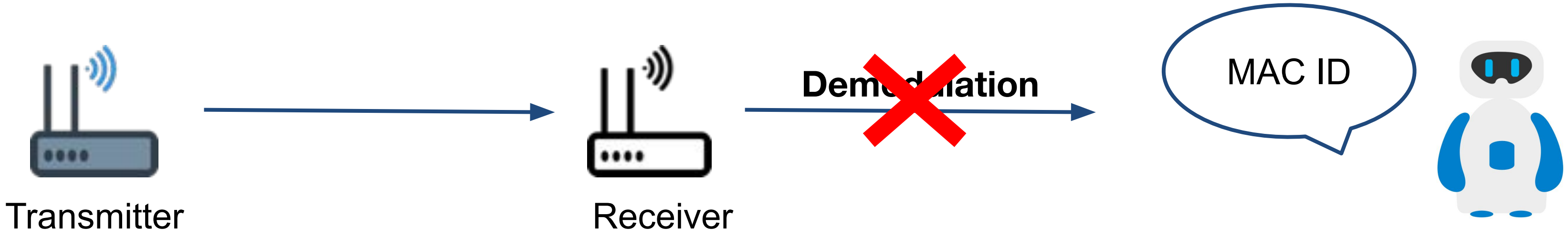
Northeastern

MAC ID Spoofing-Resistant Radio Fingerprinting

Tong Jian, Bruno Costa Rendon, Andrey Gritsenko,
Jennifer Dy, Kaushik Chowdhury, and Stratis Ioannidis

Radio Fingerprinting

- Detecting **transmission source** of signals is a key security mechanism

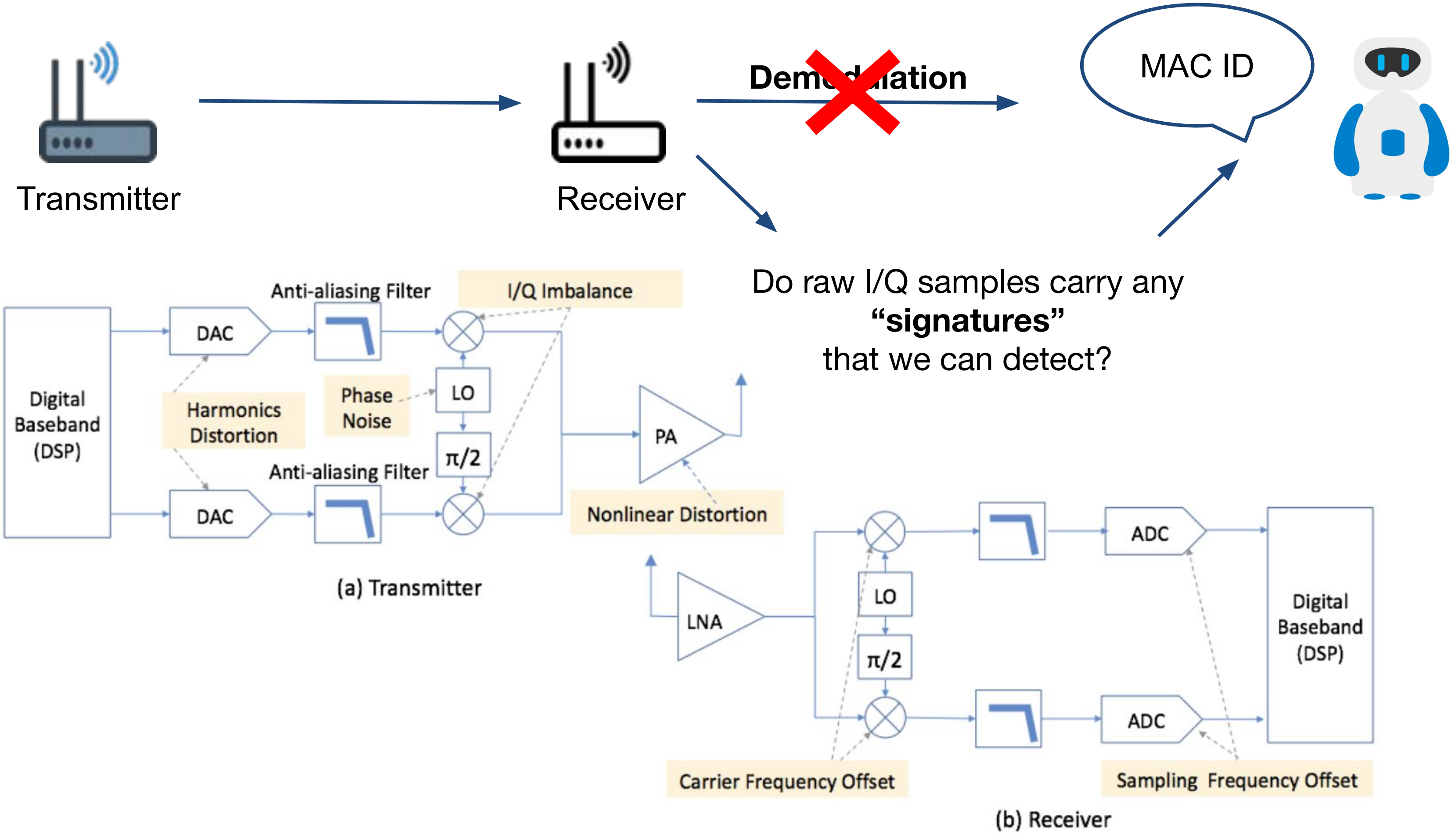


! Demodulation:

- protocol-specific
- limits ability of detection over new packet frame structures, channel bandwidths, modulation choices, and coding schemes

Radio Fingerprinting

- Detecting **transmission source** of signals is a key security mechanism



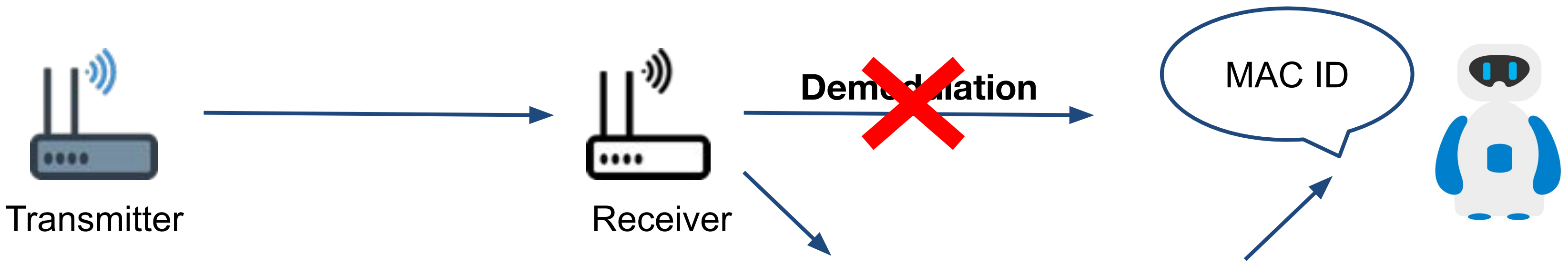
Do raw I/Q samples carry any **"signatures"** that we can detect?

(a) Transmitter

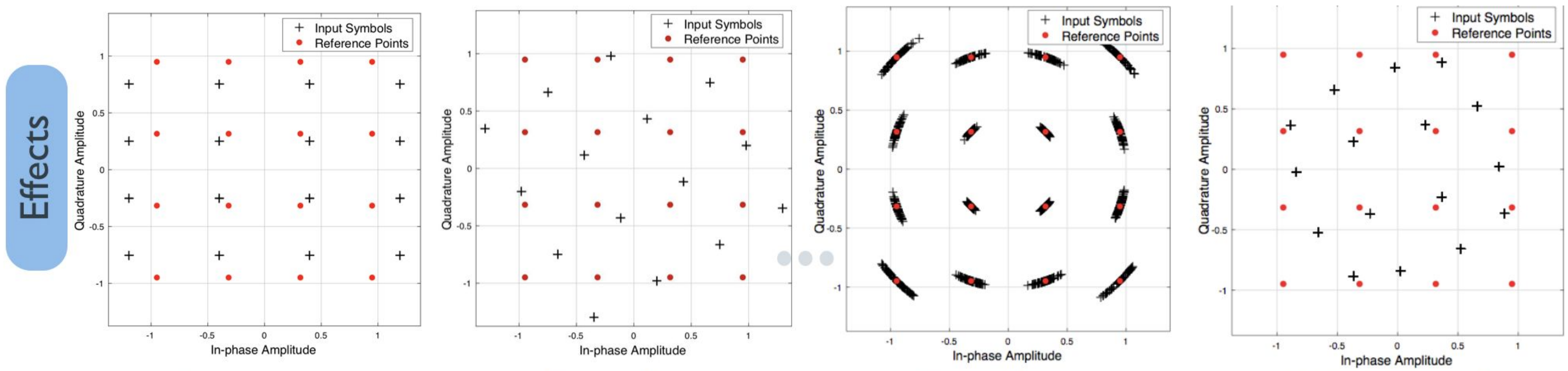
(b) Receiver

Radio Fingerprinting

- Detecting **transmission source** of signals is a key security mechanism



Do raw I/Q samples carry any
“signatures”
that we can detect?



IQ Imbalance

Phase offset

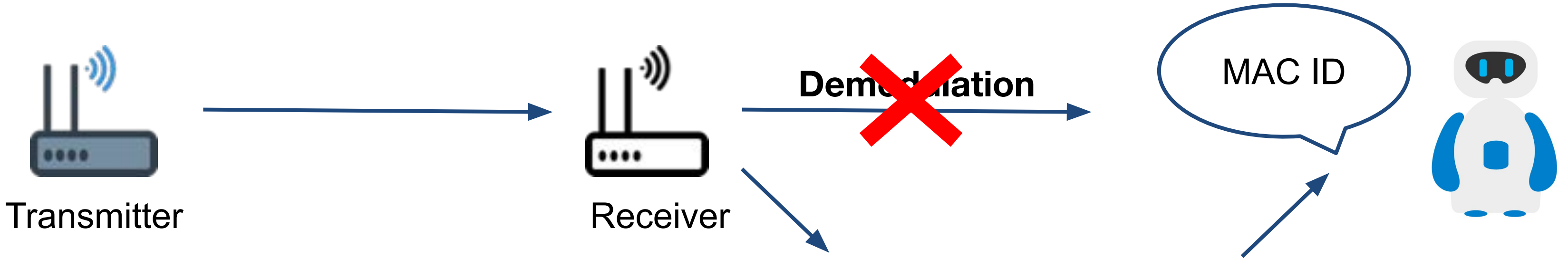
Phase Noise

AM/PM Distortions

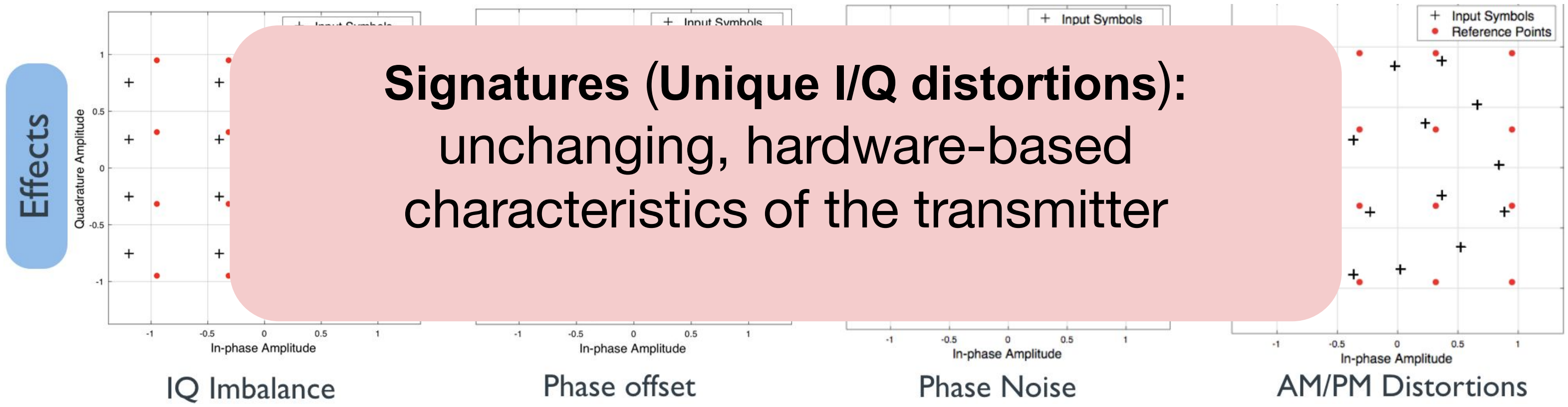
*K. Sankhe, M. Belgiovine,F. Zhou, S. Riyaz, S. Ioannidis, and K. R. Chowdhury, "ORACLE: Optimized Radio clAssification through Convolutional neuraL nEtworks,” IEEE INFOCOM 2019, Paris, France, May. 2019

Radio Fingerprinting

- Detecting **transmission source** of signals is a key security mechanism



Do raw I/Q samples carry any
“signatures”
that we can detect?



Signatures (Unique I/Q distortions):
unchanging, hardware-based
characteristics of the transmitter

IQ Imbalance

Phase offset

Phase Noise

AM/PM Distortions

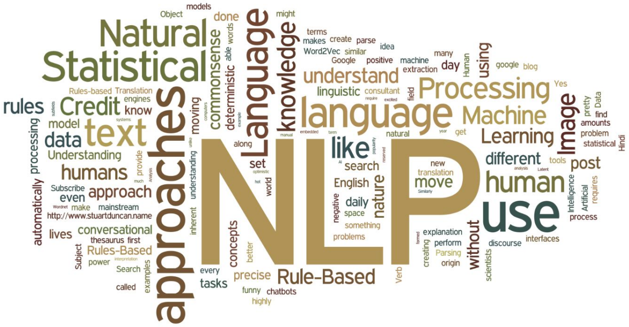
*K. Sankhe, M. Belgiovine,F. Zhou, S. Riyaz, S. Ioannidis, and K. R. Chowdhury, "ORACLE: Optimized Radio clAssification through Convolutional neuraL nEtworks,” IEEE INFOCOM 2019, Paris, France, May. 2019

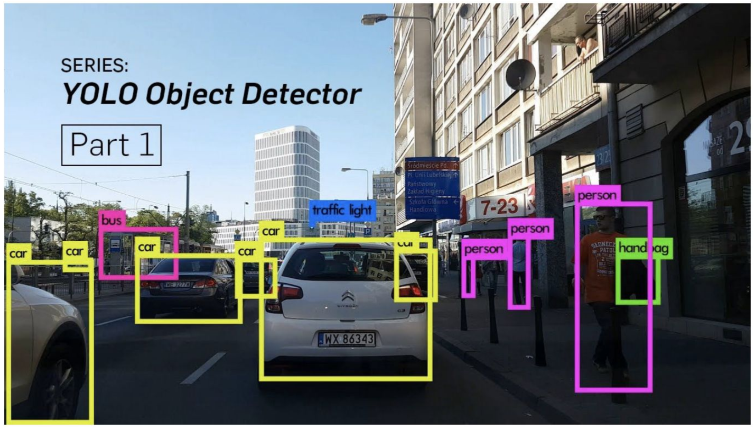
Radio Fingerprinting

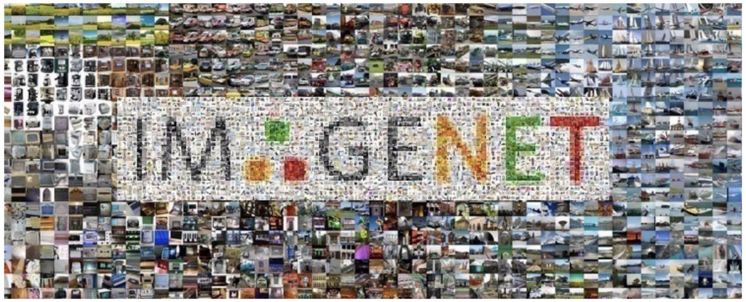
- Need for CNNs:
  - End-to-End feature interpreters -> protocol-independent

Radio Fingerprinting

- Need for CNNs:
  - End-to-End feature interpreters -> protocol-independent
  - Demonstrated performance record on numerous inference problems across application domains







upper right: NLP
lower right ImageNet
left: Object Detection

Challenges

- Features extracted by deep models cannot be easily interpreted!!!

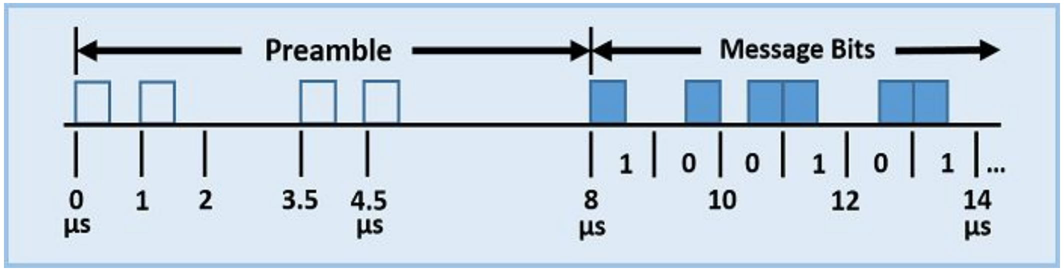
Indeed learns
unique **I/Q distortions**

OR

Simply picking up
artifacts present in the data

- Unfortunately, almost all transmissions contain a strongly discriminative artifact, **the identity of the transmitting device**, which is often included in a transmitted packet

ADS-B



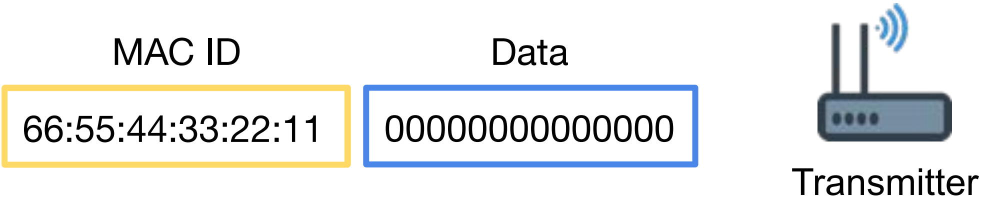
Source: https://www.mathworks.com/help/examples/xilinxzynqbasedradio_product/win64/zynqRadioHWSWADSBAD9361AD9364SL_ModeS_PPM.png

Challenges

- Features extracted by deep models cannot be easily interpreted!!!

Indeed learns unique **I/Q distortions** **OR** Simply picking up **artifacts** present in the data

- If latter…



Our Contributions

❏ Slicing technique

— makes the classifier **resistant** to learning **MAC IDs** as features

❏ Experiments on WiFi and ADS-B demonstrate slicing helps

— **100 %** -> **bitwise identical** transmissions by 19 devices

— **99.7%** -> **MAC ID** in the test set are shuffled

Overview

❑ Framework

❑ Experiments on WiFi protocol

❑ Experiments on ADS-B protocol

Overview

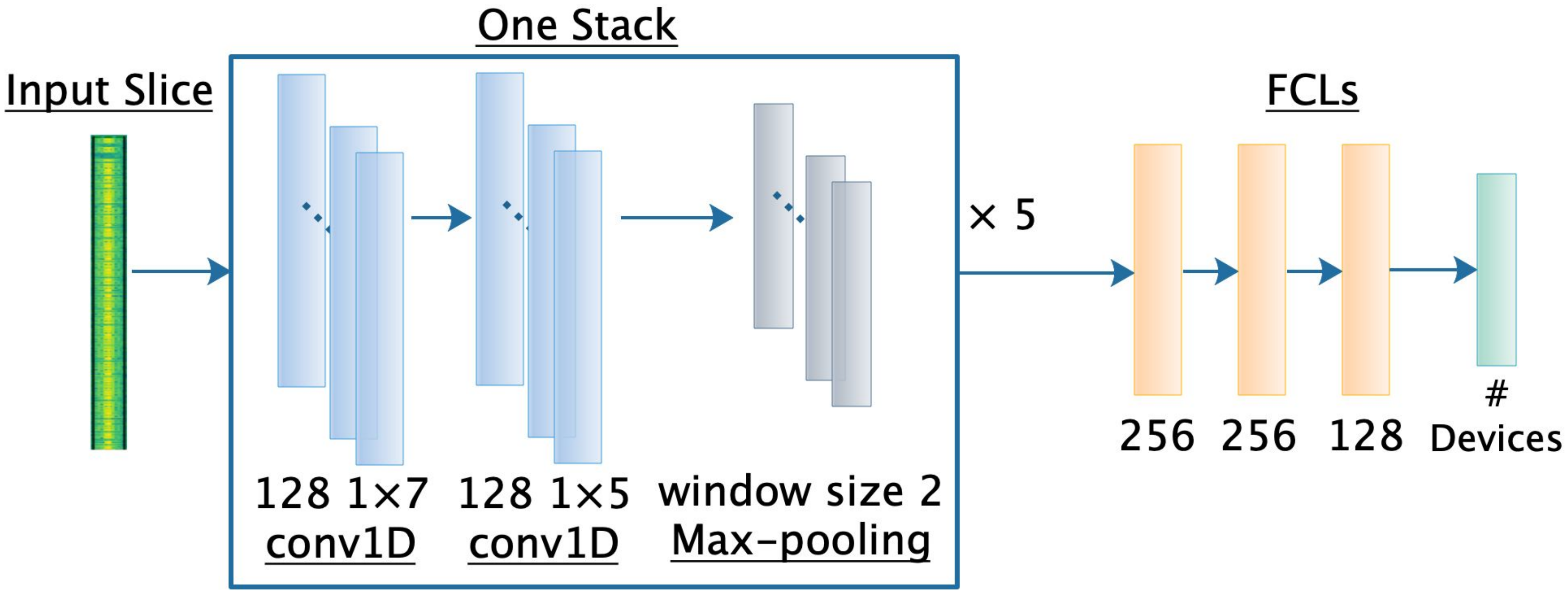
❑ Methodology

❑ Experiments on WiFi protocol

❑ Experiments on ADS-B protocol

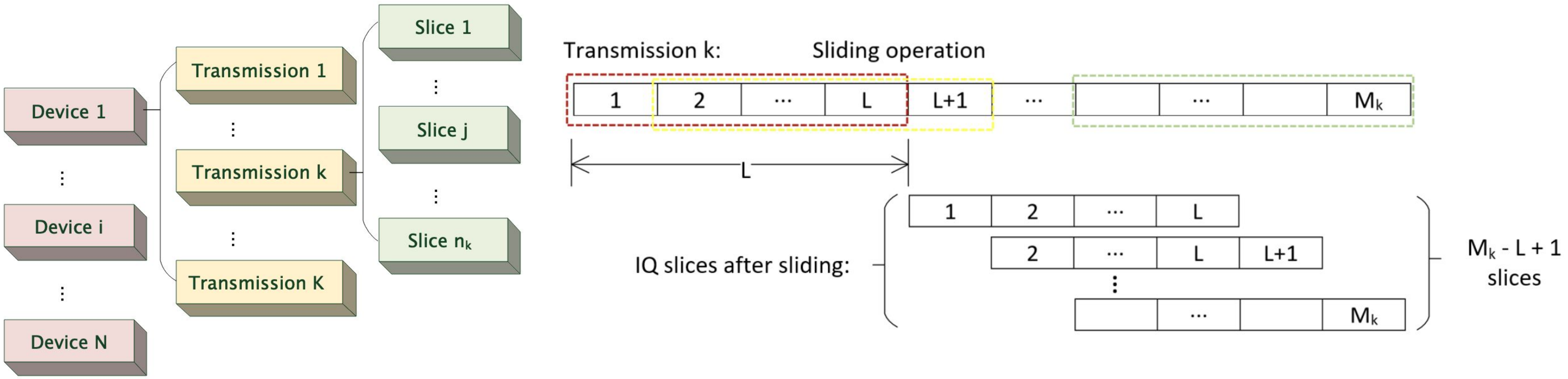
Methodology: Architecture

- CNN Architecture



Methodology: Slicing

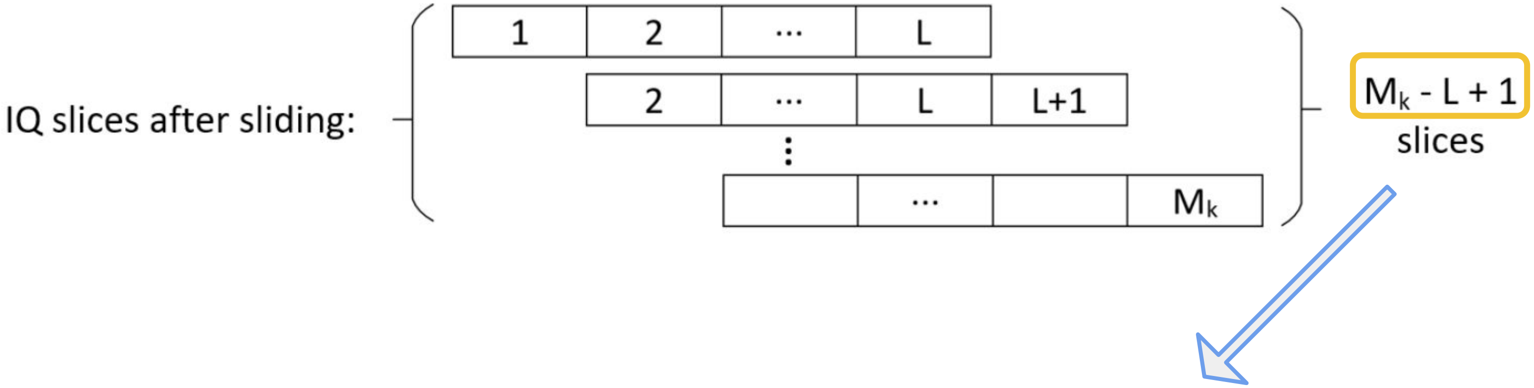
- Slicing [Riyaz et al]



*Riyaz, K. Sankhe, S. Ioannidis, and K. Chowdhury, “Deep learning convolutional neural networks for radio identification,” IEEE Communications Magazine, vol. 56, no. 9, pp. 146–152, 2018.

Methodology: Slicing

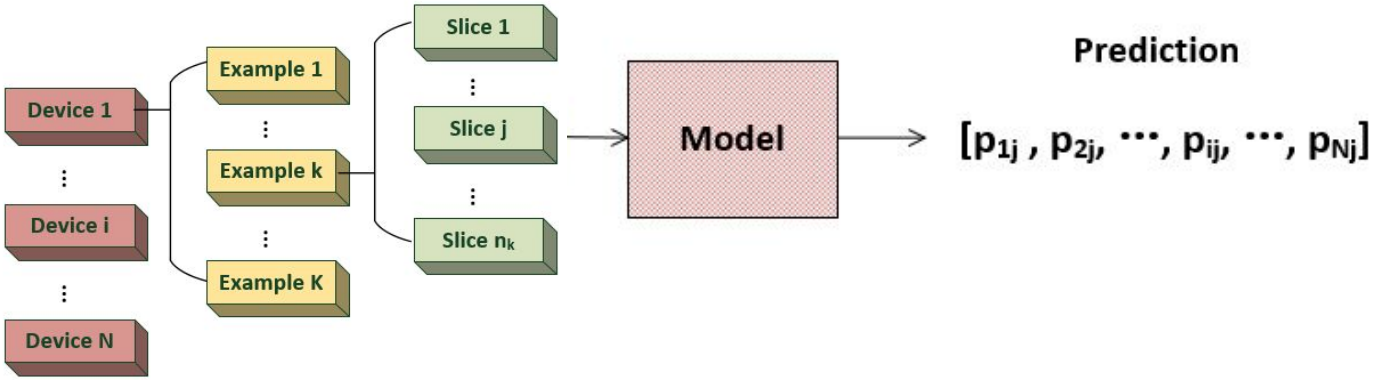
- Randomized slicing:



- For each transmission k, we sample $\boxed{n_k = \frac{M_k - L + 1}{L} \cdot \kappa}$ slices uniformly at random from all generated slices

Methodology: Testing

- We evaluate per-slice accuracy on test set
- We also evaluate per-transmission accuracy
  - Suppose there are N devices, and transmission k has n_k slices
  - p_{ij} is the probability of slice j classified as belonging to device i
    - Sum of probability over all slices: $\boxed{\hat{y} = \arg\max_i \sum_{j=1}^{n_k} p_{ij}}$



Methodology

- Advantages of randomized slicing:
  - satisfies the requirement of fixed-size input for CNNs
  - improves classifier’s ability to learn shift-invariant features
  - reduces computations during training

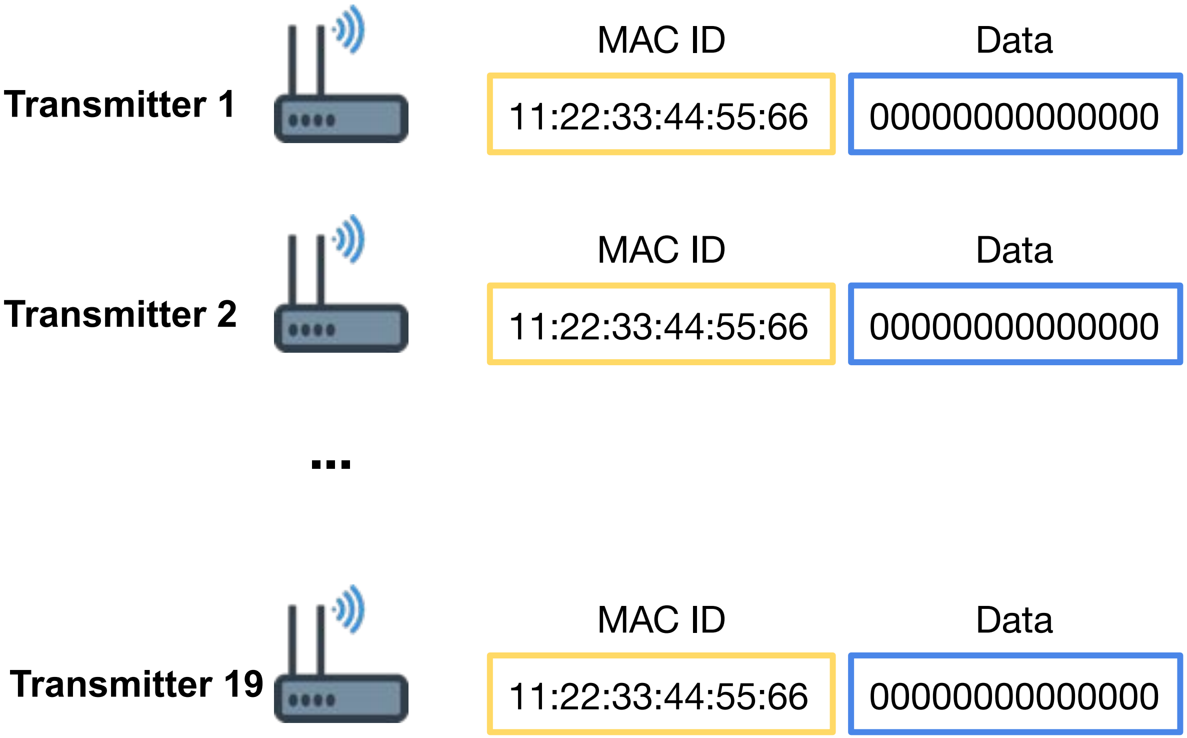
Overview

- ❑ Methodology
- ❑ Experiments on WiFi protocol
- ❑ Experiments on ADS-B protocol

Experiments on WiFi protocol

- Datasets:
  - Bitwise Identical WiFi:

    Bitwise identical WiFi transmissions by 19 devices.



Transmitter 1

MAC ID

11:22:33:44:55:66

Data

00000000000000

Transmitter 2

MAC ID

11:22:33:44:55:66

Data

00000000000000

...

Transmitter 19

MAC ID

11:22:33:44:55:66

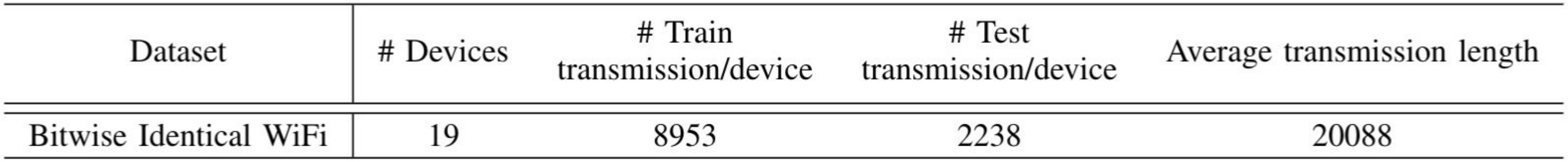
Data

00000000000000

Experiments on WiFi protocol

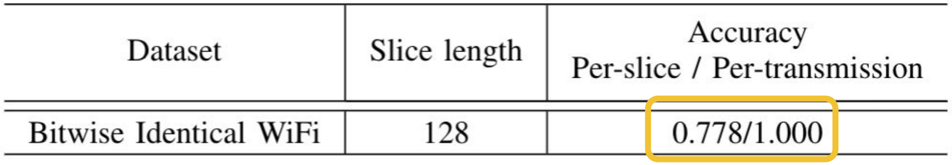
- Datasets:
  - Bitwise Identical WiFi:

    Bitwise identical WiFi transmissions by 19 devices.

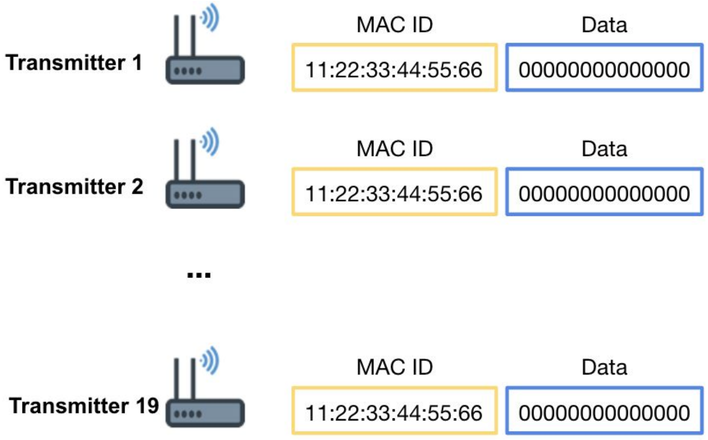


Dataset	# Devices	# Train transmission/device	# Test transmission/device	Average transmission length
Bitwise Identical WiFi	19	8953	2238	20088

- Results:



Dataset	Slice length	Accuracy Per-slice / Per-transmission
Bitwise Identical WiFi	128	0.778/1.000



Experiments on WiFi protocol

- Datasets:
  - Scrambled MAC WiFi:

    MAC IDs are randomly permuted among the signals in the test set.

	Training Set		Test Set	
	MAC ID	Data	MAC ID	Data
Transmitter 1	11:11:11	asdlkn2p	22:22:22	2ejrnlfddf
	11:11:11	23oidfkjn	33:33:33	dfaldkflkd
Transmitter 2	22:22:22	130df093	33:33:33	huhuhuhu
	22:22:22	2odfoiejo	11:11:11	omomom
Transmitter 3	33:33:33	asasasas	33:33:33	qdfqfqdq
	33:33:33	vcvcvcvc	11:11:11	bhbhbhb

Experiments on WiFi protocol

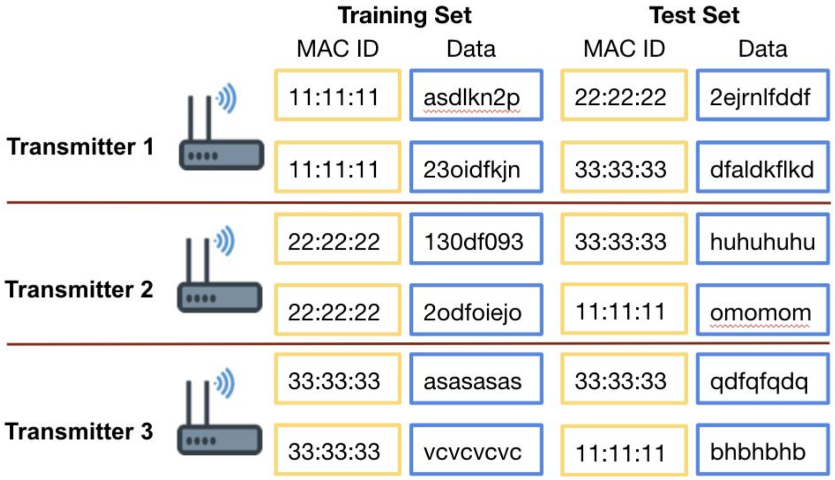
- Datasets:
  - Scrambled MAC WiFi:

    MAC IDs are randomly permuted among the signals in the test set.

Dataset	# Devices	# Train transmission/device	# Test transmission/device	Average transmission length
Scrambled MAC WiFi	100	1000	1000	45183

- Results:

Dataset	Slice length	Accuracy Per-slice / Per-transmission
Scrambled MAC WiFi	1024	0.972/0.997



	Training Set		Test Set	
	MAC ID	Data	MAC ID	Data
Transmitter 1	11:11:11	asdlkn2p	22:22:22	2ejrnlfddf
	11:11:11	23oidfkjn	33:33:33	dfaldkflkd
Transmitter 2	22:22:22	130df093	33:33:33	huhuhuhu
	22:22:22	2odfoiejo	11:11:11	omomom
Transmitter 3	33:33:33	asasasas	33:33:33	qdfqfqdq
	33:33:33	vcvcvcvc	11:11:11	bhbhbhb

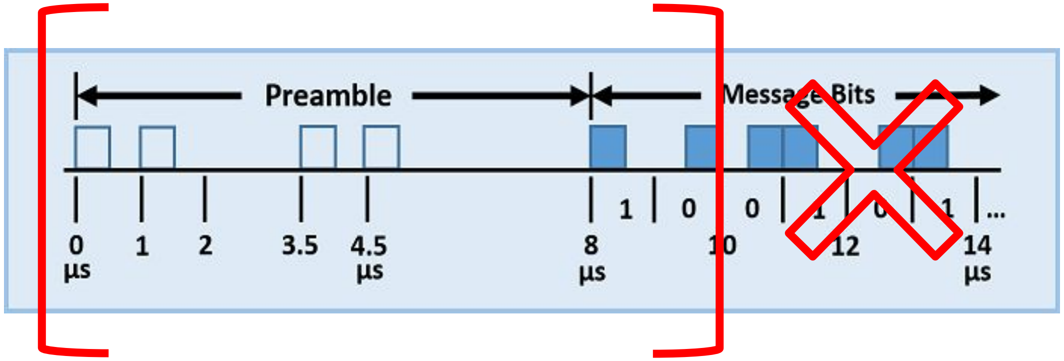
Overview

- ❑ Methodology
- ❑ Experiments on WiFi protocol
- ❑ Experiments on ADS-B protocol

Experiments on ADS-B protocol

- Datasets:

Dataset	# Devices	# Train transmission/device	# Test transmission/device	Average transmission length
ADS-B	50	141	55	9519



Crop

Experiments on ADS-B protocol

Dataset	# Devices	# Train transmission/device	# Test transmission/device	Average transmission length
ADS-B	50	141	55	9519

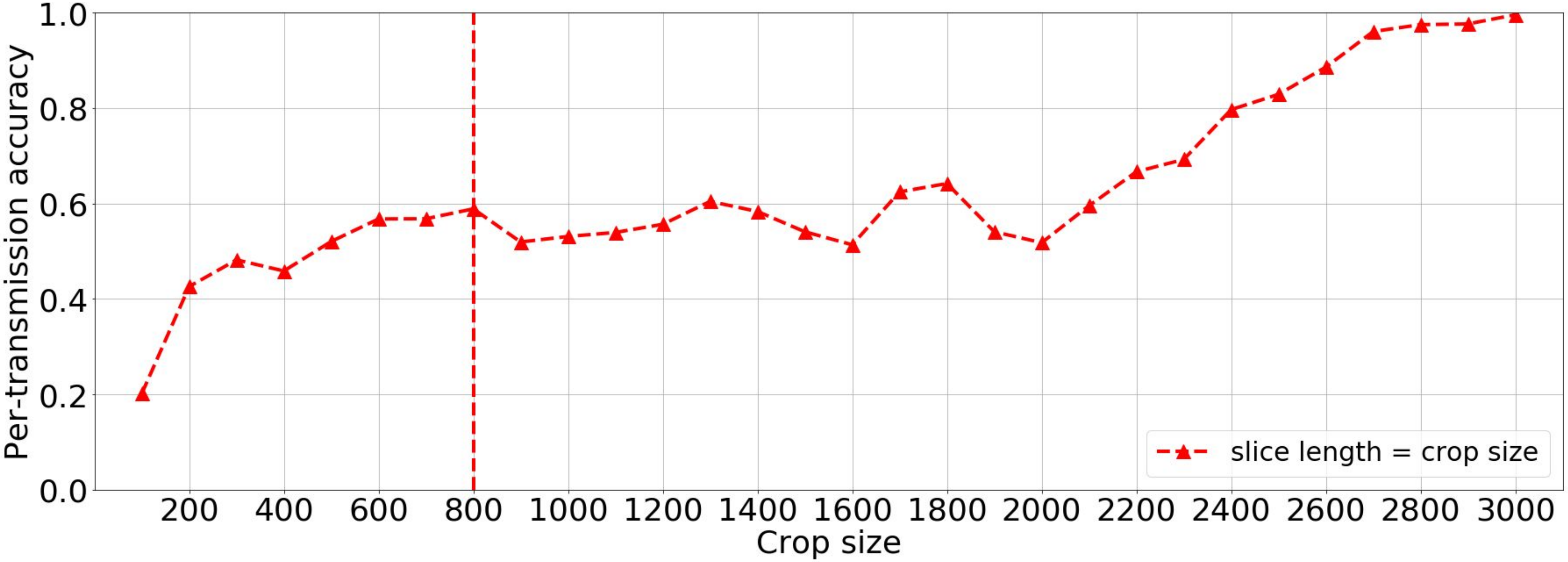
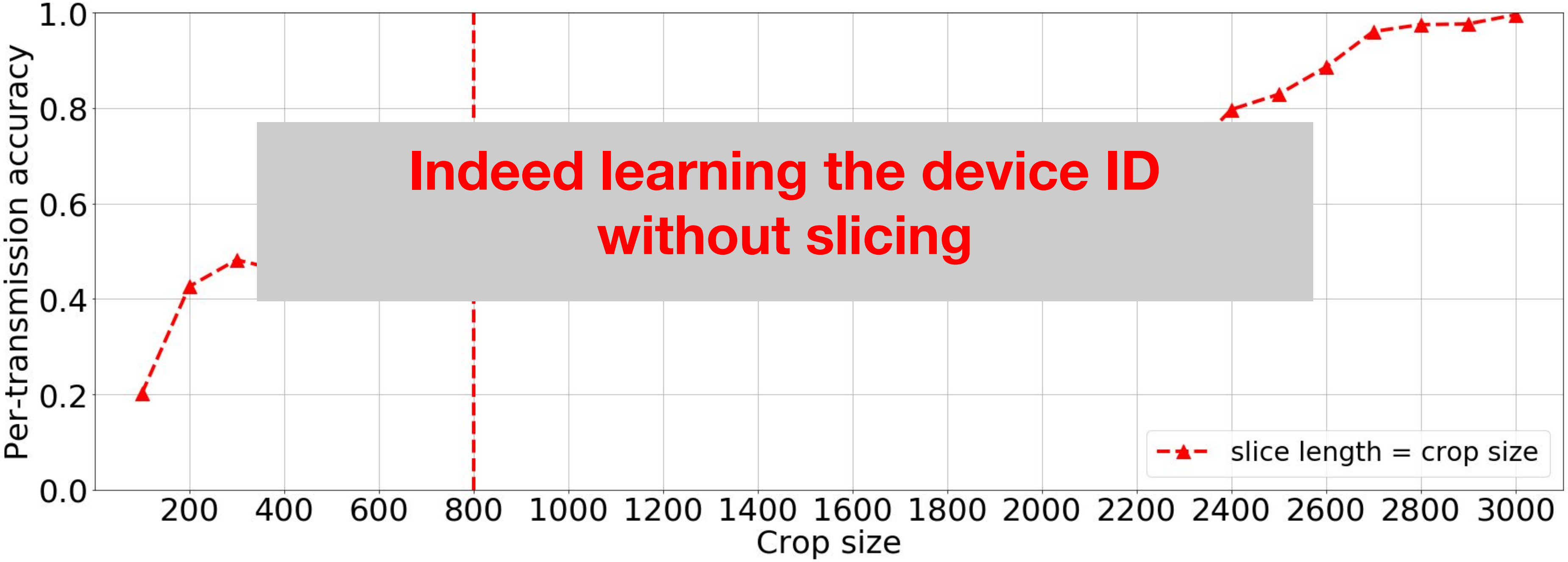


Fig. Test Accuracy without Slicing

Experiments on ADS-B protocol

Dataset	# Devices	# Train transmission/device	# Test transmission/device	Average transmission length
ADS-B	50	141	55	9519



Indeed learning the device ID without slicing

Fig. Test Accuracy without Slicing

Experiments on ADS-B protocol

Dataset	# Devices	# Train transmission/device	# Test transmission/device	Average transmission length
ADS-B	50	141	55	9519

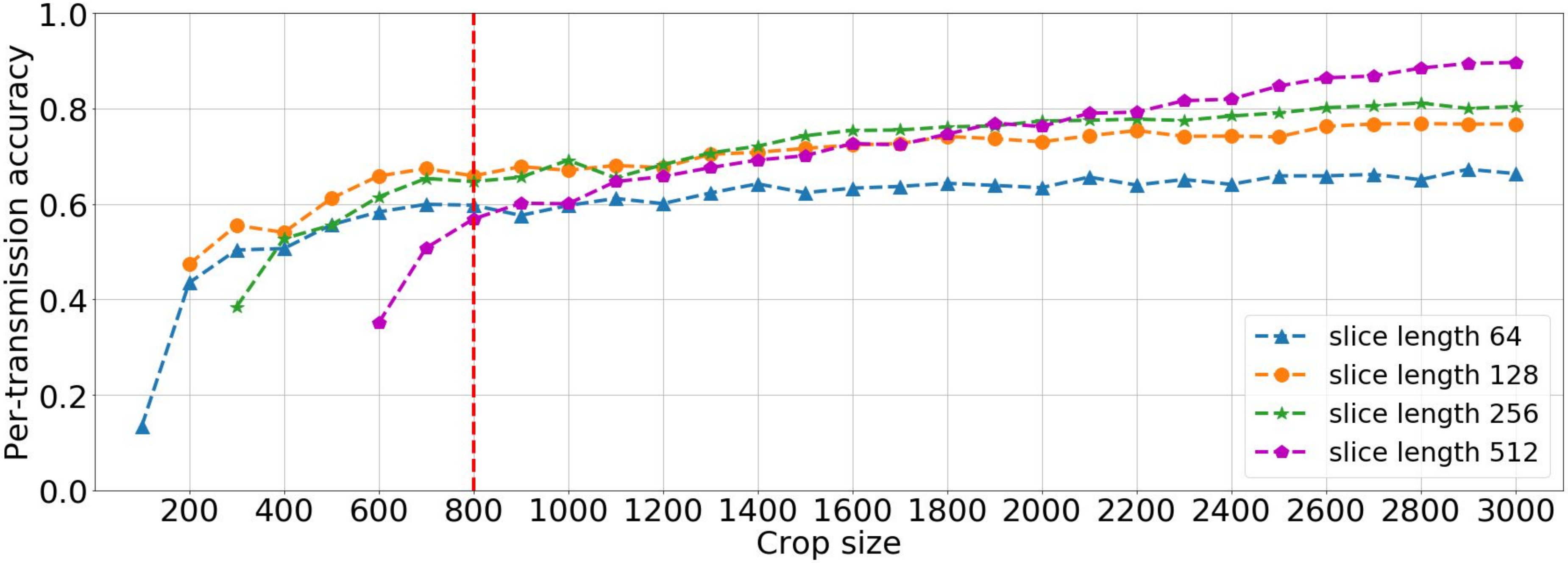


Fig. Test Accuracy with Slicing

Summary & Future Directions

❑ Classifying transmission slices
- enhances **shift-invariance**
- MAC ID spoofing-resistant
- experiments on WiFi and ADS-B protocols.

❑ We are working on...
- classification over >10K transmitters
- beating channel variations

рахмат
danke
謝謝
ngiyabonga
شكراً جزيلاً
Баярлалаа
spas
teşekkür ederim
mersi
kia ora
barka
welalin
tack
спасибо
faafetai lava
vinaka
blagodaram
dank je
misaotra
matondo
paldies
grazzi
mahalo
tapadh leat
спасибі
kiitos
dankie
nanni
nandri
dhanyavad
mauruuru
thank you
gracias
хвала
asante
manana
aciu
obrigada
tenki
bedankt
bayarlalaa
gràcie
hvala
köszönöm
dankon
akun
murakoze
enkosi
djiere dieuf
tau
mochchakkeram
дякую
mamnun
dziękuję
chnorakaloutioun
gratias ago
gràcies
sulpáy
taiku
go raibh maith agat
chokrane
sobodi
dėkuji
sagolun
sukriya
kop khun krap
obrigado
mèsi
didi madloba
najis tuke
ありがとう
tanemirt
grazie
arigatô
takk
dakujem
trugarez
kam sah hamnida
rahmat
terima kasih
rahmet
diolch
dhanyavadagalu
shukriya
merce
мерси
তোমাকে ধন্যবাদ
감사합니다
xièxie
ευχαριστώ
merci