Semantic Communications: Principles and Challenges

Huiqiang Xie, Zhijin Qin, Geoffrey Ye Li, Biing-Hwang Juang

Queen Mary University of London, UK
Imperial College London, UK
Georgia Institute of Technology, GA, USA

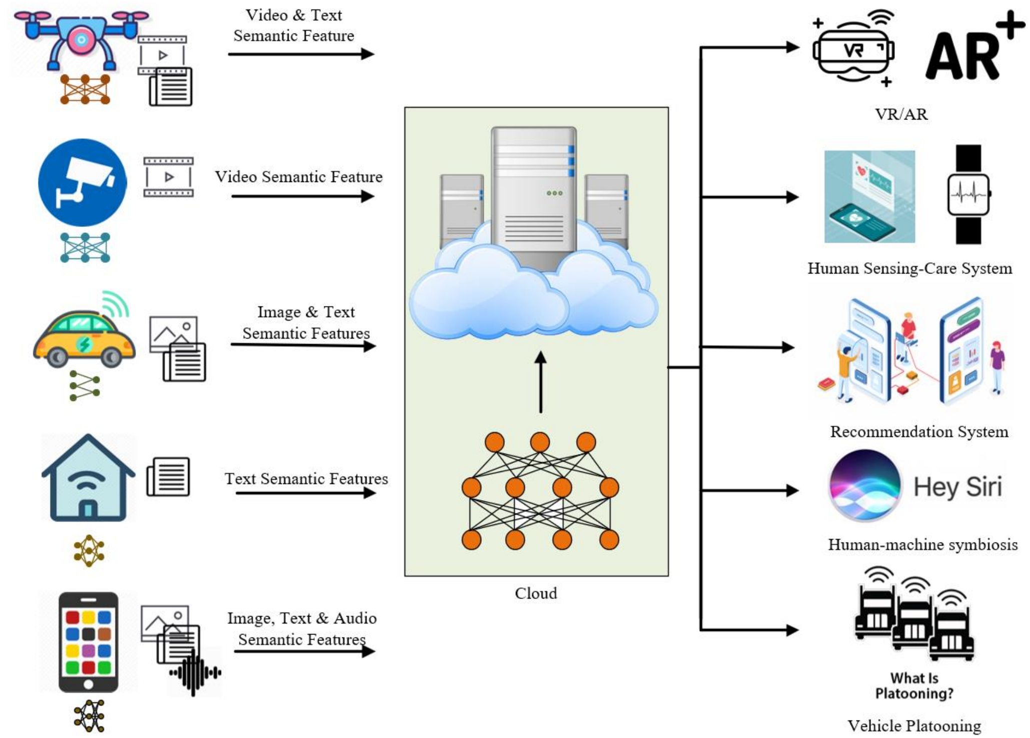
H. Xie, **Z. Qin**, G. Y. Li, and B.-H. Juang, "Deep learning enabled semantic communication systems," *IEEE TSP*, Apr. 2021.

Outline

- State-of-the-art
- Deep Learning enabled Semantic Communications
- DeepSC Variants
- Open Questions

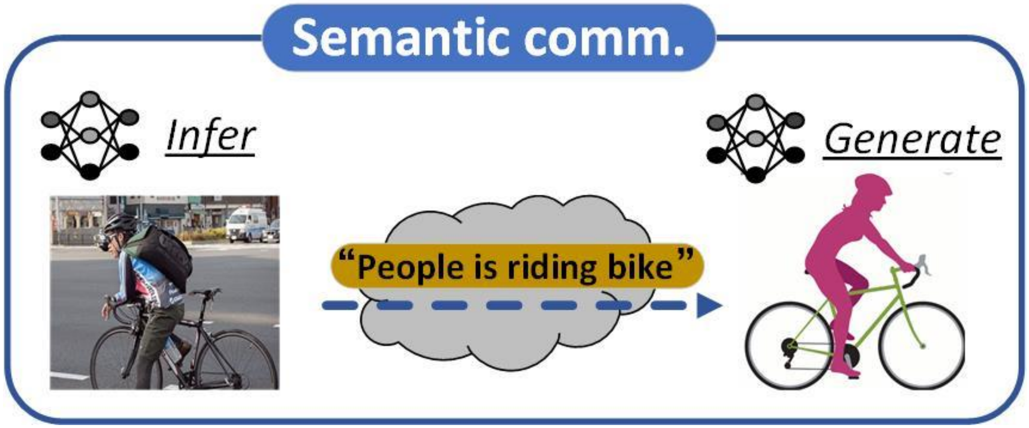
Semantic Communications in 6G and Beyond

- Research on 6G on the way
- Key role: intelligence transmission
  - Semantic communications
- Applications
  - Machine-to-machine communications
  - Human-to-machine communications
  - Human-to-human communications
- State-of-the-art: in its infancy
  - UKRI, NSF, ERC, NSFC
  - Huawei, Nokia, China Mobile



Conventional vs. Semantic Communications

- **Conventional communications**
  - A tube for accurate symbol transmission.
  - Regardless of content in source.
- **Semantic communications**
  - Transmitted symbols convey the desired meaning.
  - Transmitting semantic features relevant to **task** only.
  - Significantly improved transmission efficiency.



C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*, The University of Illinois Press, 1949.

Semantic Communications

STATE-OF-THE-ART

Bottlenecks

- **Initial semantic communication works**
  - Logic probability based semantic communication [1,2].
  - Word-level based semantic communication [3].
  - Cannot fully understand the meaning behind texts.

- **Derive semantic capacity of a discrete memoryless channel** [2]

$$C_s = \sup_{P(X|W)} \{I(X;Y) - H(W|X) + \overline{H_s(Y)}\}$$

  - I(X;Y): the mutual information.
  - $\overline{H_s(Y)}$: average logical information of received messages, the ability to interpret.
  - P(X | W): conditional probabilistic distribution of a semantic coding strategy.

[1] R. Carnap et al., "An outline of a theory of semantic information," Res. Lab. Electronics, Massachusetts Inst. Technol., Cambridge MA, Oct. 1952.
[2] J. Bao et. al, "Towards a theory of semantic communication," in IEEE Network Science Workshop, West Point, NY, USA, Jun. 2011.
[3] B. Guler et. al, "The semantic communication game," IEEE Trans. Cogn. Comm. Networking, vol. 4, no. 4, pp. 787– 802, Sep. 2018.

Power of Deep Learning

- **Deep learning enabled semantic communications**
  – **Avoid requirement on a general mathematical model.**
  – Provide strong capability of feature abstraction.
  – Improve communication system performance.

- **Challenges**
  – How to define the meaning behind bit sequences?
  – How to design metrics for semantic communications?
  – How to design systems at semantic level?

DeepSC and Variants

- **DeepSC** [1]
  - First work on deep learning enabled semantic communication.
  - Trained by maximizing **mutual information** and minimizing **semantic errors.**
- **DeepSC-S** [2]
  - Joint semantic-channel coding for speech transmission.
- **MU-DeepSC** [3]
  - Multi-user semantic communications.
  - Multimodal data transmission.
- **L-DeepSC** [4]
  - A lite DeepSC with smaller model size and low complexity.

[1] H. Xie, Z. Qin, G. Y. Li, and B.-H. Juang, "Deep learning enabled semantic communication systems," *IEEE TSP*, Apr. 2021. (Top 1 popular article of TSP)
[2] Z. Weng and Z. Qin, "Semantic communication systems for speech transmission," *IEEE JSAC*, Aug. 2021. (Popular Article of JSAC)
[3] H. Xie, Z. Qin, and G. Y. Li, "Task-oriented multi-user semantic communications for VQA", *IEEE Wireless Lett.*, 2022.
[4] H. Xie and Z. Qin, "A lite distributed semantic communication system for Internet of Things," *IEEE JSAC*, Jan. 2021. (Popular Article of JSAC)

Text Transmission

DEEPSC

System Model

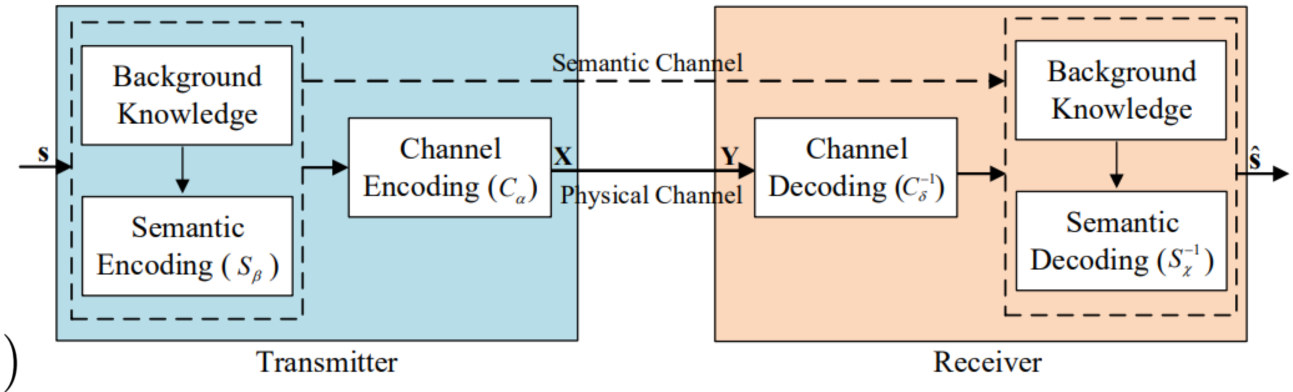
- **Transceiver**
  - Transmitter

$$\mathbf{X} = C_{\boldsymbol{\alpha}}\left(S_{\boldsymbol{\beta}}\left(\mathbf{s}\right)\right),$$

  - Receiver

$$\mathbf{Y} = \mathbf{HX} + \mathbf{N},$$

$$\hat{\mathbf{s}} = S_{\boldsymbol{\chi}}^{-1}\left(C_{\boldsymbol{\delta}}^{-1}\left(\mathbf{Y}\right)\right)$$



- **Channels**
  - Physical channel noise is caused by **physical channel impairment**

    i.e., AWGN, fading channels…
  - Semantic channel noise refers to **misunderstanding**

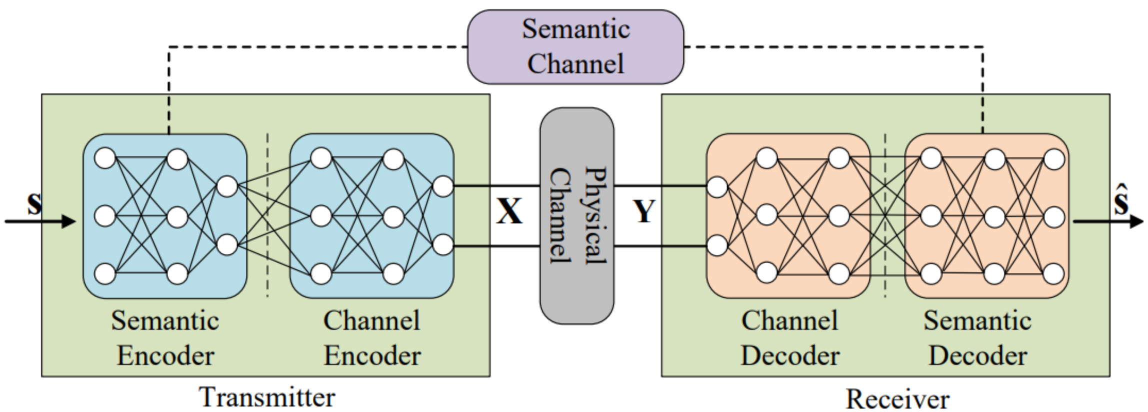
    i.e., interpretation errors and disturbance in estimated information.

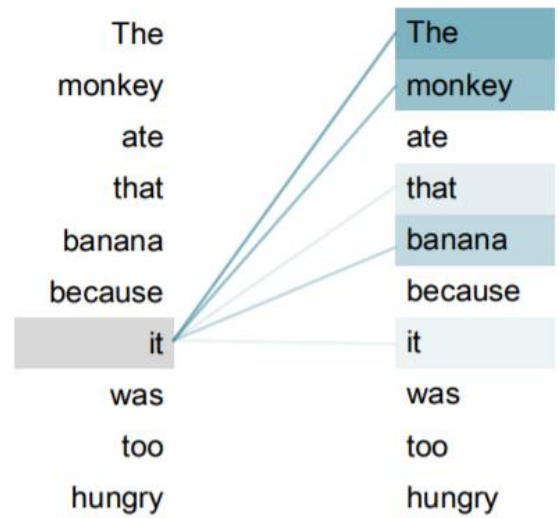
H. Xie, **Z. Qin**, G. Y. Li, and B.-H. Juang, "Deep learning enabled semantic communication systems," *IEEE TSP*, Apr. 2021. **(Most popular article of TSP)**
Code: https://github.com/HQXie0910/The-implementations-of-DeepSC

Proposed DeepSC-1

- **Transformer based semantic coding**
  - Merge conventional communications and semantic into neural networks.
  - Transformer [1] can learn the semantic in text.

    e.g., "it" completes pronoun reference "the animal".





[1] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," in *Proc. Advances Neural Info. Process. Systems (NIPS'17),* Long Beach, CA, USA. Dec. 2017, pp. 5998–6008.

Proposed DeepSC-2

- **Loss function design**
  - Total loss function to train the transceiver is

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{CE}}(\mathbf{s}, \hat{\mathbf{s}}; \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\chi}, \boldsymbol{\delta}) - \lambda \mathcal{L}_{\text{MI}}(\mathbf{x}, \mathbf{y}; T, \boldsymbol{\alpha}, \boldsymbol{\beta})$$

- **Cross-entropy**: Through reducing CE, the network can learn **syntax, phrase, the meaning of text.**

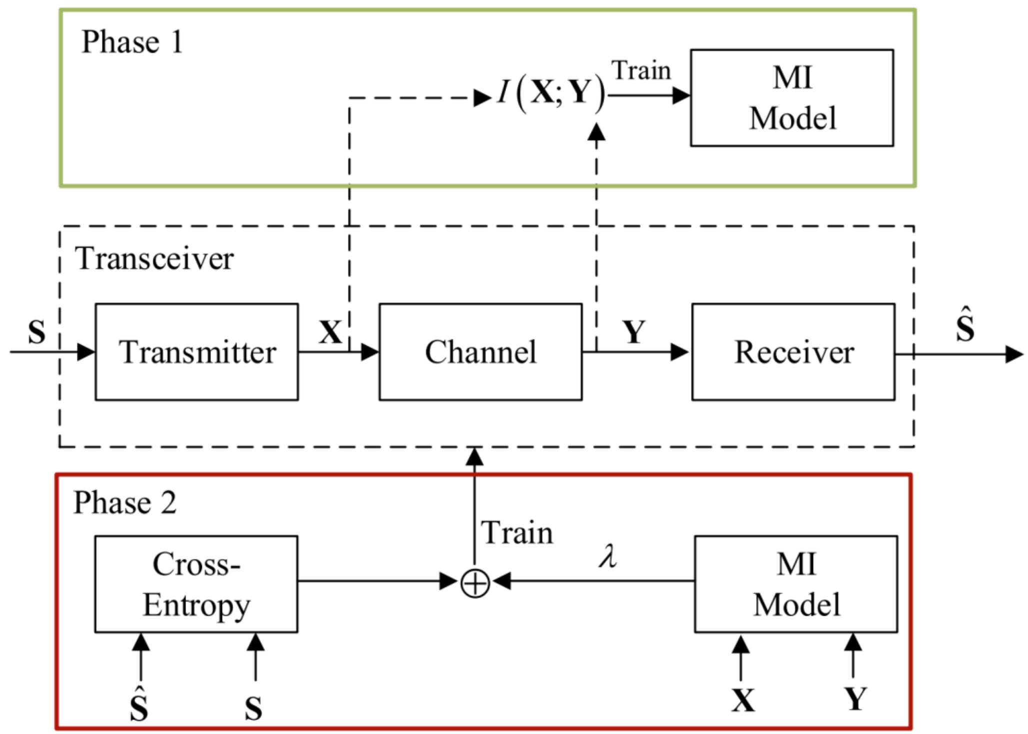
$$\mathcal{L}_{\text{CE}}(\mathbf{s}, \hat{\mathbf{s}}; \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\chi}, \boldsymbol{\delta}) = \\ -\sum_{i=1} q(w_i) \log(p(w_i)) + (1 - q(w_i)) \log(1 - p(w_i))$$

- **Mutual information:** **maximize** the data rate

$$\mathcal{L}_{\text{MI}}(\mathbf{X}, \mathbf{Y}; T) = \mathbb{E}_{p(x,y)}[f_T] - \log\left(\mathbb{E}_{p(x)p(y)}\left[e^{f_T}\right]\right)$$

Proposed DeepSC-3

- **Two-step training**
  - Maximize mutual information.
  - Train the whole model.



- **Performance gain**
  - Semantic encoder.
  - End-to-end communications.

Performance Metrics

- **BLEU score**
  - Compare the difference between words in two sentences

$$\log \mathrm{BLEU} = \min\left(1 - \frac{l_{\hat{\mathbf{s}}}}{l_{\mathbf{s}}}, 0\right) + \sum_{n=1}^{N} u_n \log p_n$$

  - ➢ $l_{\mathbf{s}}$ is the length of sentence **s**, $l_{\hat{\mathbf{s}}}$ is the length of sentence $\hat{\mathbf{s}}$.
  - ➢ p_n is the n-grams score, u_n is the weights of n-grams.

- **Sentence similarity**
  - Use **siamese network** to compute semantic similarity by:

$$\operatorname{match}(\hat{\mathbf{s}}, \mathbf{s}) = \frac{\boldsymbol{B}_{\boldsymbol{\Phi}}(\mathbf{s}) \cdot \boldsymbol{B}_{\boldsymbol{\Phi}}(\hat{\mathbf{s}})^{T}}{\left\|\boldsymbol{B}_{\boldsymbol{\Phi}}(\mathbf{s})\right\| \left\|\boldsymbol{B}_{\boldsymbol{\Phi}}(\hat{\mathbf{s}})\right\|}$$

  - ➢ $\mathbf{B}_{\mathbf{\Phi}}(\cdot)$ is the BERT model.
  - Sentence, **s**, will be mapped into **semantic vector space**, $\mathbf{B}_{\mathbf{\Phi}}(\mathbf{s})$, by BERT model.
  - Similarity is computed by measuring **distance** between $\mathbf{B}_{\mathbf{\Phi}}(\mathbf{s})$ and $\mathbf{B}_{\mathbf{\Phi}}(\hat{\mathbf{s}})$.

Simulation Setting

- **Dataset**
  - The proceedings of the European Parliament
- **The proposed network architecture**
  - Transmitter: 3 layers of Transformer encoder and 2 dense layers
  - Receiver: 2 dense layers and 3 layers of Transformer decoder
- **Benchmark**
  - Deep Learning based joint source-channel coding (DL based JSC coding)
  - Traditional methods
    - Source coding: Huffman coding and fixed-length coding
    - Channel coding: Reed Solomon and Turbo coding
    - Modulation: 64-QAM

Simulation Results-1

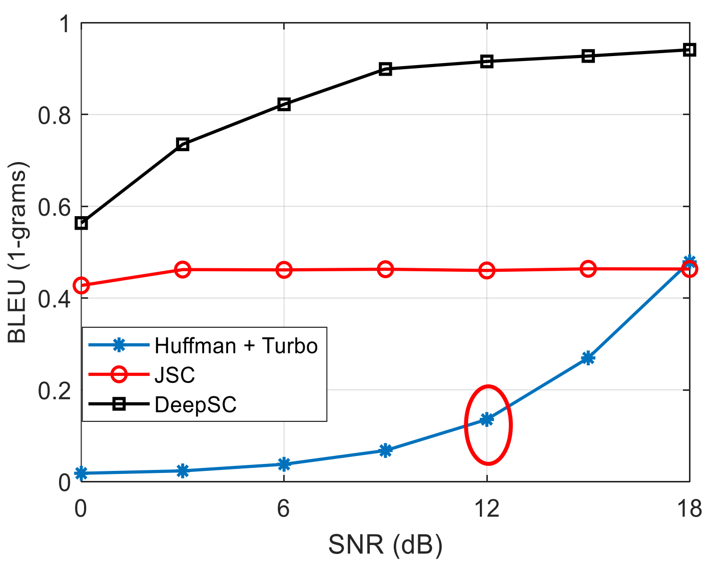
- ## Observations

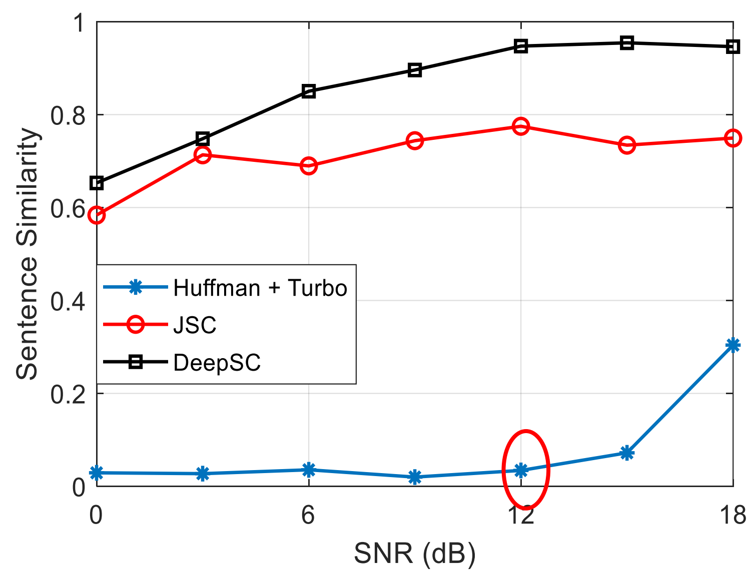
  – Deep learning enabled approaches are more competitive in **low SNR region**.

  – Tendency in sentence similarity is much closer to **human judgement**.

  ➢ SNR = 12 dB, **20%** BLEU score = **approximate 0** sentence similarity.

  ➢ People are **unable** to understand the meaning of **texts full of errors**.





Simulation Results-2

- **Recovered sentences comparison**
  – Rayleigh fading

THE SAMPLE SENTENCES BETWEEN DIFFERENT METHODS OVER RAYLEIGH FADING CHANNELS WHEN SNR IS 18 dB

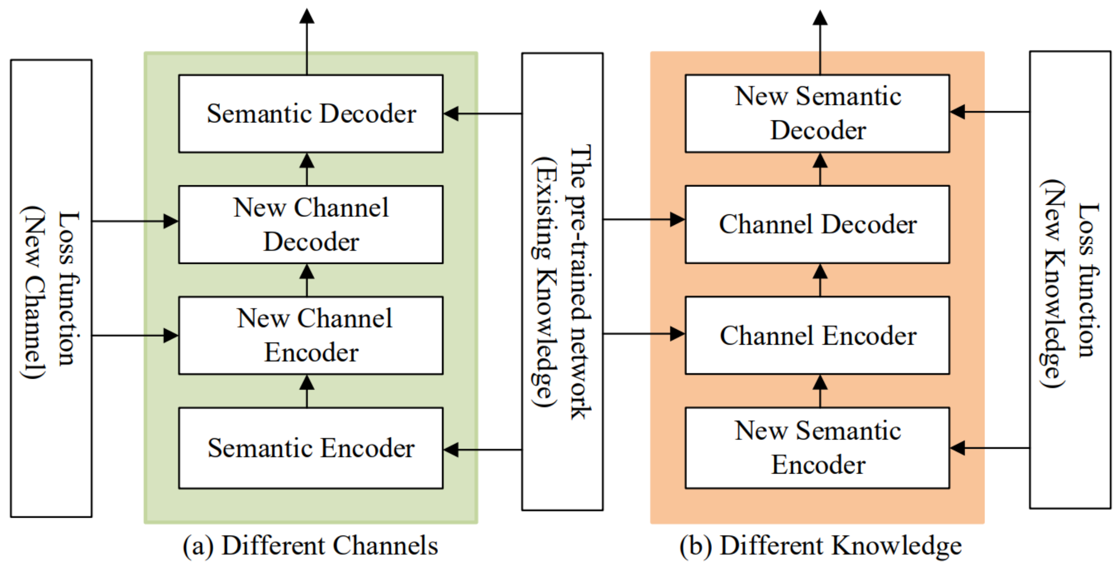
Transmitted sentence	it is an important step towards equal rights for all passengers.
DeepSC	it is an important step towards equal rights for all passengers.
JSCC-[22]	it is an essential way towards our principles for democracy.
Huffman + Turbo coding	rt is a imeomant step tomdrt equal rights for atp passurerrs.
Huffman + RS coding	it is an important step towards ewiral rlrsuo for all passengess.
Bit5 + Turbo coding	it is an yoportbnt ssep sowart euual qighd fkr ill passeneers.
Bit5 + RS coding	it iw an ymp!rdbnd stgo to!atds eq.al ryghts dkr alk passengers.

[22] N. Farsad *et. al*, "Deep learning for joint source-channel coding of text," *in Proc. IEEE ICASSP'18,* Calgary, AB, Canada, Apr. 2018, pp.2326–2330

Transfer Learning for Dynamic Environment

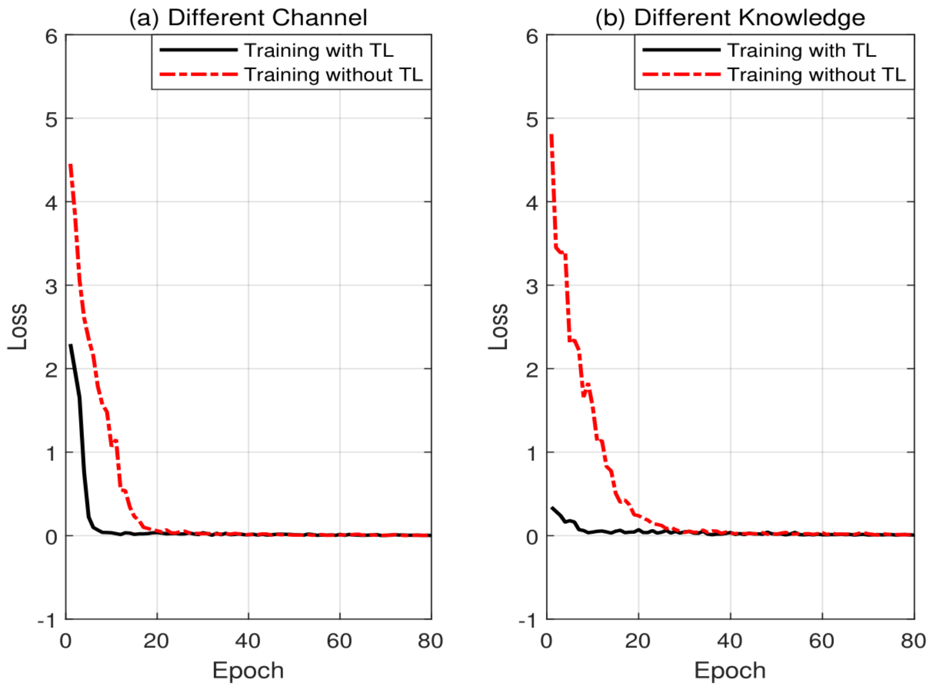
- **Core idea**
  - Store knowledge gained and apply it to a different but related problem.

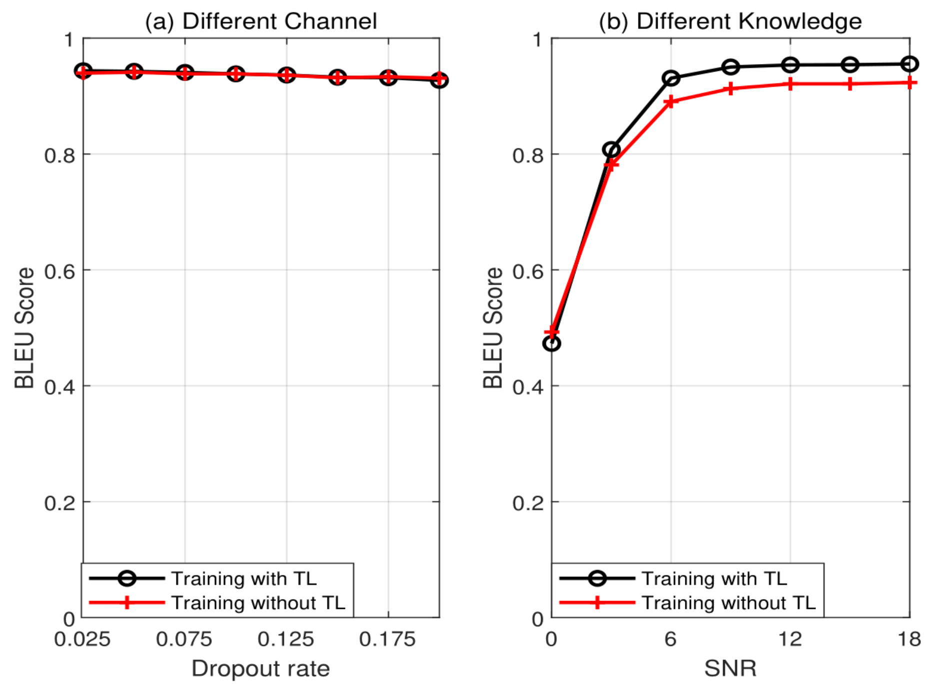
- **Two types of dynamic environment**
  - Different background knowledge
    - ➢ Freeze **channel encoder**
    - ➢ Update **semantic encoder**
  - Different channel conditions
    - ➢ Freeze **semantic encoder**
    - ➢ Update **channel encoder**



Simulation Results-3

- **Observations**
  - Transfer learning can accelerate the training with **fewer number of epochs.**
  - Transfer learning achieves **similar or higher performance.**



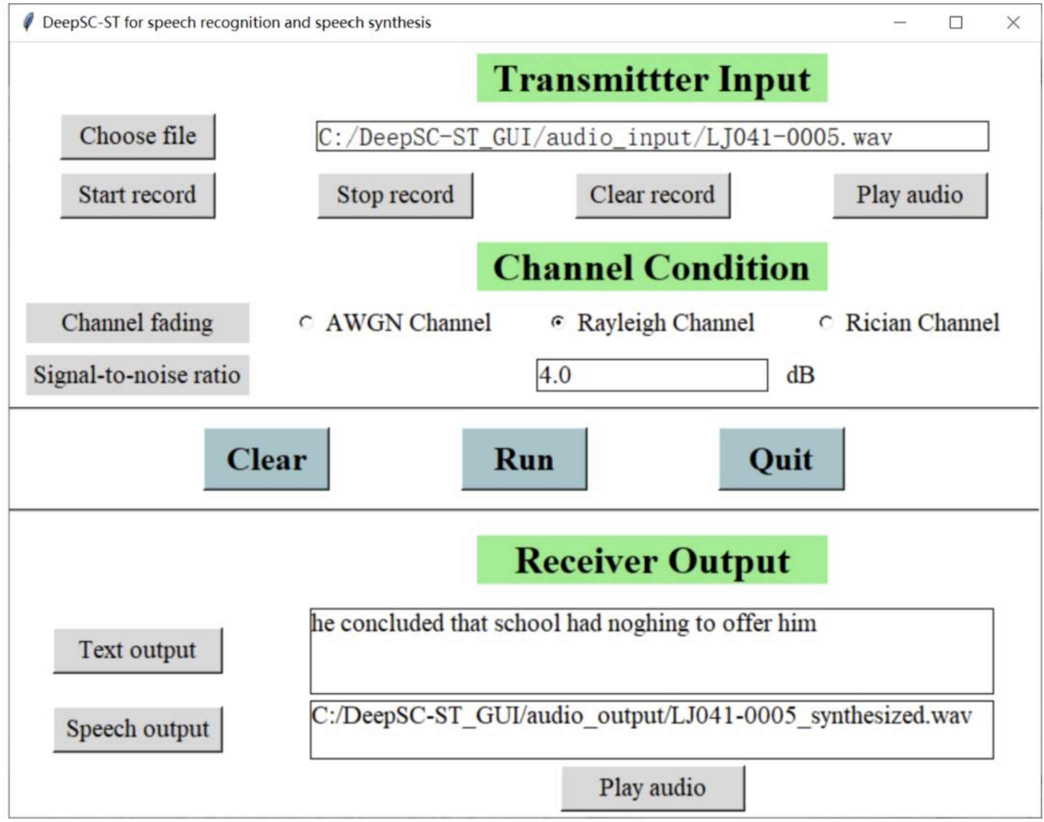


Variants

DEEPSC

DeepSC-ST: Speech Recognition and Synthesis

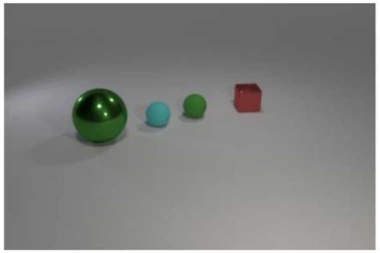
- **Challenges**
  - – Speech transmission.
  - – Task-oriented transmission.
  - – Separate text semantic from speech.
- **Novelty and contribution**
  - – Transmit text features only.
  - – Speech recognition & synthesis.
  - – Efficient speech synthesis.
- **Achieved performance**
  - – Outperform in low SNR regime.
  - – Lower network traffic by 92.1%.
  - – A demonstration for the proof-of-concept of DeepSC-ST.



- Z. Weng, and **Z. Qin**, "Semantic Communication Systems for Speech Transmission", *IEEE JSAC*, Jun. 2021.
- Z. Weng, **Z. Qin,** X. Tao, C. Pan, G. Liu and G. Y. Li, "Deep Learning Enabled Semantic Communications with Speech Recognition and Synthesis ", *Arxiv 2205.04603*, May. 2022.

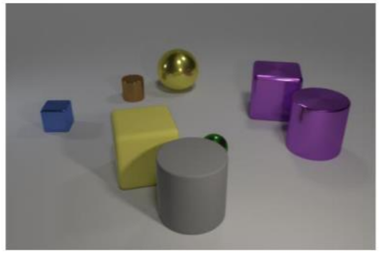
MU-DeepSC

- **Challenges**
  - – Multi-user scenario.
  - – Multimodal data transmission.
  - – Task: visual question answering.
- **Novelty and contribution**
  - – Transmit semantic features only.
  - – Semantic features separation at receiver.
  - – Output the answer directly.
- **Achieved performance**
  - – Computational complexity reduced by 80%.
  - – Transmitted symbols reduced by 70%.



Q: *Are there any other things that are the same shape as the small red shiny object?*

A (DeepSC-VQA): *no* ✓
A (JPEG+LDPC+UTF-8+Turbo): *yes* ✗



Q: *Is the color of the metal block that is right of the yellow rubber object the same as the large metal cylinder?*

A (DeepSC-VQA): *yes* ✓
A (JPEG+LDPC+UTF-8+Turbo): *no* ✗

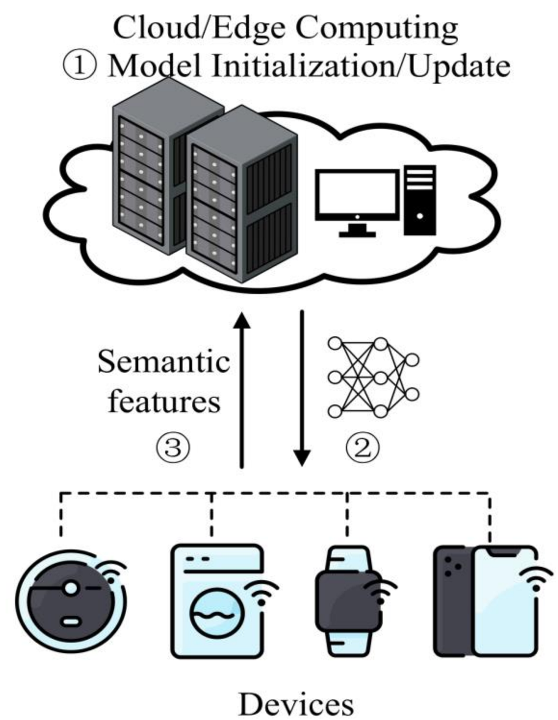
- H. Xie, **Z. Qin**, and G. Y. Li, "Task-oriented multi-user semantic communications for VQA", *IEEE Wireless Lett.*, 2021.

Lite DeepSC (L-DeepSC)

- ## Challenges
  - – Distributed power-constrained IoT devices.
  - – Model update communication costs.
  - – Finite constellation points.

- ## Novelty and contribution
  - – Prune non-essential model parameters by 90%.
  - – Quantize essential model parameters.

- ## Achieved performance
  - – 40x compression ratio without performance degradation.
  - – Reduced transmission data.
  - – Lowered power consumption at devices.



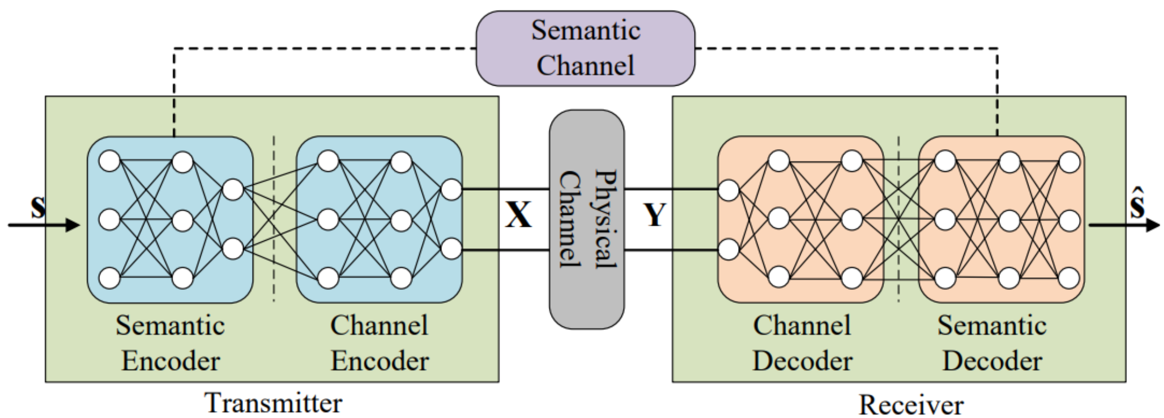
• H. Xie and **Z. Qin,** "A lite distributed semantic communication system for internet of things," *IEEE JSAC,* Jan. 2021 (**Popular Article of IEEE JSAC**).

Further Readings

- **A comprehensive survey**
  - Z. Qin, X. Tao, J. Lu, T. Weng, and G. Y. Li, "Semantic communications: principles and challenges", https://arxiv.org/abs/2112.10255, 2022.
- **Multi-user multimodal data**
  - H. Xie, Z. Qin, X. Tao, K. B. Letaief, "Task-oriented multi-user semantic communications", JSAC, 2022, accepted to appear.
- **Semantic communication robust to semantic noise**
  - Q. Hu, G. Zhang, Z. Qin, Y. Cai, G. Yu, "Robust semantic communications against semantic noise", https://arxiv.org/abs/2202.03338, 2022.
- **Semantic-aware network management**
  - L. Yan, Z. Qin, R. Zhang, Y. Li, and G. Y. Li, "Resource allocation for text semantic communications," IEEE Wireless Commun. Lett., 2022.
  - L. Yan, Z. Qin, R. Zhang, Y. Li, G. Y. Li, "Resource allocation for semantic-aware networks", https://arxiv.org/abs/2201.06023, 2022.

Summary

- **DeepSC and variants**
  - Support text, speech, and multimodal data transmission.
  - Extend from single user case to multi-user case.
- **Achieved performance**
  - Robust to low SNR region.
  - Significantly reduced transmission data size.
  - Lowered power consumption.



Thank you for the attention!

To be Solved

- **Theory-oriented challenge**
  - Semantic compression limit.
  - Semantic information representation.

- **Application-oriented challenge**
  - Knowledge base.
  - Transceiver for various tasks.
  - Transceiver for various sources.

- **Semantic-aware network management**
  - New formulation.
  - Semantic resource allocation.