SLNet: A Spectrogram Learning Neural Network for Deep Wireless Sensing

Zheng Yang[1], Yi Zhang[1], Kun Qian[2], Chenshu Wu[3]
[1]Tsinghua University
[2]University of California San Diego
[3]The University of Hong Kong

NSDI 2023

汇报人：陈宇杰

Author

RESEARCH INTERESTS

- Internet of Things
- Industrial Internet
- sensing and positioning
- edge computing

Zheng Yang | 杨铮
https://tns.thss.tsinghua.edu.cn/~yangzheng/

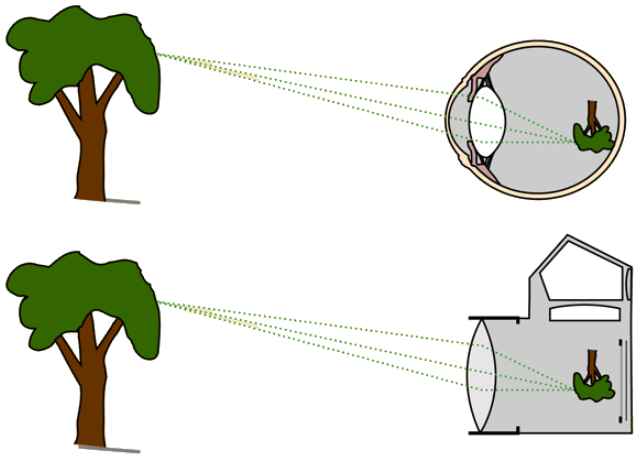
Computer Vision technologies

How human perceive the world

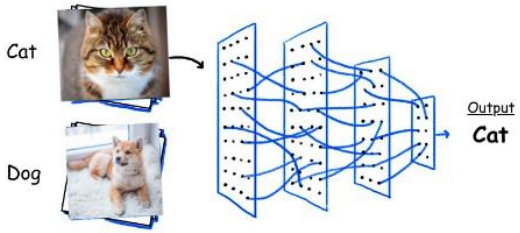
Vision is our dominant sense.

Eyes manage 80% of all information you will ever take in.

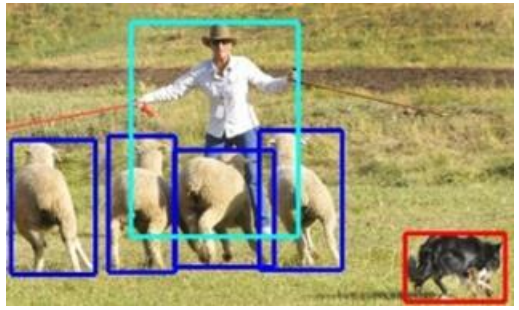
Computer Vision technologies



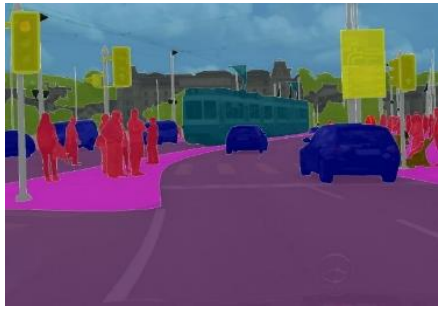
Computer Vision (CV) technologies endow computers with an effective way of SENSE.



Classification

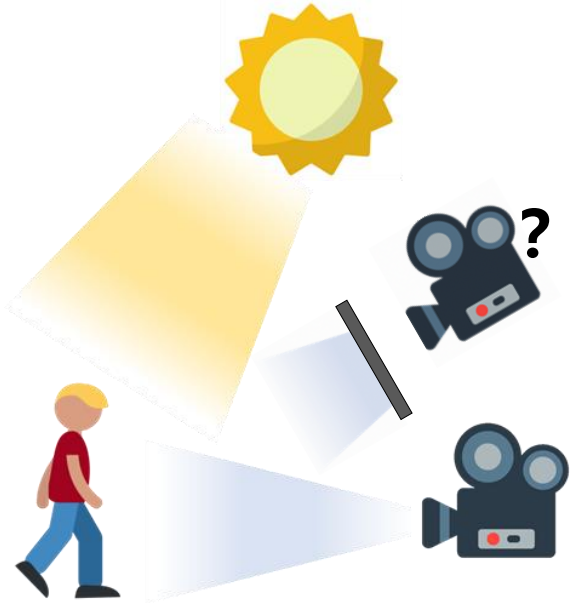


Detection



Segmentation

The limitations in vision systems



Occlusions

Privacy

Despite being effective, CV technologies still has limitations including occlusions, environmental interference, privacy leakage, etc.

The chance in wireless signals





Powerful application potential

Widely commercialized

Gesture recognition

Gait detection

Fall detection

Respiration detection

Related work

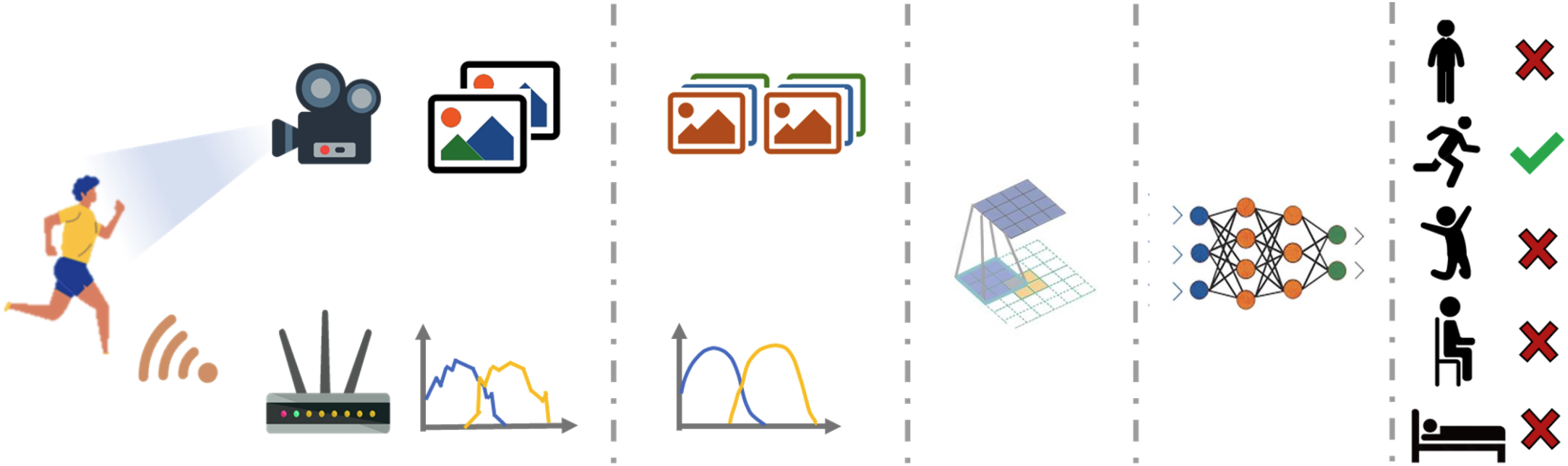
➢ Model-based wireless sensing

- QGesture: Quantifying Gesture Distance and Direction with WiFi Signals.(IMWUT 2018)
- GaitWay: Monitoring and Recognizing Gait Speed Through the Walls.(MobiCom 2020)
- SiFall: Practical Online Fall Detection with RF Sensing.(SenSys 2022)
- SMARS: Sleep Monitoring via Ambient Radio Signals.(MobiCom 2019)

➢ Learning-based wireless sensing

- Widar3.0: Zero-Effort Cross-Domain Gesture Recognition With Wi-Fi.(TPAMI 2021)
- Zero-Effort Cross-Domain Gesture Recognition with Wi-Fi.(MobiSys 2019)
- mSilent: Towards General Corpus Silent Speech Recognition Using COTS mmWave Radar.(IMWUT 2023)

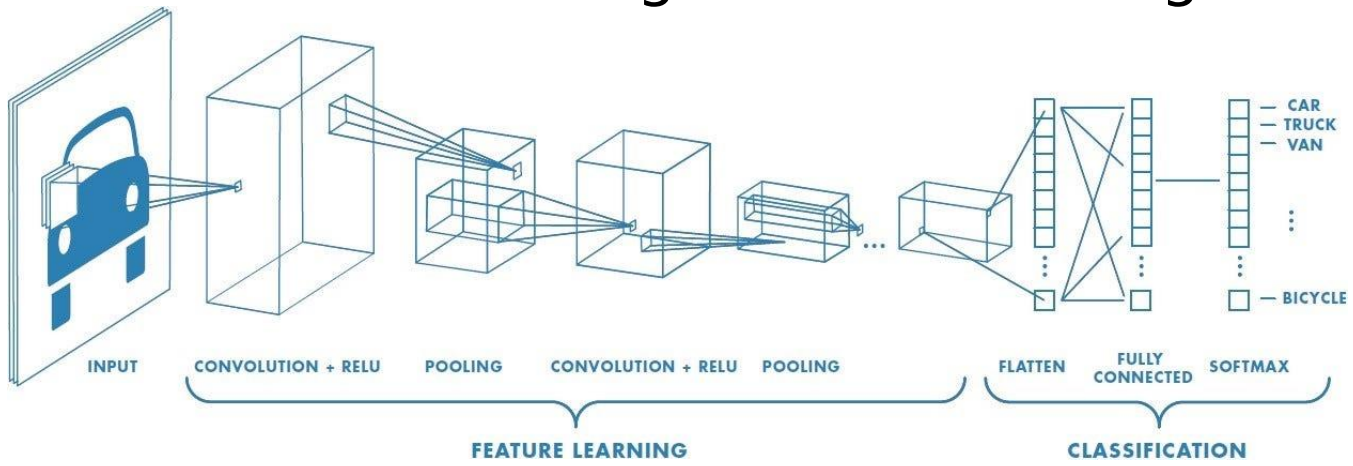
Motivation

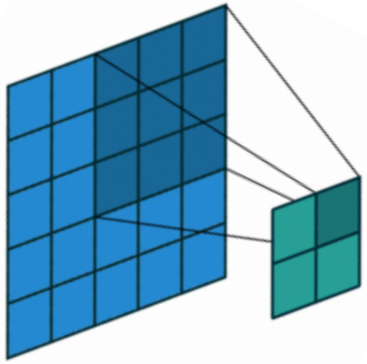


- Non-visual: RF data contains physical and geometric connotations in time, space, and frequency domains that are not visually intelligible.
- Complex: RF data is complex-valued with both amplitude and phase information.

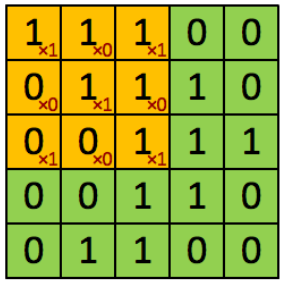
Motivation

Convolutional neural network is not designed for wireless signals.

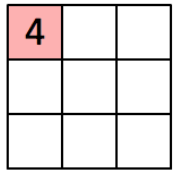




Share weights across different locations



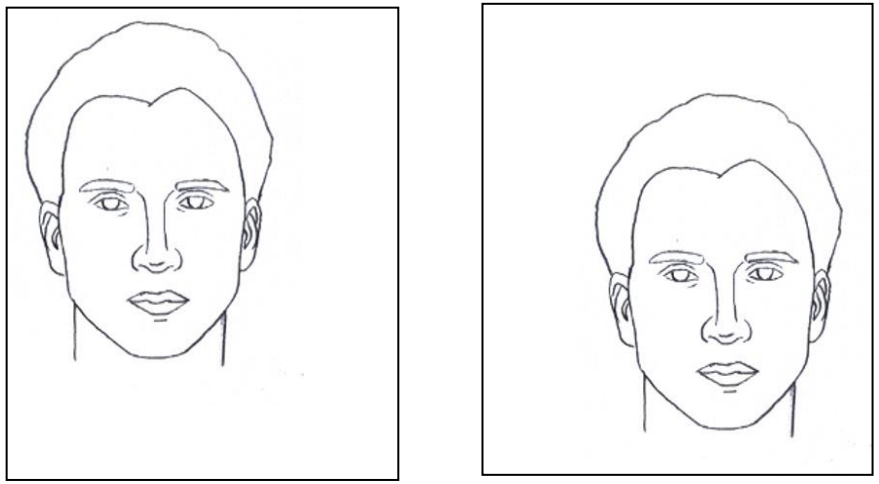
Image

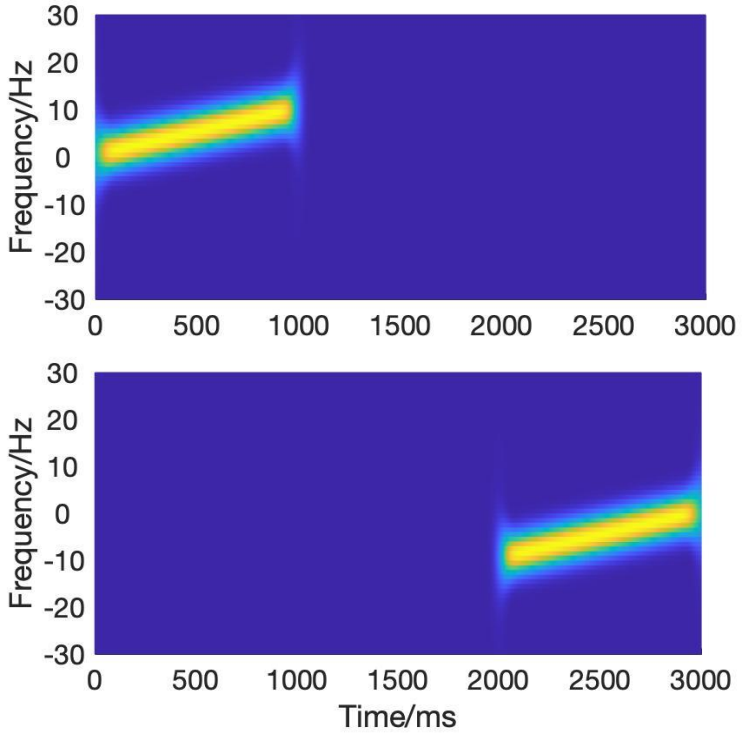


Convolved Feature

The convolution operation takes a local field of the input, discarding the global locations.

Motivation



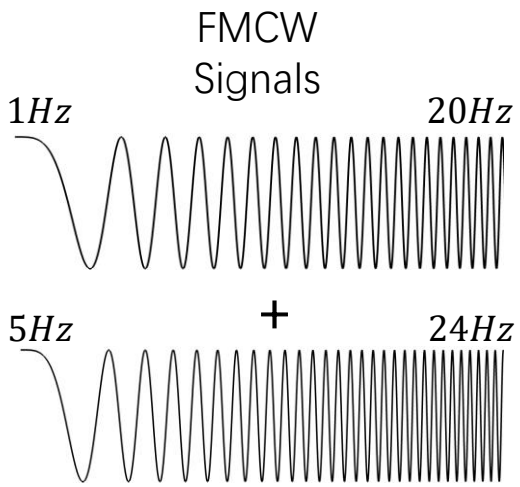


CNN is tailored for images since it is invariant to shifts.

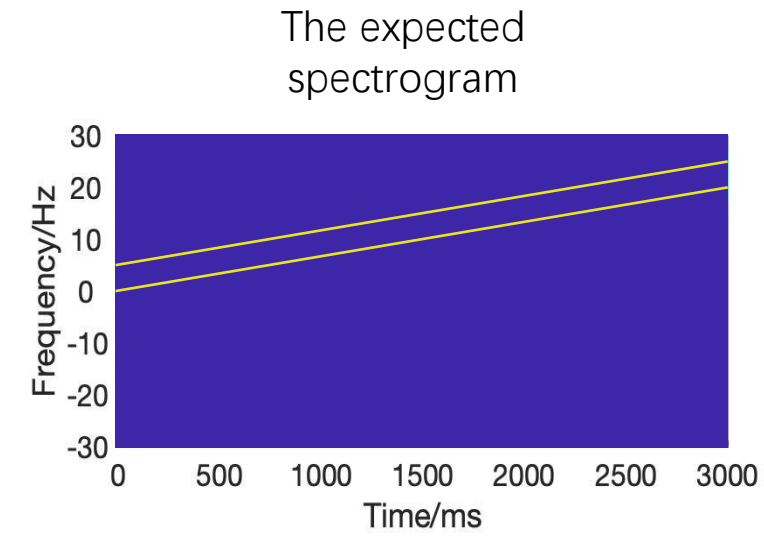
Wireless signals require global discriminations.

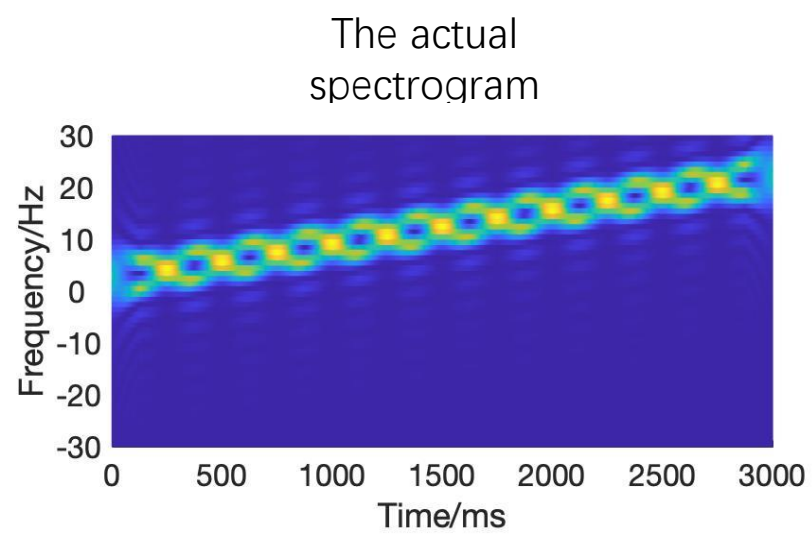
Motivation

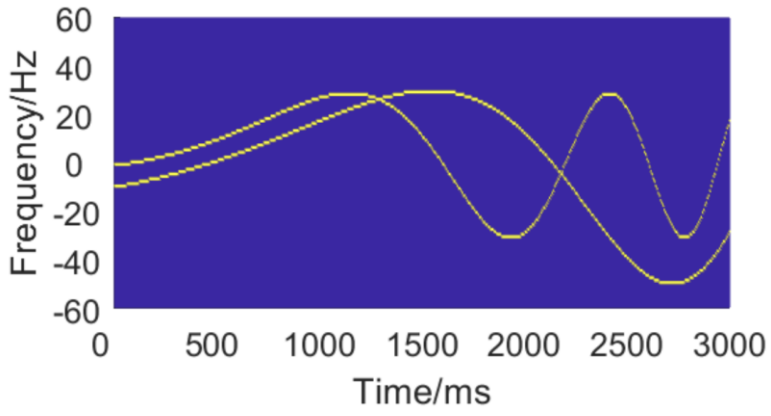
The blur in spectrograms cause difficulty for recognition.

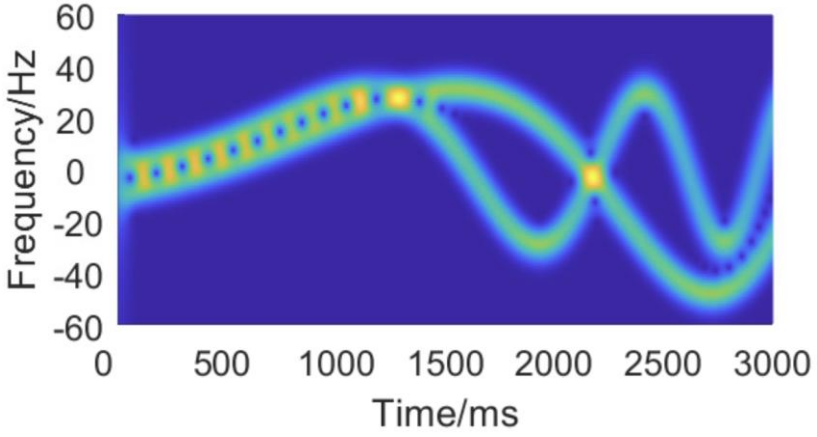


The interference is more significant for smaller frequency difference.



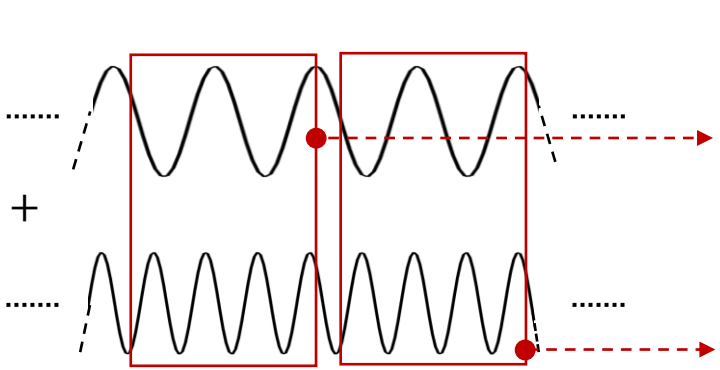


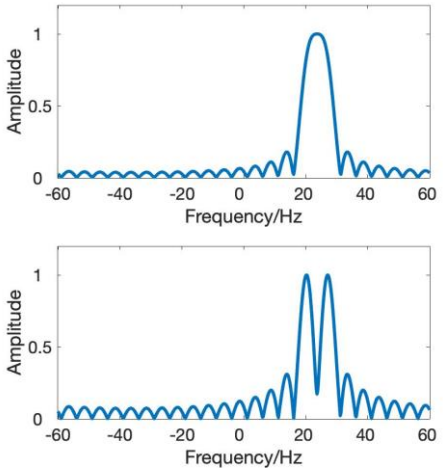


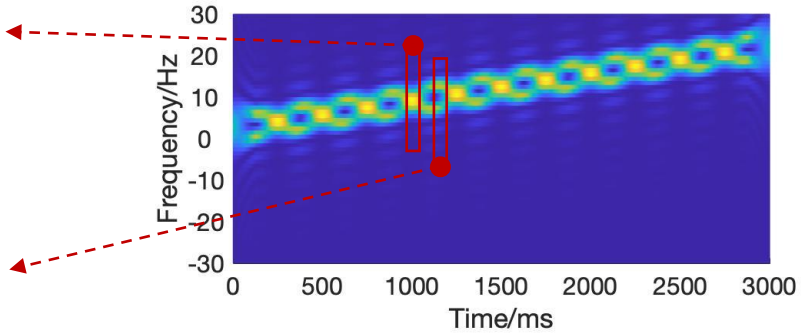


Motivation

The interference caused by spectral leackage is **unstable**.







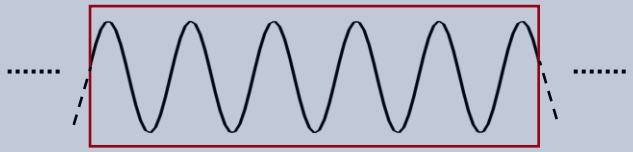
Different initial phase cause different interference patterns.

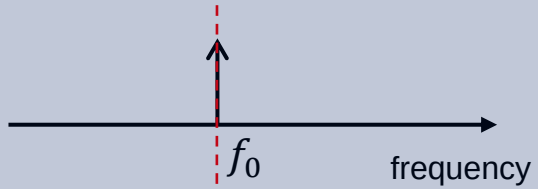
Motivation

Joseph Fourier

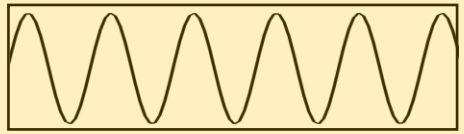
It is all about Fourier Transform (FT)

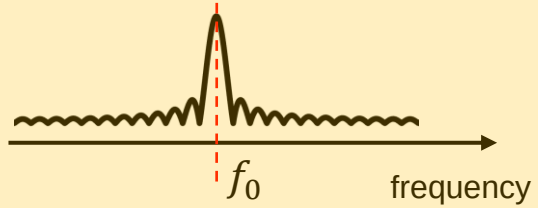
FT in theory

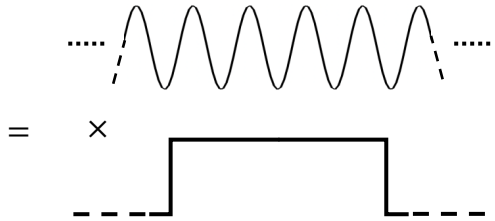




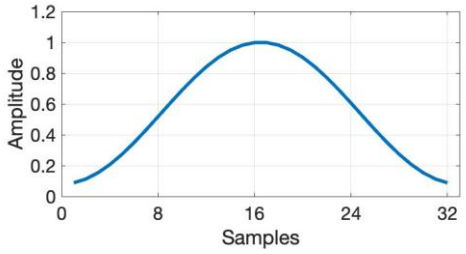
FT in practice

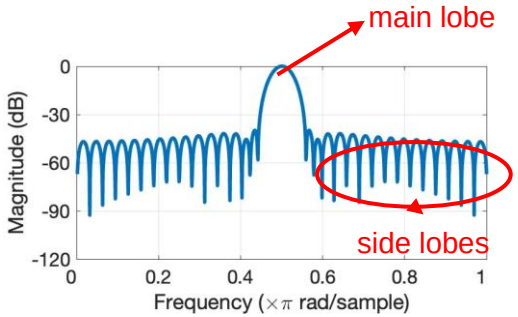




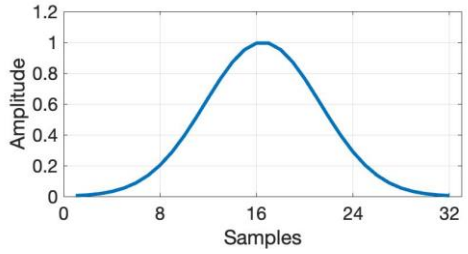


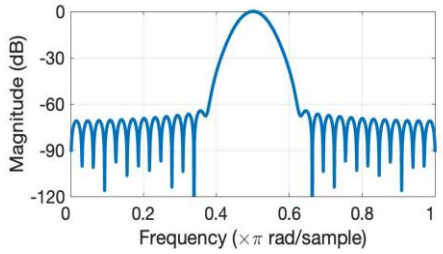
Spectral leakage





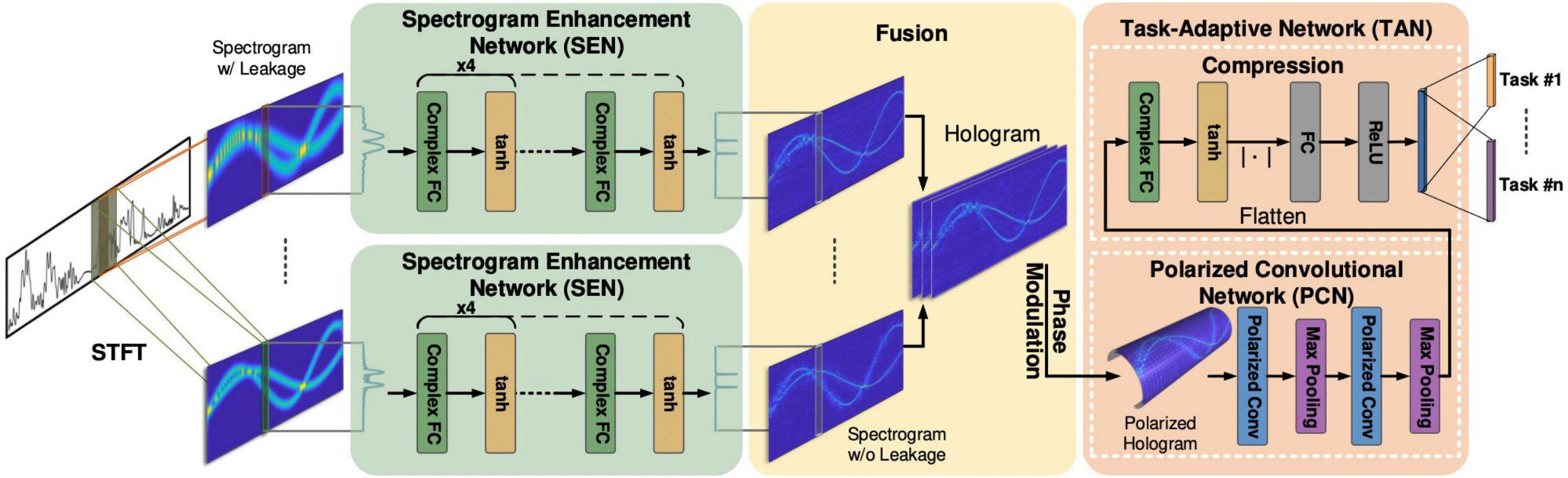
Hamming





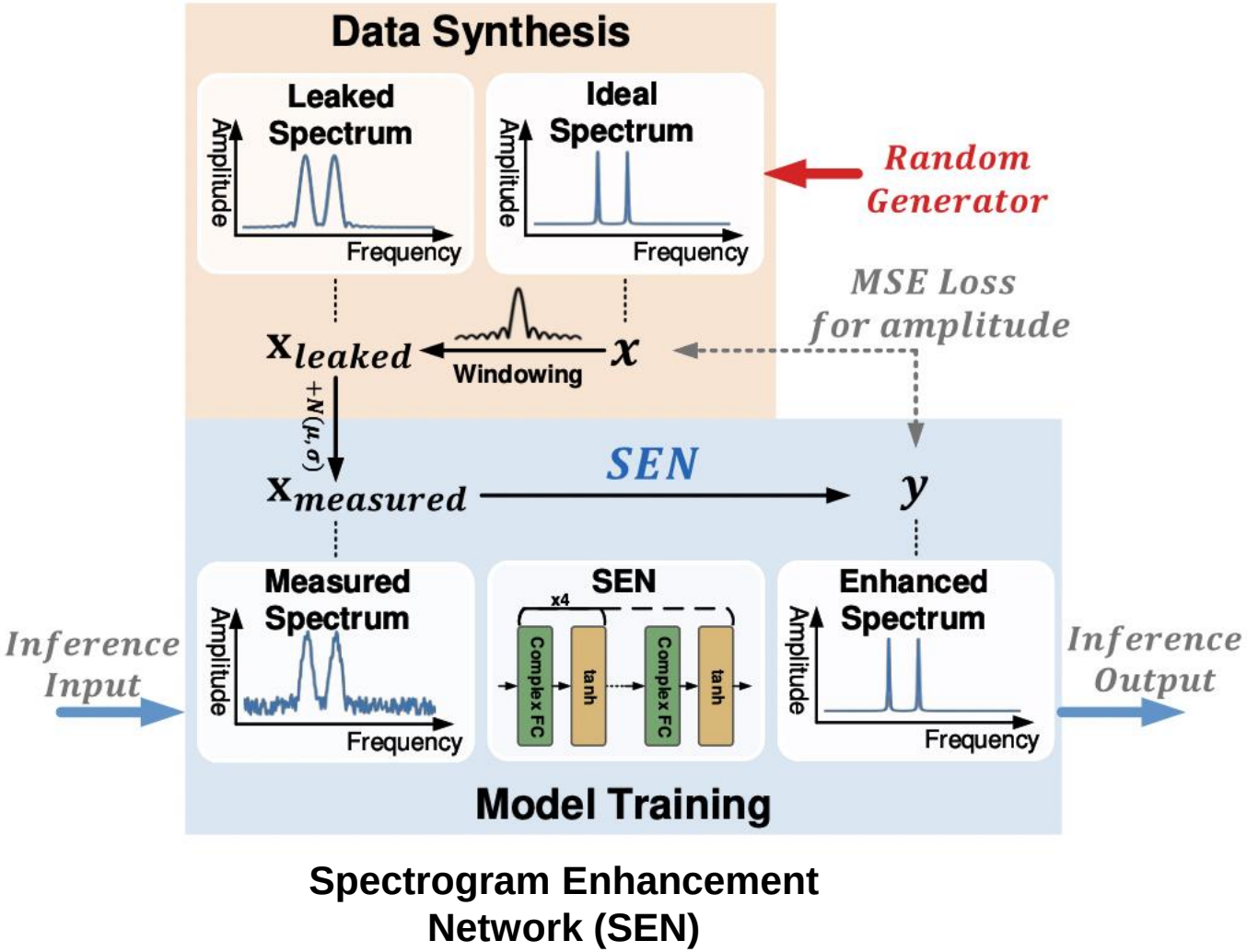
Gaussian

Design


Spectrogram
w/ Leakage
STFT
Spectrogram Enhancement Network (SEN)
x4
Complex FC
tanh
Complex FC
tanh
Spectrogram Enhancement Network (SEN)
x4
Complex FC
tanh
Complex FC
tanh
Fusion
Hologram
Spectrogram
w/o Leakage
Phase Modulation
Task-Adaptive Network (TAN)
Compression
Complex FC
tanh
|·|
FC
ReLU
Flatten
Task #1
Task #n
Polarized Convolutional Network (PCN)
Polarized Hologram
Polarized Conv
Max Pooling
Polarized Conv
Max Pooling

Design

Learning-Assisted Spectrogram Enhancement



Spectrogram Enhancement Network (SEN)

Generate an ideal spectrum with **1 to 5** frequency components randomly.

amplitudes	phases	frequencies
[0,1]	[0,2∏]	[-60,60]Hz

$$\hat{s} = As + n$$

s: ideal frequency spectrums

$\hat{s}$: estimated frequency spectrums

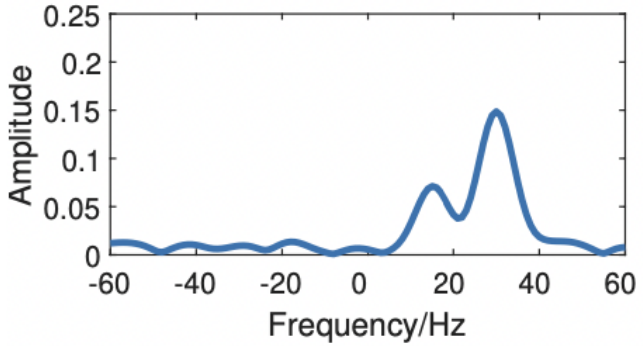
A: windowing effect

n: random complex noises

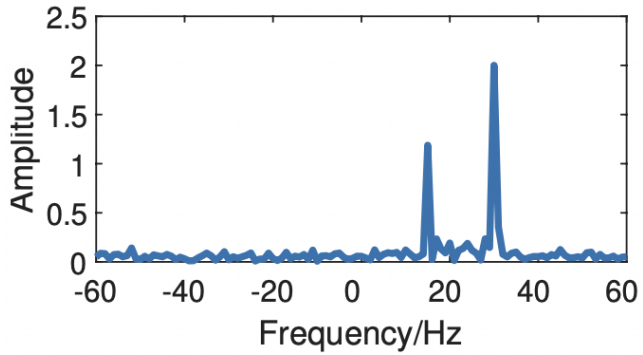
Design

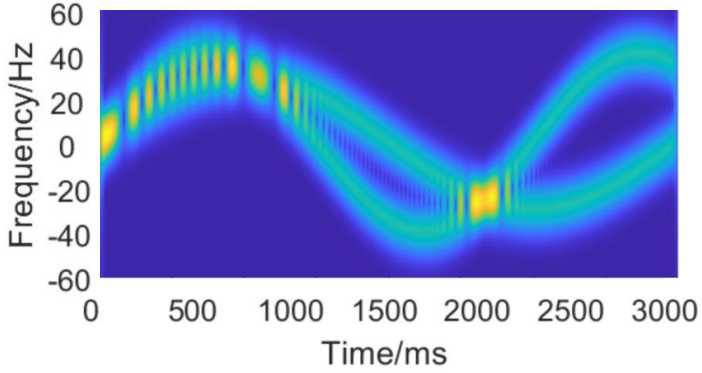
Learning-Assisted Spectrogram Enhancement

SEN results

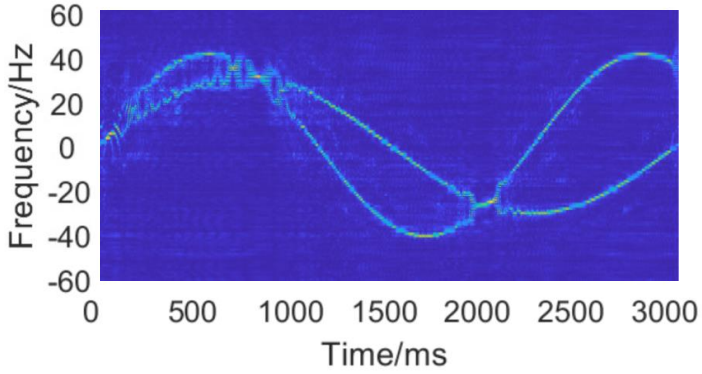


SEN



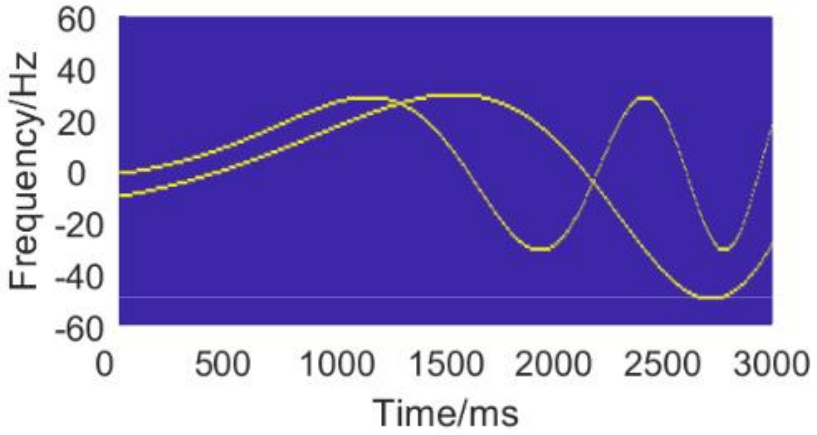


SEN

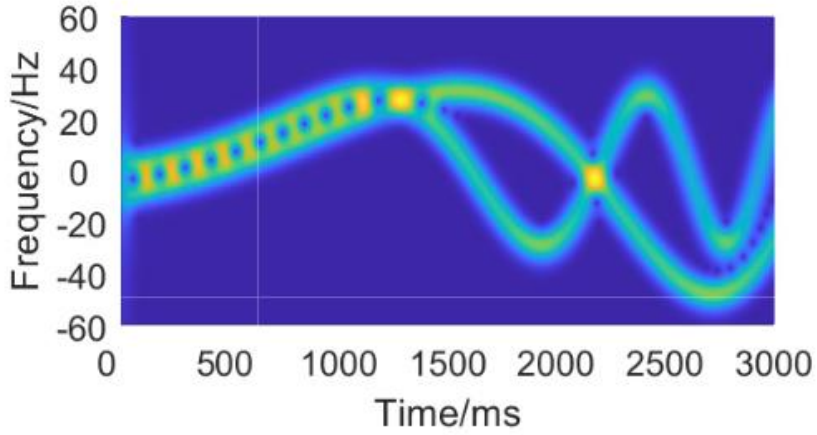


Frequency components are more distinguishable after SEN enhancement.

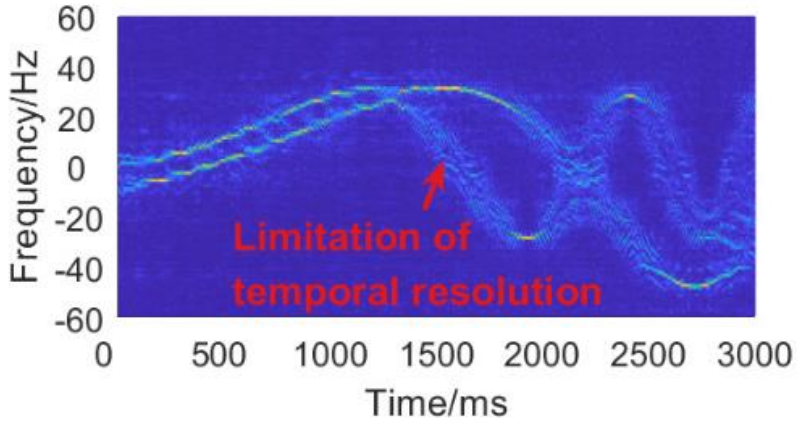
Design



(a)



(b)

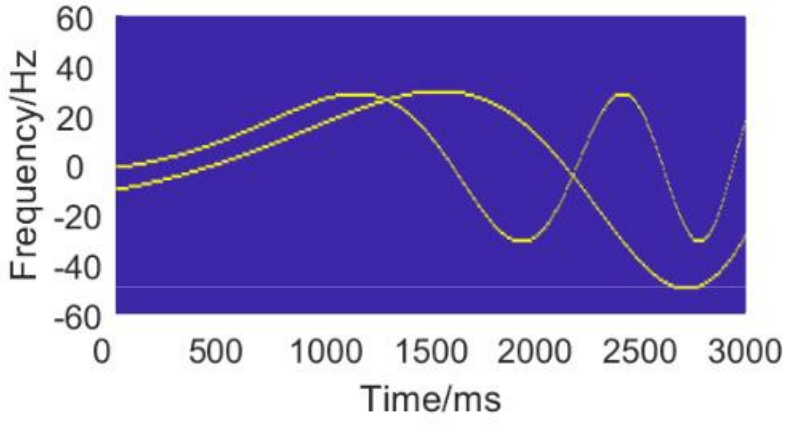


(c)

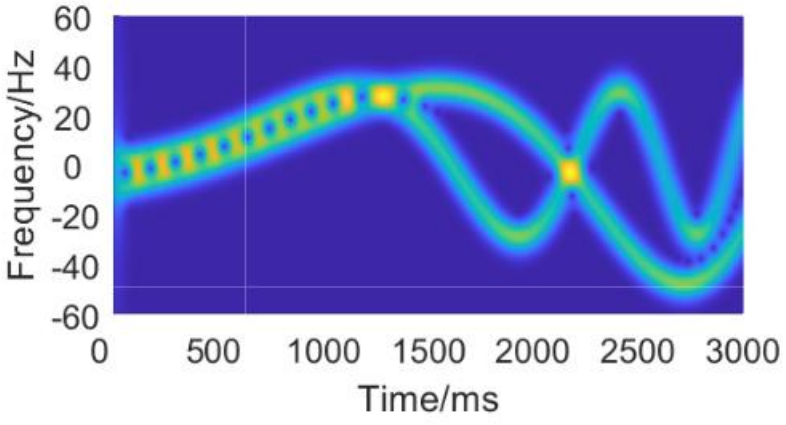
SEN assumes the frequency components remain quasi-static during the sliding window period of generating the spectrum.

Design

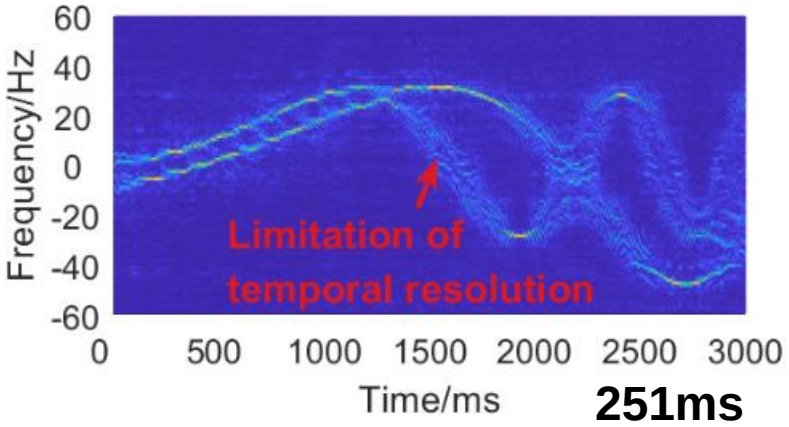
Multi-Resolution Spectrogram Fusion



(a)

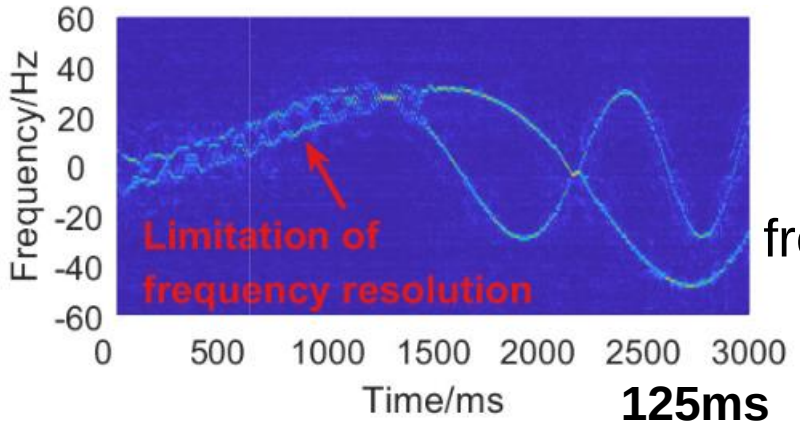


(b)



251ms

(c)



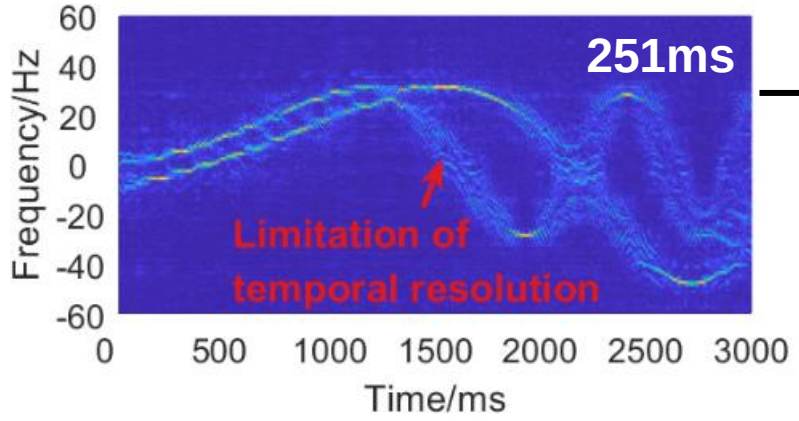
125ms

(d)

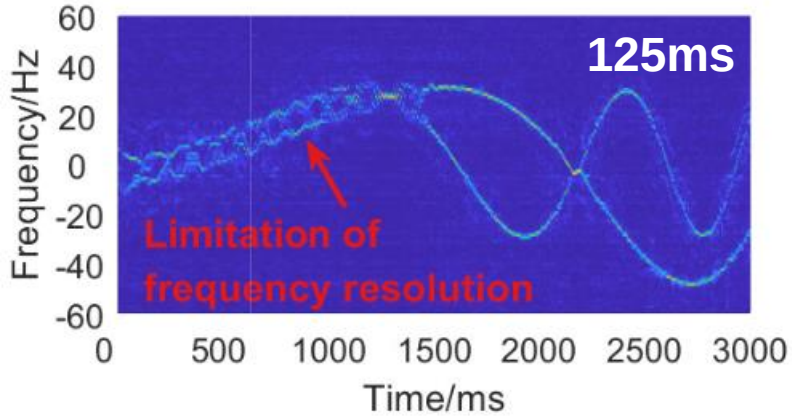
sliding window ↓ frequency resolution

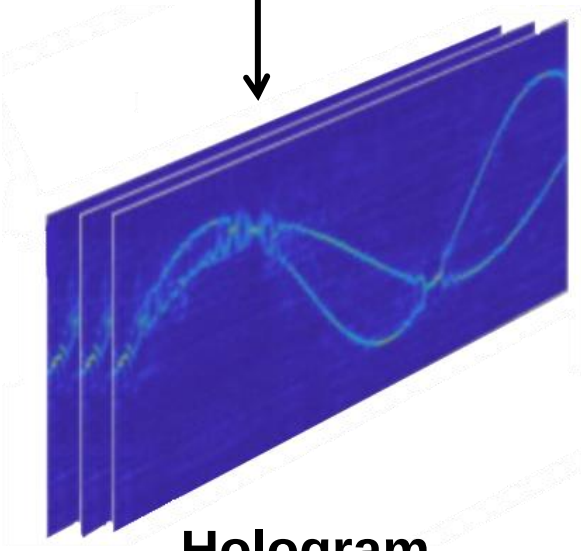
Design

Multi-Resolution Spectrogram Fusion



sliding windows with different lengths

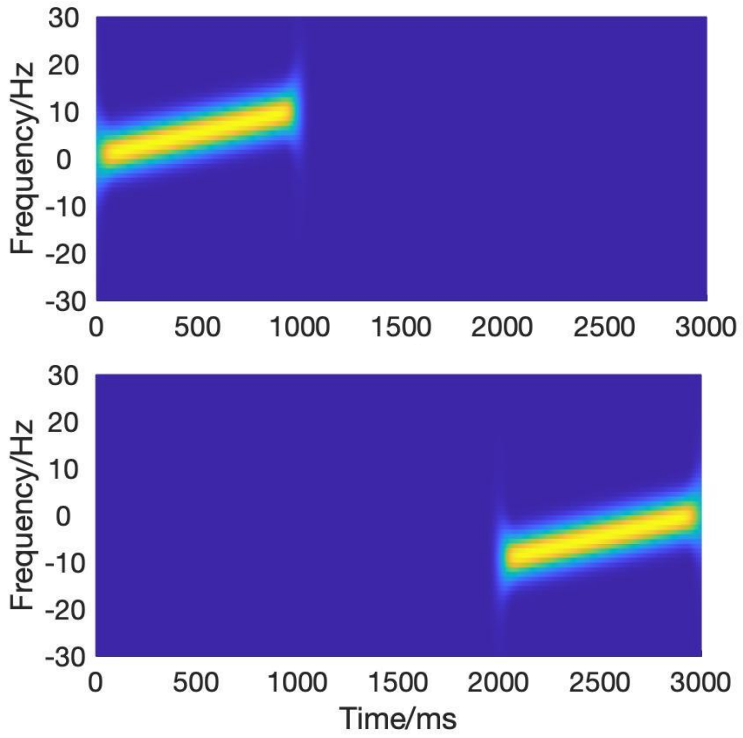




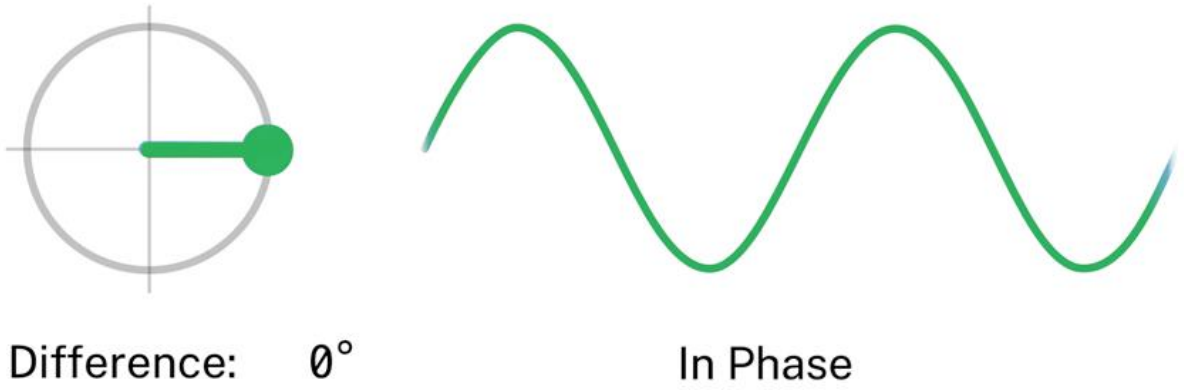
Hologram

Each sliding window's spectrogram is processed using a SEN pre-trained on synthesis spectrograms of the same length.

Design



Wireless signals require global discriminations.



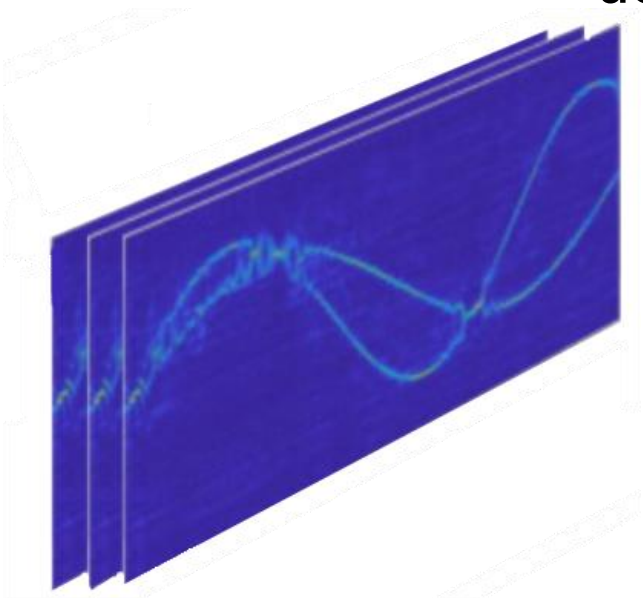
$$f(t)=A\sin(\omega t+\theta)$$

Phase refers to the offset of a waveform in time or space; different phase values result in the wave appearing at different positions.

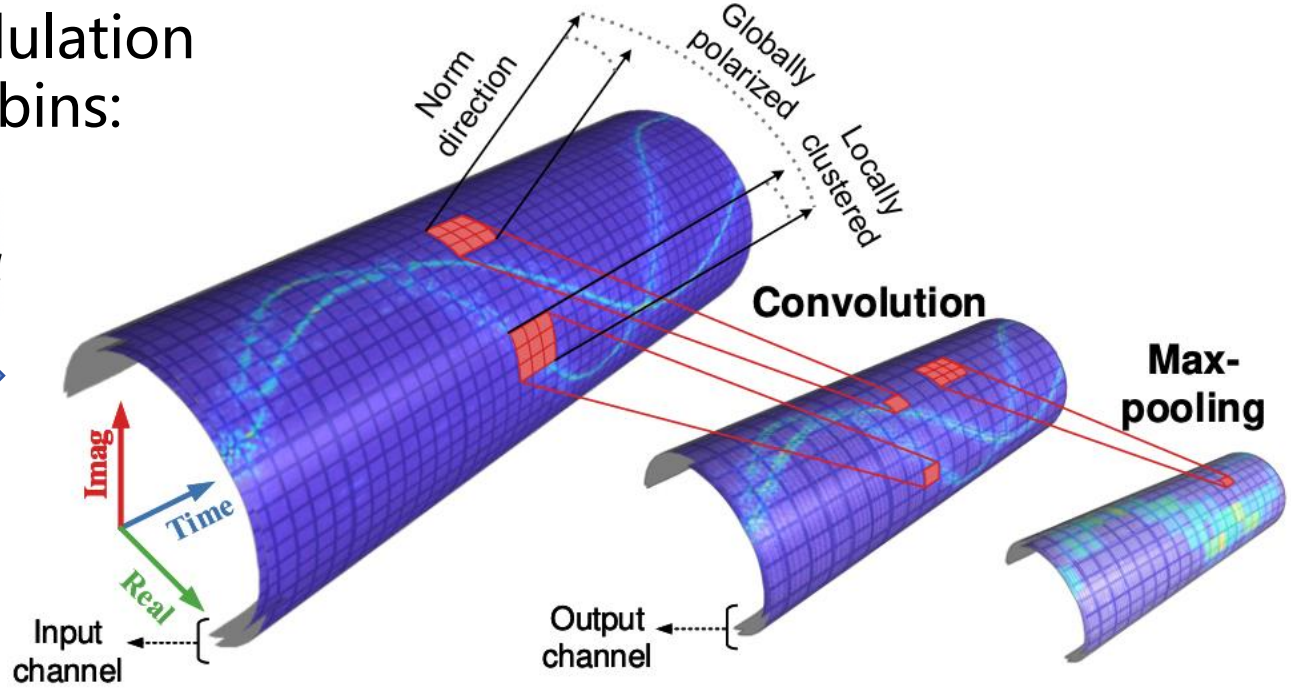
Design

Phase-polarized feature extraction

Linear phase modulation across frequency bins:



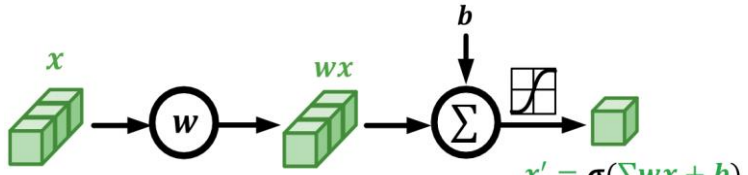
$$\phi_i = i\frac{\phi_h - \phi_l}{M} + \phi_l$$



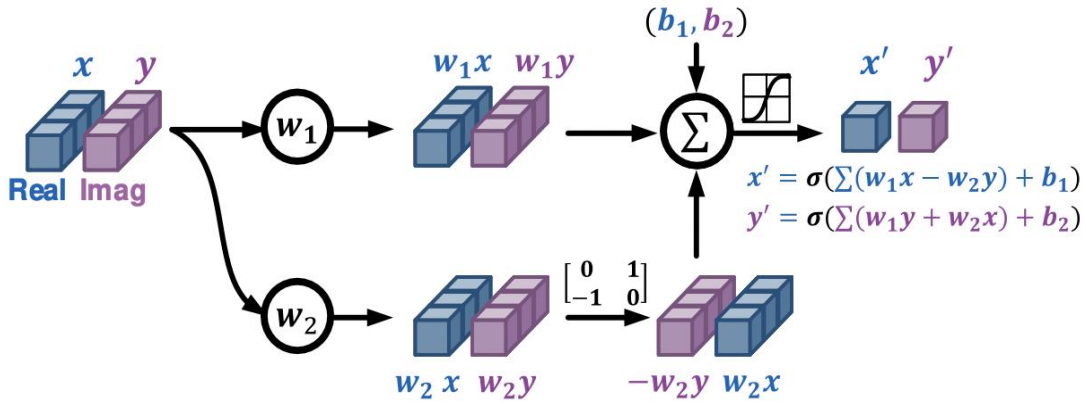
global discrimination is introduced along the frequency dimension

Design

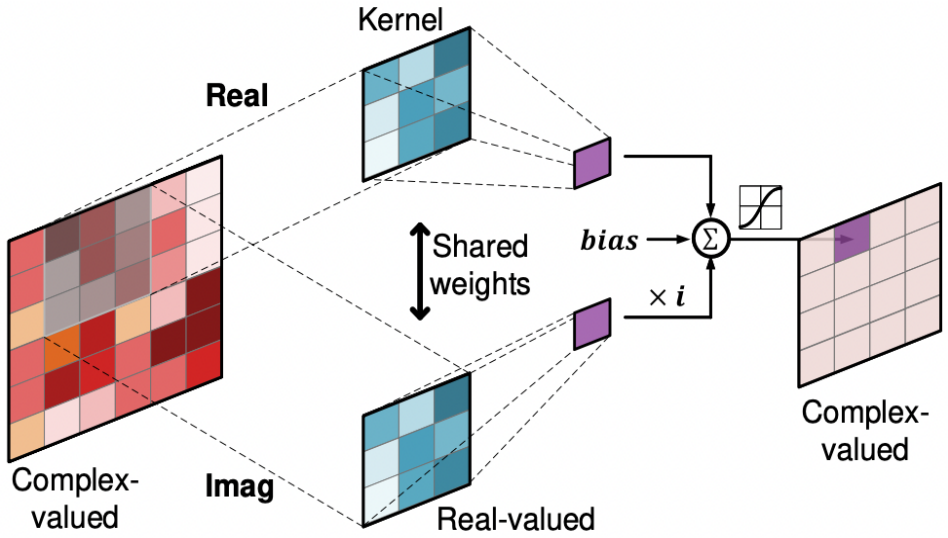
Complex-Valued Neural Network



Real-valued neurons



Complex-valued neurons



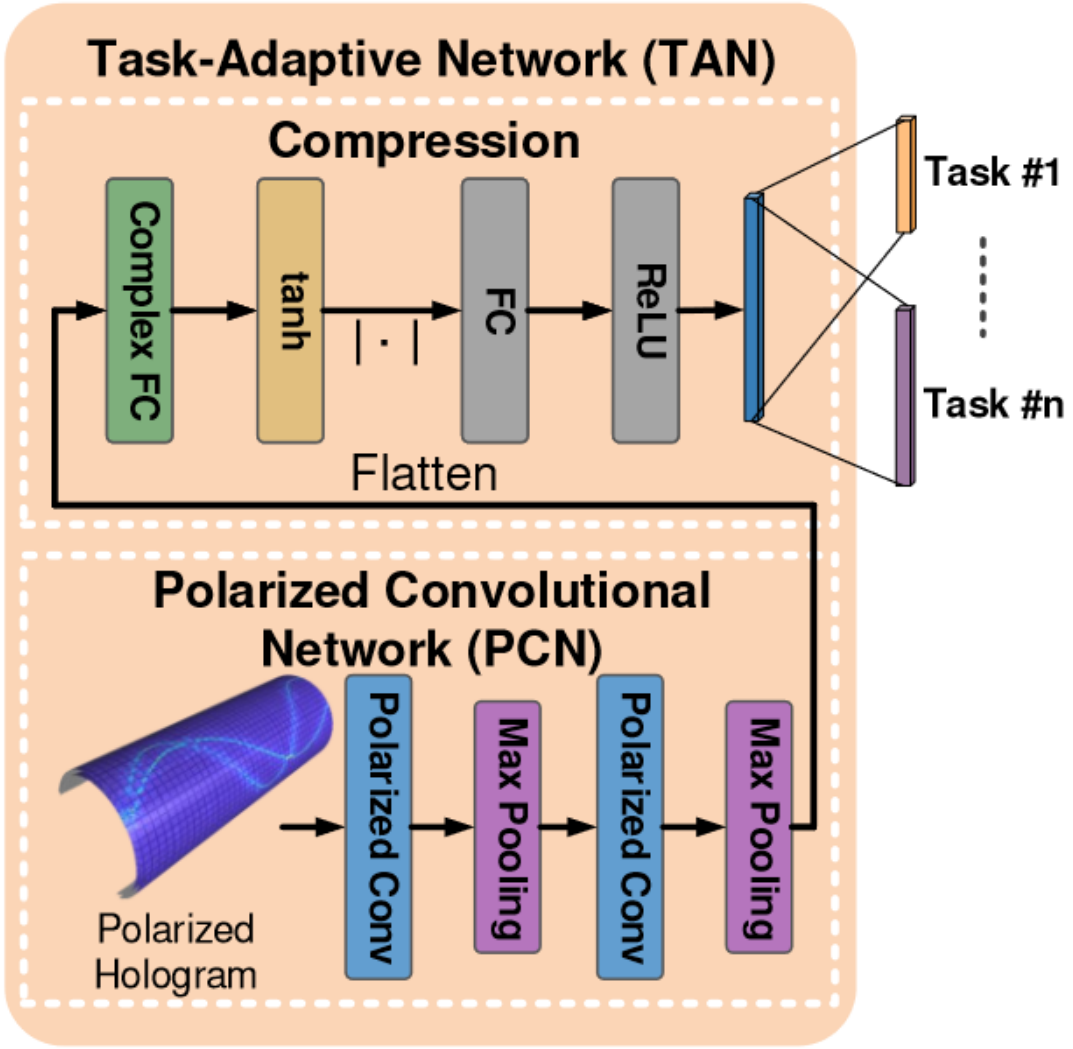
Complex-valued convolution

Design

Task-Adaptive Network

SLNet calculates the absolute value of the output from the complex-valued FC layer and inputs it to the real-valued FC layer.

additional FC		TASK
+ softmax	⟷	gesture classification
+ sigmoid	⟷	Fall detection



Evaluation

- **Hardware**:

  Intel 5300 WiFi Network Interface Cards (NICs)

  NVIDIA GeForce 2080Ti GPU

  Ubuntu 10.04

  Linux CSI Tool（log CSI readings）

- **Dataset**:

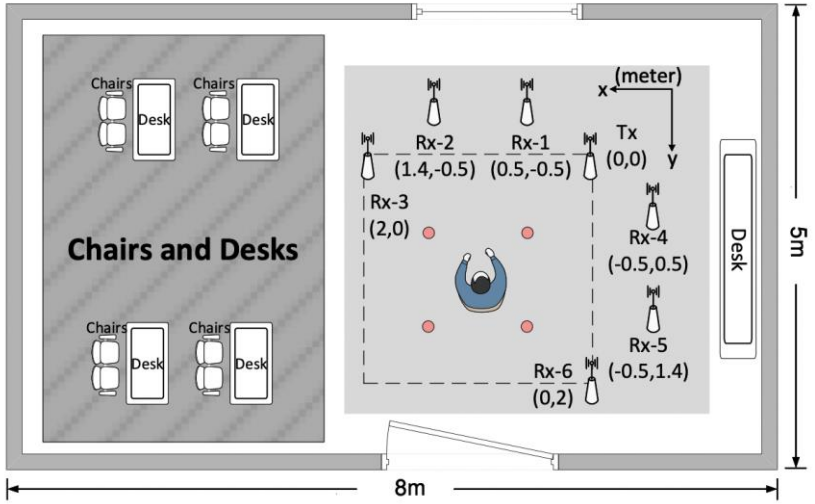
  Widar 3.0: WiFi-based Activity Recognition Dataset（Gait&Gesture）

  https://ieee-dataport.org/open-access/widar-30-wifi-based-activity-recog

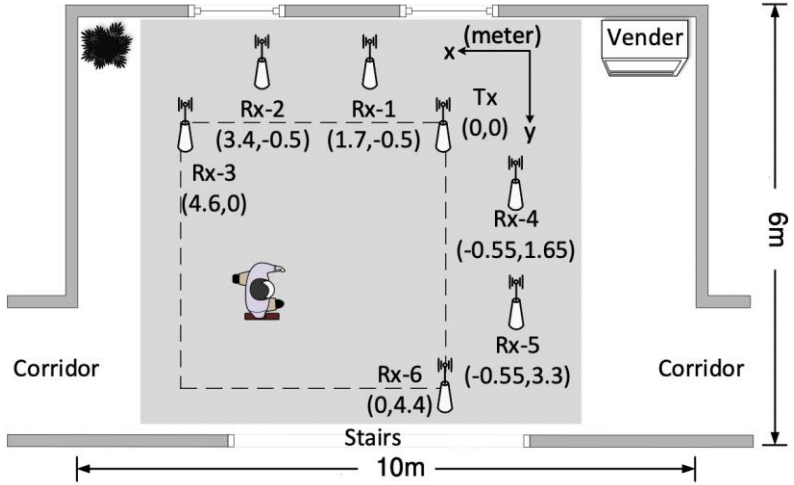
- **Code:**

  https://github.com/SLNetRelease/SLNetCode

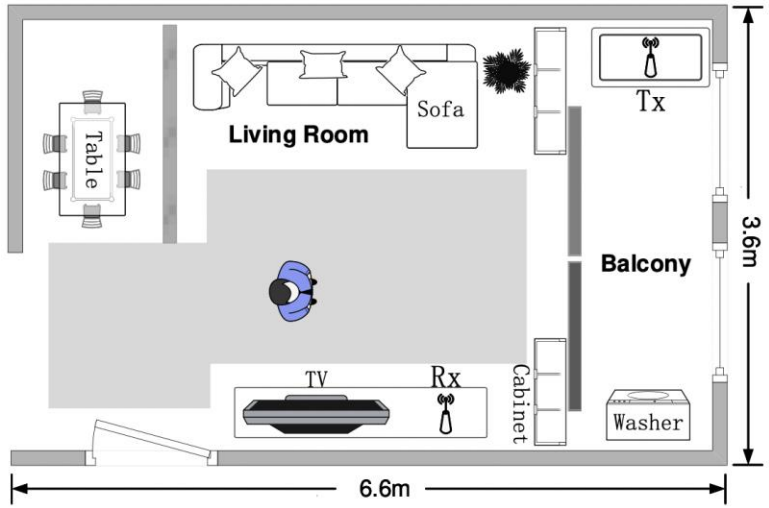
Evaluation



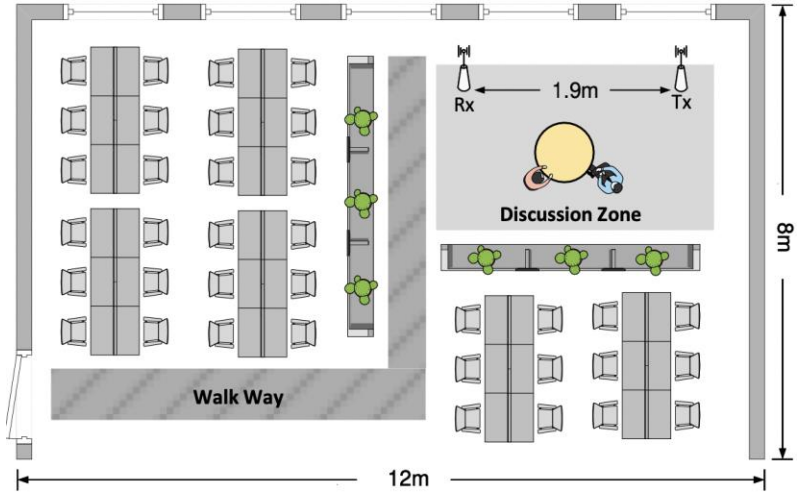
Human gesture recognition with Wi-Fi



Human gait identification with Wi-Fi



Fall detection within home environment with Wi-Fi

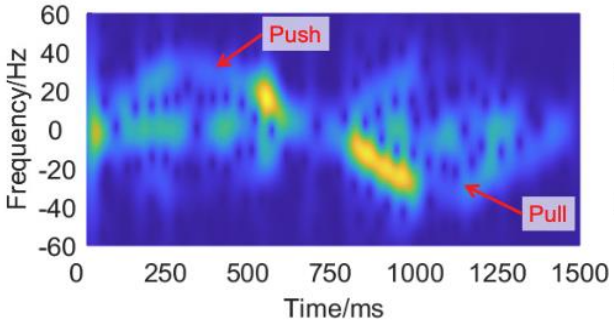


Monitoring the breath rate of two users with Wi-Fi

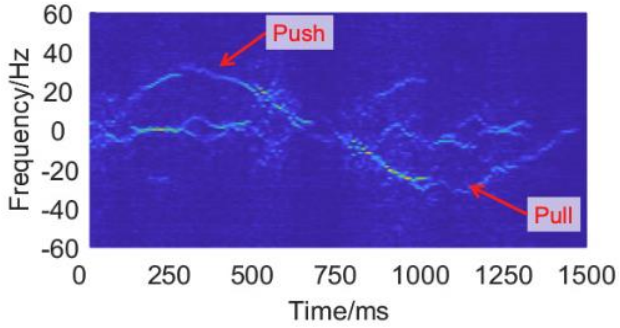
Evaluation

Human gestures

Raw FFT

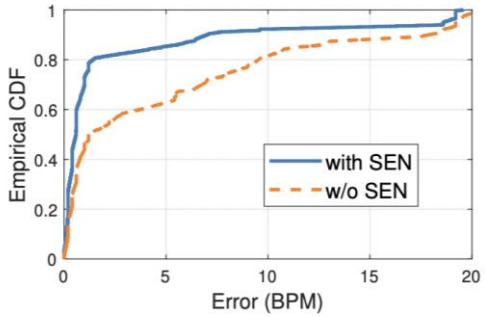


Enhanced

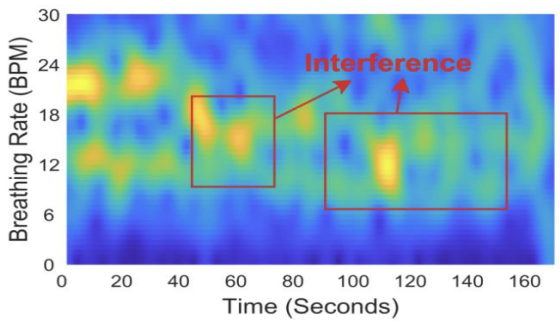


Human breath

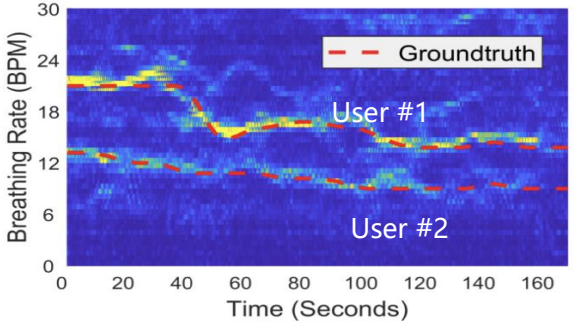
Estimation Error



Raw FFT



Enhanced

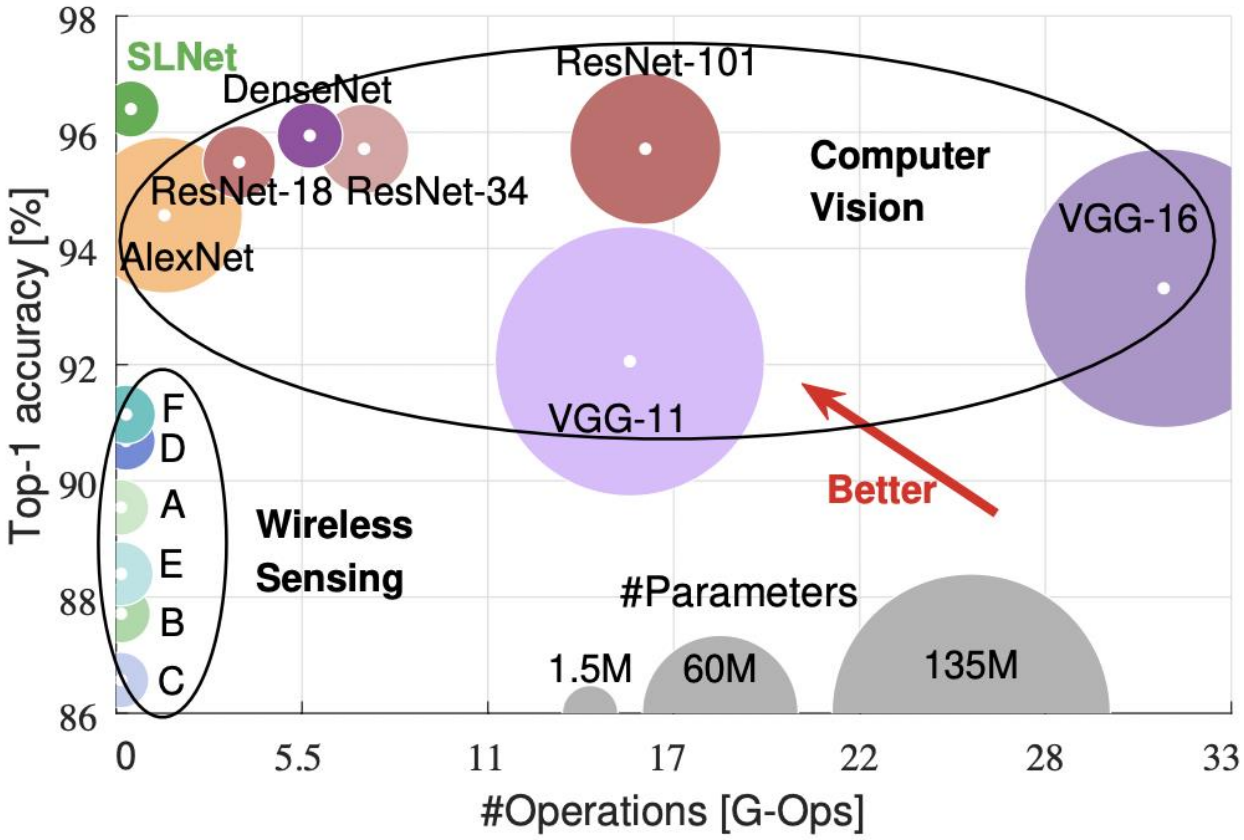


Frequency components becomes more distinguishable.

Evaluation

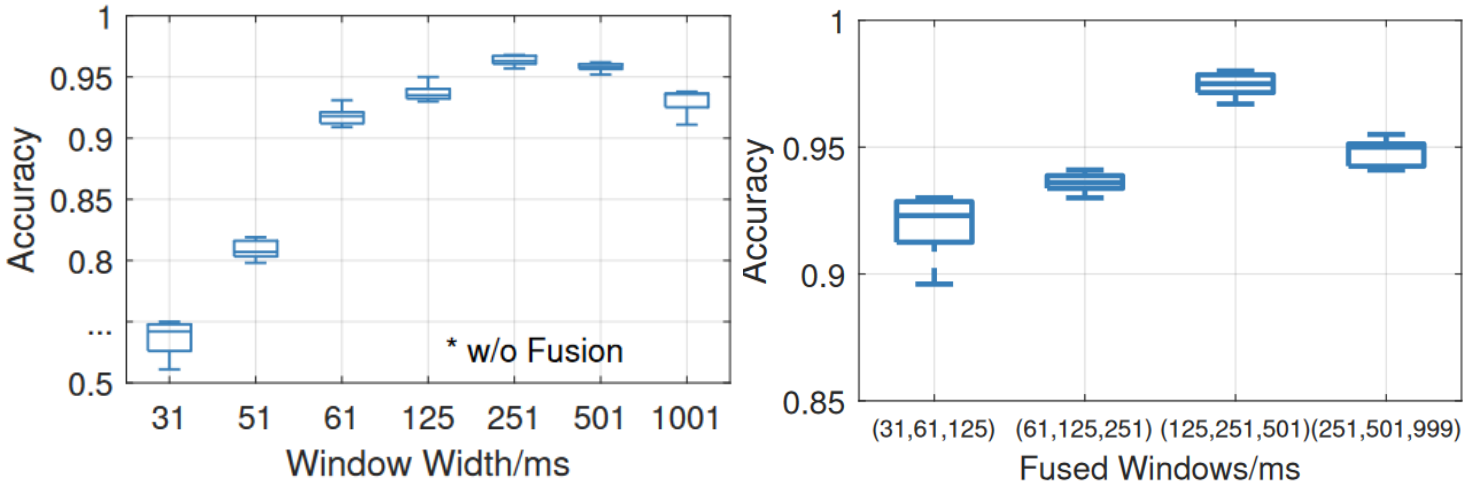
Recognition tasks

Modality	Ref.	Gesture	Gait	Fall[1]	Para[2]
WiFi	[23, 90]	90.6%	95.1%	92.8%, 96.3%	1.07M
	[8, 22]	89.0%	96.6%	96.4%, 84.3%	2.72M
	[39, 79]	84.3%	83.3%	96.8%, 93.8%	5.77M
	[73][3]	78.9%	70.9%	95.5%, 96.8%	0.06M
FMCW	[87]	88.0%	95.4%	96.0%, 96.0%	1.06M
	[84, 86]	91.6%	96.4%	99.7%, 95.7%	2.76M
Acoustic	[30]	89.6%	95.4%	90.6%, 98.3%	6.08M
Vision	[40]	88.3%	90.1%	95.3%, 95.3%	128.8M
	[15]	91.9%	96.6%	97.0%, 95.6%	11.18M
	[20]	91.0%	97.7%	99.8%, 96.3%	6.96M
CVNN	[17, 32]	72.3%	96.0%	95.2%, 93.7%	115.6M
	[46]	92.0%	96.3%	98.4%, 93.8%	2.94M
WiFi	**SLNet**	**96.6%**	**98.9%**	**99.8%, 97.2%**	**1.48M**

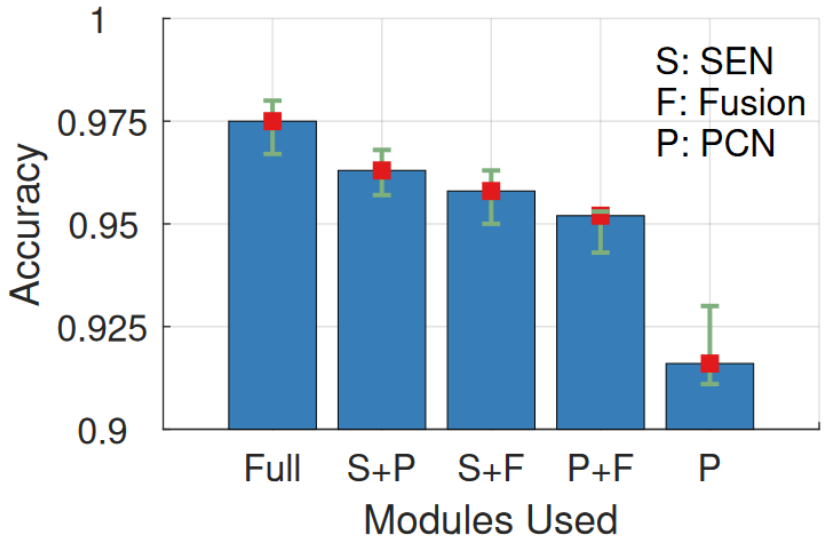


Less parameters, but higher performance.

Evaluation



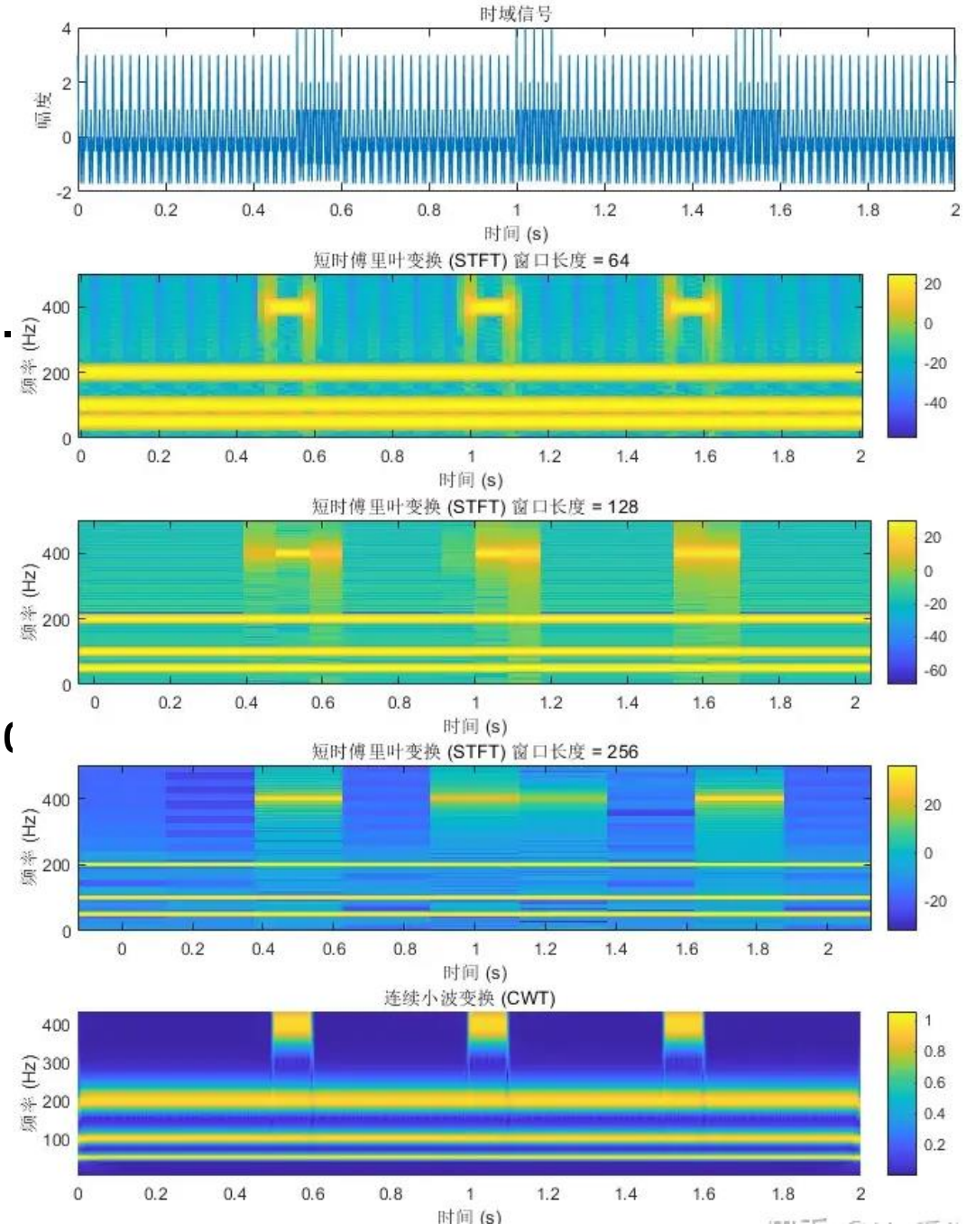
A single-resolution spectrogram for wireless sensing tasks is not the optimal solution



Spectral leakage problem cannot be neglected in wireless sensing tasks

idea

- Utilize **phase** information in signal processing.
- Mitigate the impact of **spectral leakage** on spectral analysis.
- Use **wavelet transform** to solve the multi-win problem.



请老师和同学们批评指正！

汇报人：陈宇杰