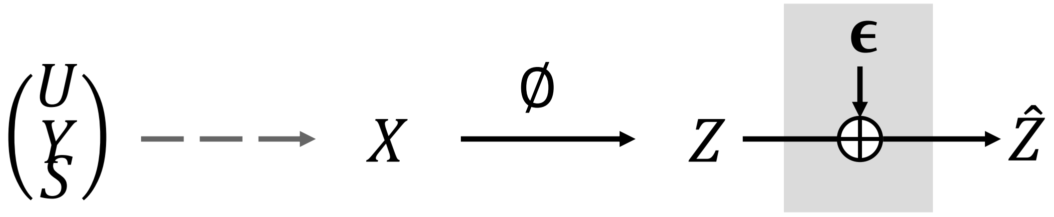


(c) $\mathrm{PSNR}_{\mathrm{train}} = 16$ dB

Conclusion

- Information Bottleneck needs more investigation
  - Generalization gap in amortized inference VS single-letter JSCC
  - Variational prior
- Case-by-case design for learning-based communication systems.
- There exists a connection between representation learning and communication system design.

Representation Learning (RL) and Communication



Research Topics:

- **From RL to Communication**:
  - Task-Oriented Communication
- **From Communication to RL**:
  - Privacy-preserving RL with a capacity-limited channel

Thanks!