Multi-View Images Over-the-Air Aggregation and Classification

Fan Yang

Electrical and Computer Engineering, University of Toronto

Motivation

- **Background:** In the context of modern muti-device sensing wireless networks, multi-view classification relying on the fusion of image features from multiple cameras is considered a promising objective detection technology.
- **Challenge:** Most recent methods deploy the AI models either on the edge devices or on the central server. However, on-device inference causes huge computation overhead especially for deep neural networks, while on-server inference causes huge communication overhead by transmitting the high-dimensional raw data.
- **Goal:** Develop an efficient multi-view classification approach that can be deployed in muti-device sensing wireless networks.

Related Work

- **MVCNN** (Multi-View Convolutional Neural Network) is a deep learning model that extracts features from multiple 2D views of a 3D object using CNNs, aggregates these features through methods like pooling, and performs classification [1,2].
- **Split inference** paradigm divides an AI model into two parts: one deployed on resource-limited devices for feature extraction, and the other at an edge server for completing the remaining computation-intensive inference task [3].
- **AirComp** (Over-the-air computation) exploits the waveform-superposition property of a multi-access channel to realize over-the-air aggregation of extracted features simultaneously transmitted by multiple devices [4].

References

[1] H. Su, S. Maji, E. Kalogerakis, and E. Learned-Miller, “Multi-view convolutional neural networks for 3D shape recognition,” in Proc. IEEE Int. Conf. Comput. Vision (ICCV), Santiago, Chile, Dec. 7–13 2015.

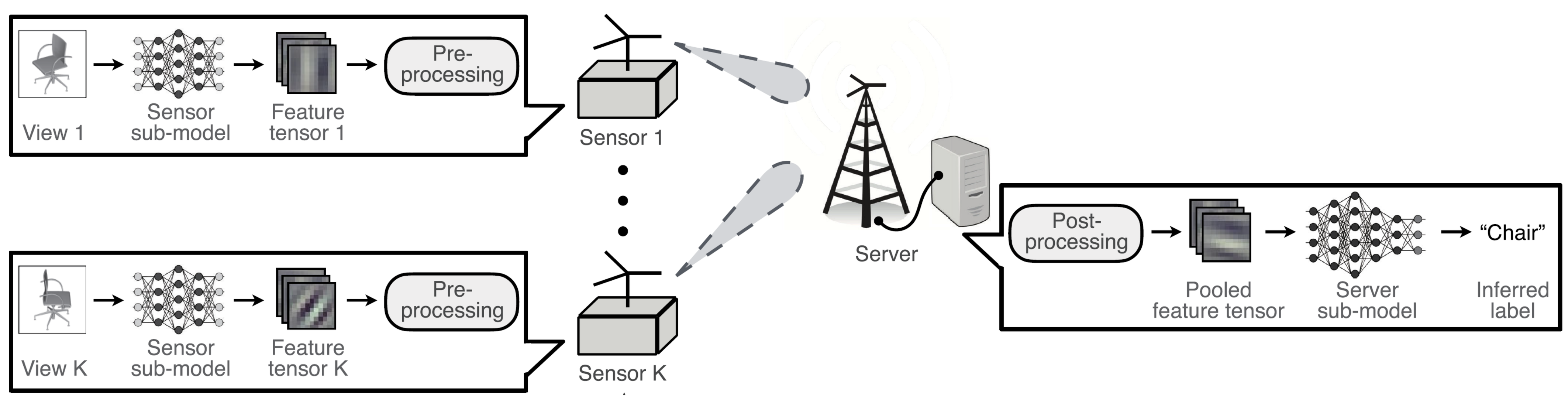
[2] J.-C. Su, M. Gadelha, R. Wang, and S. Maji, “A Deeper Look at 3D Shape Classifiers,” in Second Workshop on 3D Reconstruction Meets Semantics, ECCV, 2018.

[3] J. Shao, Y. Mao, and J. Zhang, “Task-oriented communication for multidevice cooperative edge inference,” IEEE Trans. Wireless Commun., vol. 22, no. 1, p. 73–87, Jan. 2023.

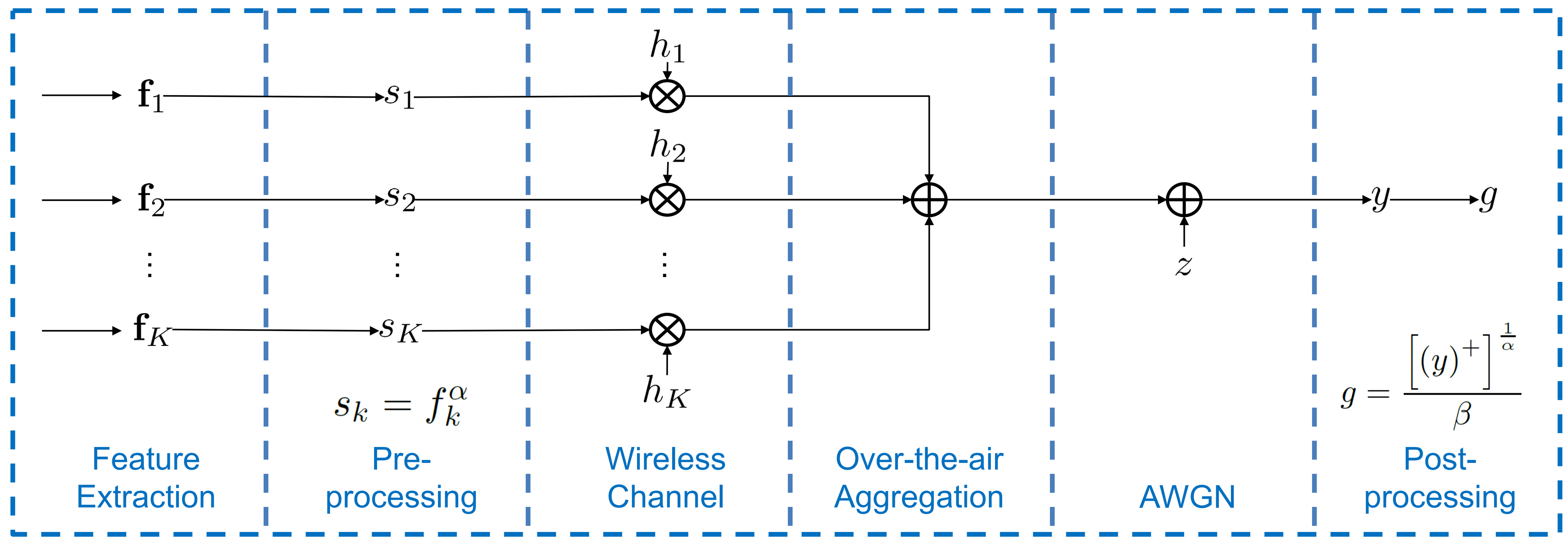
[4] X. Chen, K. B. Letaief, and K. Huang, “On the View-and-Channel Aggregation Gain in Integrated Sensing and Edge AI,” IEEE J. Sel. Areas Commun., pp. 1–1, 2024.

Method

- **System model:**



- **Signal Processing:**

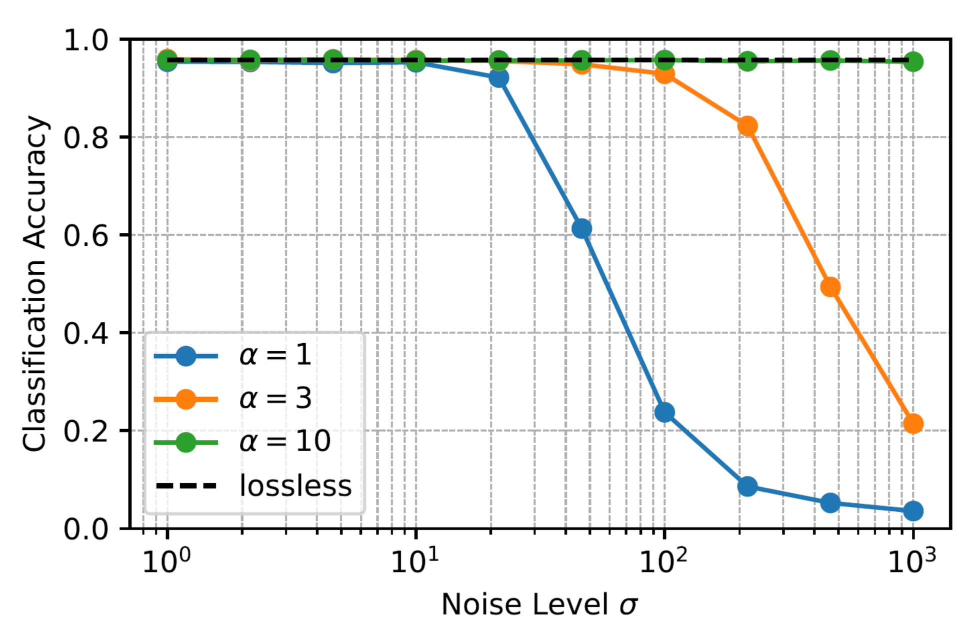


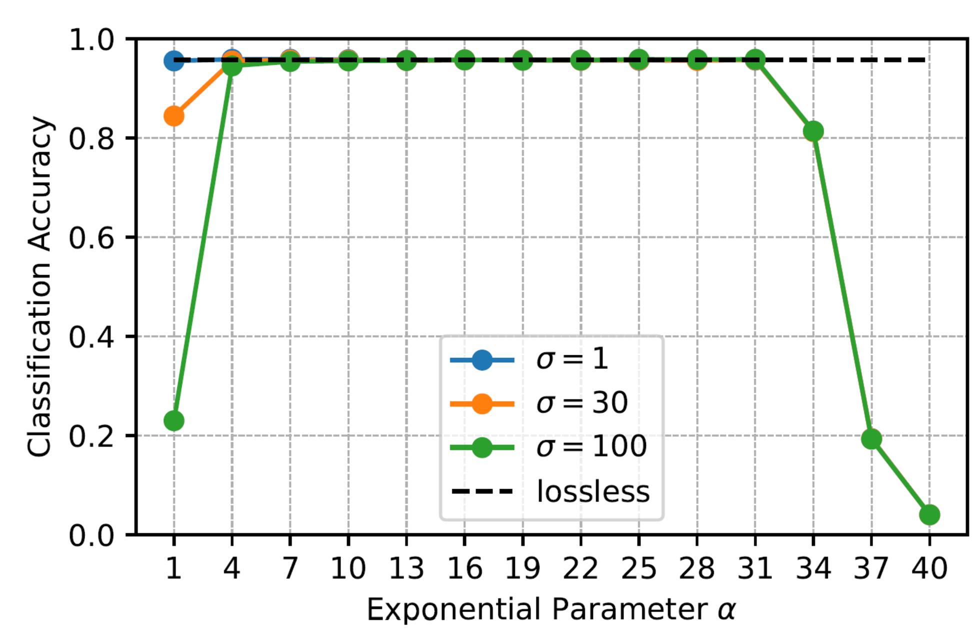
- **Steps:**

1. Pretrain the MVCNN model (RestNet-18) on a noise-free wireless channel.
2. Deploy the pretrained AI model on a noisy wireless channel.
3. Tune the AirComp factors α and β to implement average-pooling and max-pooling for multi-view features aggregation.

Experimental Results

- **Classification performance**





- **Communication performance**