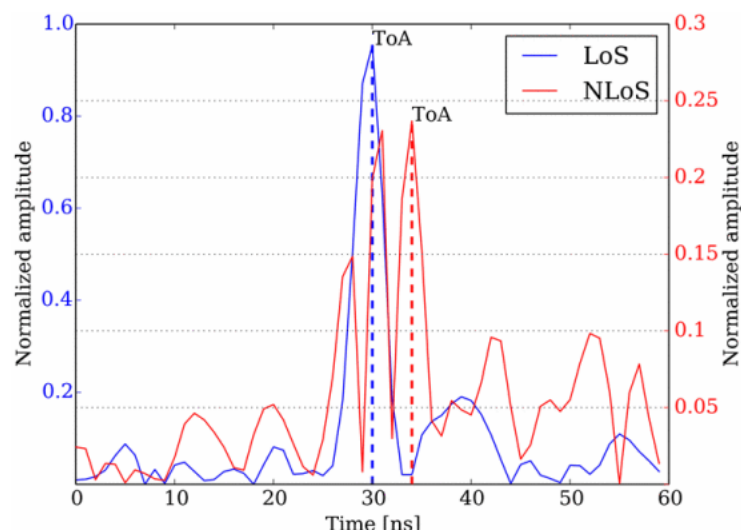
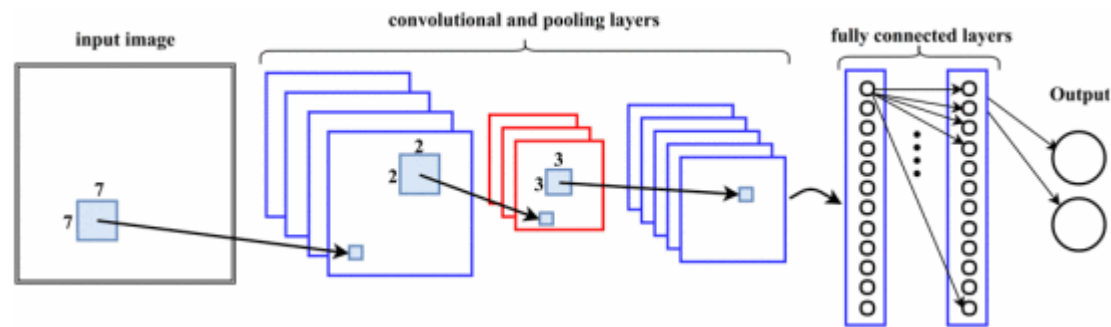




# *TDOA 논문 요약*

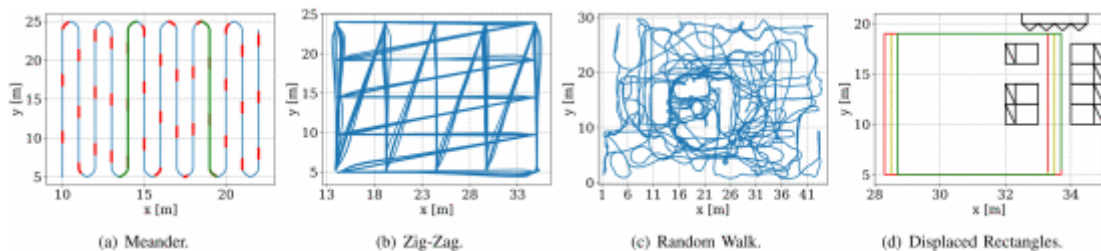


인풋 이미지 예시



모델 구조 예시

- CIR을 통해 왼쪽 그림과 같은 그래프를 생성
- 해당 이미지를 인풋으로 하는 CNN 모델을 거쳐 타겟의 위치를 아웃풋으로 도출

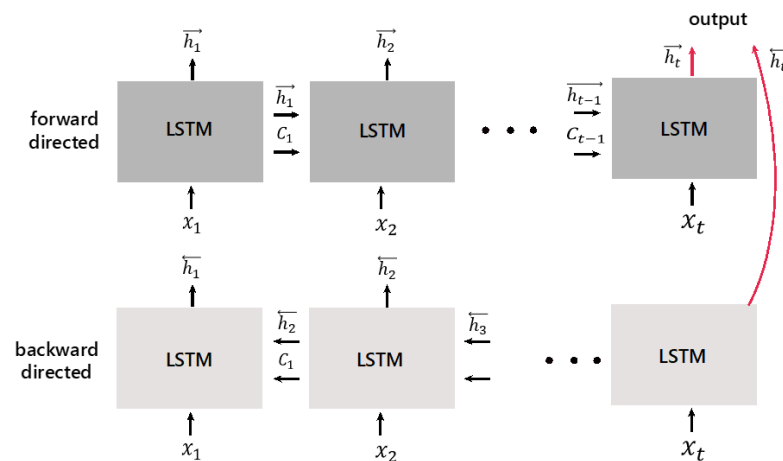


Dataset	CEP	CE95	MAE
Meander	15.6cm	36.4cm	17.4cm
Zig-Zag	24.1m	67.3m	29.1cm
Random Walk	30.3cm	86.8cm	36.2cm
Displaced Rectangles	10.2cm	24.3cm	11.6cm

- 실험 과정에서 다양한 형태의 이동 형태를 보여줌
- 해당 실험의 결과는 다음과 같이 나왔다
- 이때 상황은 Noise가 전혀 존재하지 않는 개활지에서 진행되었다
- 이러한 환경을 본인의 실험으로 적용한 결과 CNN은 50cm, Conv-lstm은 12cm 정도의 성능을 보였다

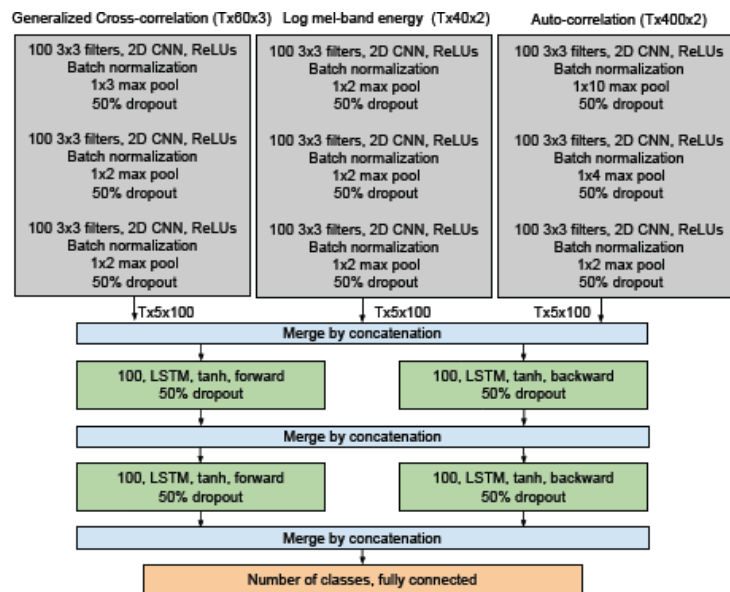
- Time-Frequency Masking : 오디오 부분을 강화, 축소 또는 격리하기 위해 시간 주파수 표현의 노이즈에 가중치를 적용하는 프로세스
  - => 해당 방법을 활용해 GCC-PHAT, Harmonic-Percussive Source Separation에 활용
- GCC-PHAT : 주파수 영역에서 두 센서의 신호들의 상호 상관(신호 섞이고 변형)을 구한 다음 백색화(잡음을 없애는) 가중치를 곱하는 과정을 거쳐 신호의 스펙트럼 영향을 배제하는 과정
  - => 이때, 신호의 지연 정보가 들어 있는 위상만 남고 크기는 언제나 1로써 일정하게 유지
  - => 가중치를 위의 TF 마스크로 설정 가능
  - => 해당 방식을 기반으로 두 센서 사이의 최종적인 지연 시간을 구할 수 있다.
- 즉, noise를 제거하고 pure한 신호를 얻기 위한 방법!

- 해당 논문은 Steering Vectors(위상 지연(필터에 신호가 입력되고 같은 신호가 출력될 때까지의 지연시간)을 모음 벡터)에서 TDoA 값을 구하려고 함
- 이때 신호의 파형 반복 및 기타 노이즈에 의해 변형된 신호를 TF-Masking GCT-PHAT 수식을 통해 지연 시간(TDOA)를 구함
- 본 논문에서는 pure한 신호를 구하는데 필요한 TF mask를 BLSTM(이후의 step 또한 앞의 step에 영향을 주는 lstm)을 이용해 학습
- 해당 방식은 SNR(신호와 잡음 비율)이 낮은 복잡한 환경에서 높은 강건성을 보였다.



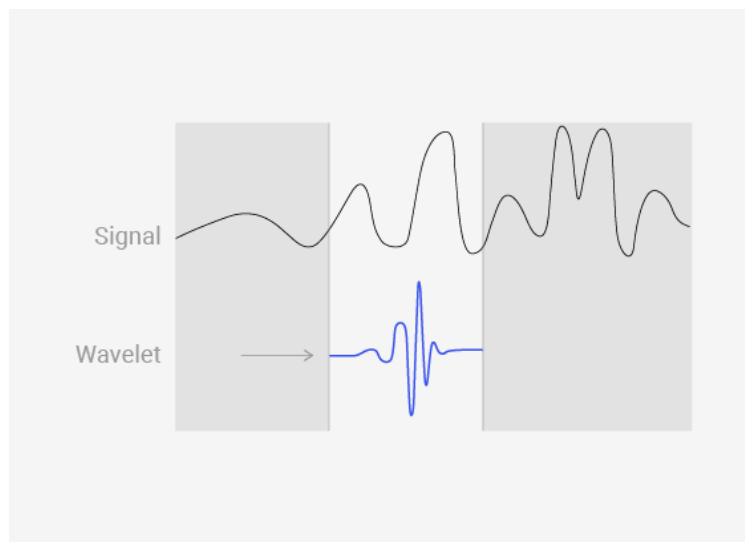
- Binaural Mel-Band Energies : 서로 다른 주파수 대역에서 거리에 따라 각 파장마다 다른 세기를 보이는 현상
- 파장의 조합을 알면 세기의 차이를 조사하여 신호로부터 거리 계산이 가능하다
- 사운드 이벤트 감지(SED)를 위한 신호 분리를 위해 기본 신호에서 GCC-PHAT을 이용
- Cross correlation : 신호처리 분야에서는 단지 두 신호가 얼마나 형태적으로 유사한지를 수치적으로 계산, 두 신호를 어떻게 겹치는지(displacement)에 따라서 pairing 관계가 달라진다.

- 모델에서는 Cross-correlation, Long mel-band, Auto correlation 을 인풋으로 한다
- 이때 각각에서 Convolution 과정을 거친 후, 특징 벡터를 추출해 이를 모아 LSTM 모델에 들어간다.
- 이때 논문에서 Time series는 5이며 아웃풋은 모든 사운드 이벤트 클래스에 대한 확률이 나와있는 벡터 (다수의 음향 이벤트가 발생하는 환경에서 이를 인식하고 각각의 발생과 소멸 시점을 판단하는 기술)





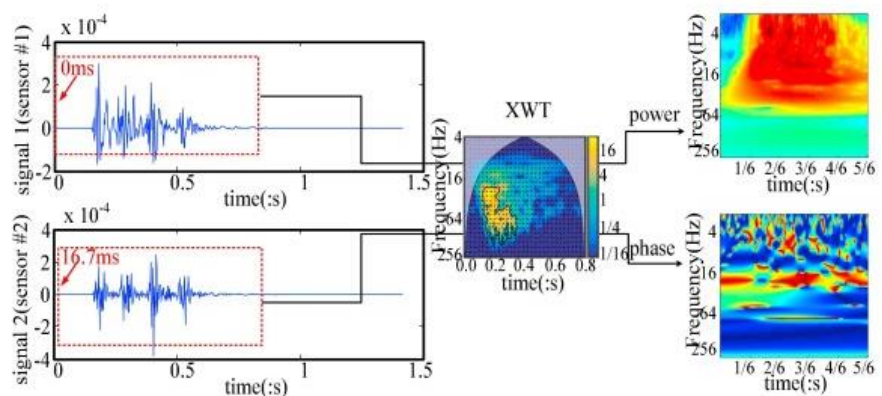
- 웨이블릿(wavelet) : 0을 중심으로 증가와 감소를 반복하는 진폭을 수반한 파동 같은 진동을 말한다.  
Ex) 지진계나 심박 체크에 기록되어 보이는 것과 같은 전형적인 "짧은 진동"의 형태
- 웨이블릿 변환(Wavelet Transform) : 임의의 신호를 웨이블릿(Wavelet)으로 정의되는 함수들로 분해하는 방법입니다. 푸리에 변환(Fourier Transform)이 무한히 진동하는 sine, cos 함수를 기저함수로 사용해 신호를 분해하는 것과는 달리 웨이블릿 변환은 진동하는 시간이 제한되는 함수를 기저함수로 사용해 시간 당 포함되는 주파수 성분의 크기를 보게 됩니다.



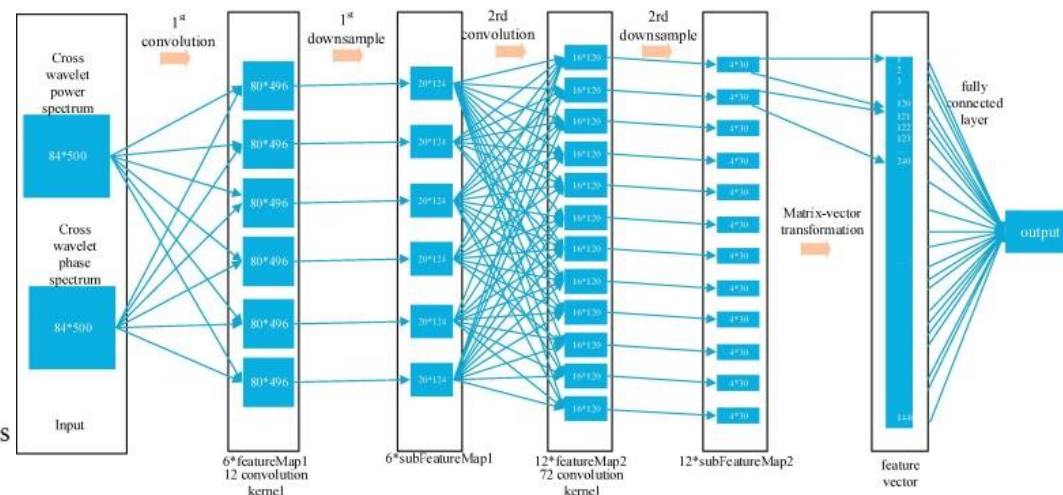




- 본 논문은 먼저 센서에서 측정된 두 신호(예 : 지진의 p파, s파)를 교차 웨이블릿 변환(Cross Wavelet Transform)을 진행하여 'power' 이미지와 'phase' 이미지로 변환
- 그런 다음 두 이미지를 인풋으로 하는 CNN 모델에 넣어 특정 사건(지진 같은)에서 측정된 여러 신호파가 도달하는 시간 차이를 구하게 된다.
- 그렇게 구한 신호 차이를 이용해서 사건이 일어난 위치 좌표를 추론
- 해당 실험을 통해 미세 지진원 위치 식별을 더욱 정교하게 강건하게 구할 수 있었다.

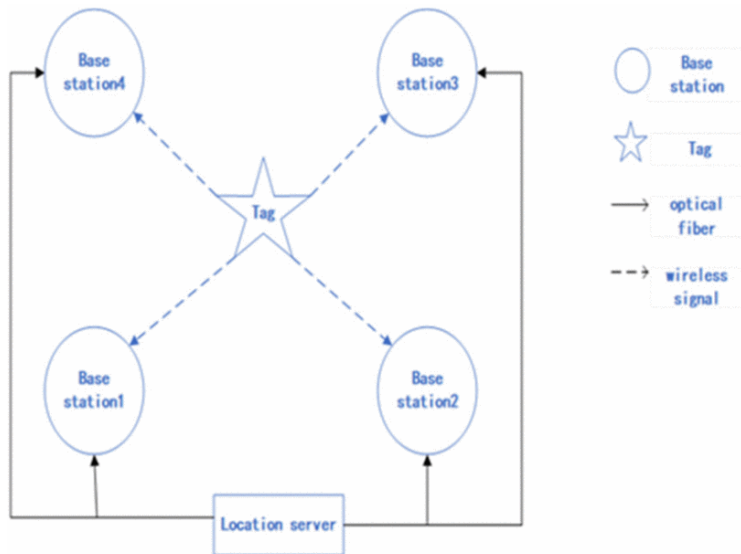


(b) The power and phase spectra of cross wavelet transform between two signals with 16.7ms delay





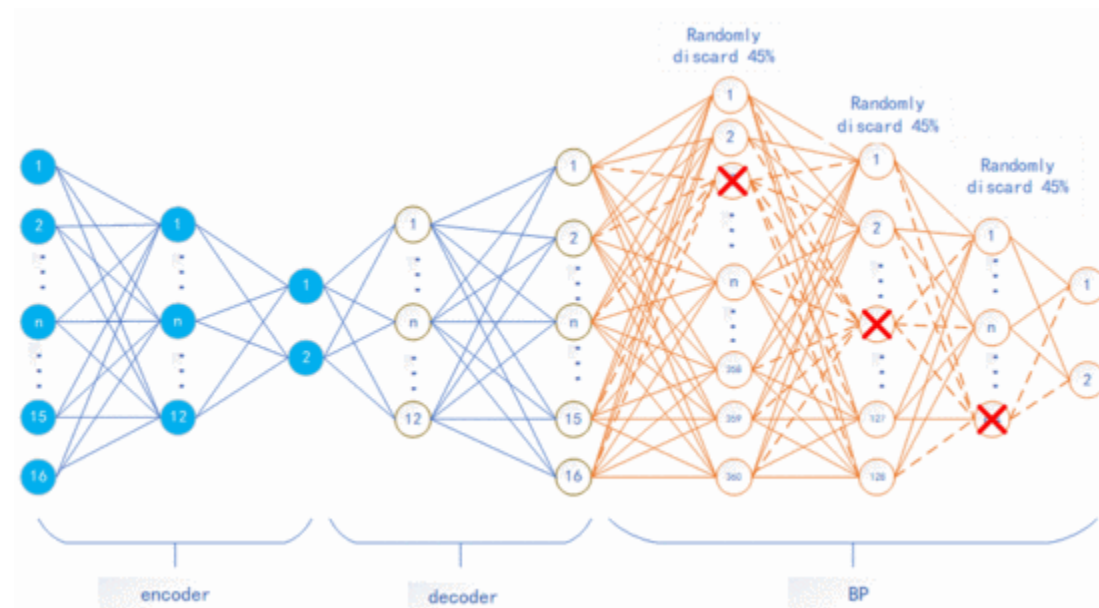
- 본 논문은 신호기의 좌표와 측정한 TDOA 값을 인풋으로 하고 태그의 위치 좌표를 리턴하는 모델
- 실험은 빈 공간에서 사각형 모양으로 모양을 만든 다음 진행한다.
- 타겟은 이동하지 않고 고정된 좌표에서 얻은 신호를 바탕으로 데이터셋을 구축한다.



Base station1 (Slave)			Base station2 (Slave)			Base station3 (main)			Base station4 (Slave)			TDOA				Label real coordinates		
X	Y	Z	X	Y	Z	X	Y	Z	X	Y	Z	s1-s3	s2-s3	s3-s4	s4-s1	X	Y	Z
22.302	3.555	2.868	15.24	3.555	2.868	21.905	7.731	2.909	15.382	7.731	2.868	-2.12381	1.133829	0	1.783097	20.702	3.555	1
22.302	3.555	2.868	15.24	3.555	2.868	21.905	7.731	2.909	15.382	7.731	2.868	-2.07709	1.248502	0	1.884232	20.702	3.555	1
22.302	3.555	2.868	15.24	3.555	2.868	21.905	7.731	2.909	15.382	7.731	2.868	-2.13919	1.184917	0	1.851934	20.702	3.555	1
22.302	3.555	2.868	15.24	3.555	2.868	21.905	7.731	2.909	15.382	7.731	2.868	-2.1895	0.963719	0	1.902693	20.702	3.555	1
22.302	3.555	2.868	15.24	3.555	2.868	21.905	7.731	2.909	15.382	7.731	2.868	-2.21732	1.171066	0	1.881963	20.702	3.555	1
22.302	3.555	2.868	15.24	3.555	2.868	21.905	7.731	2.909	15.382	7.731	2.868	-2.33764	1.150257	0	1.841645	20.702	3.555	1
22.302	3.555	2.868	15.24	3.555	2.868	21.905	7.731	2.909	15.382	7.731	2.868	-2.14684	1.131835	0	1.767489	20.702	3.555	1
22.302	3.555	2.868	15.24	3.555	2.868	21.905	7.731	2.909	15.382	7.731	2.868	-2.25189	1.200168	0	1.918839	20.702	3.555	1
22.302	3.555	2.868	15.24	3.555	2.868	21.905	7.731	2.909	15.382	7.731	2.868	-2.12501	1.247086	0	1.810173	20.702	3.555	1
22.302	3.555	2.868	15.24	3.555	2.868	21.905	7.731	2.909	15.382	7.731	2.868	-2.24661	1.03572	0	1.905558	20.702	3.555	1



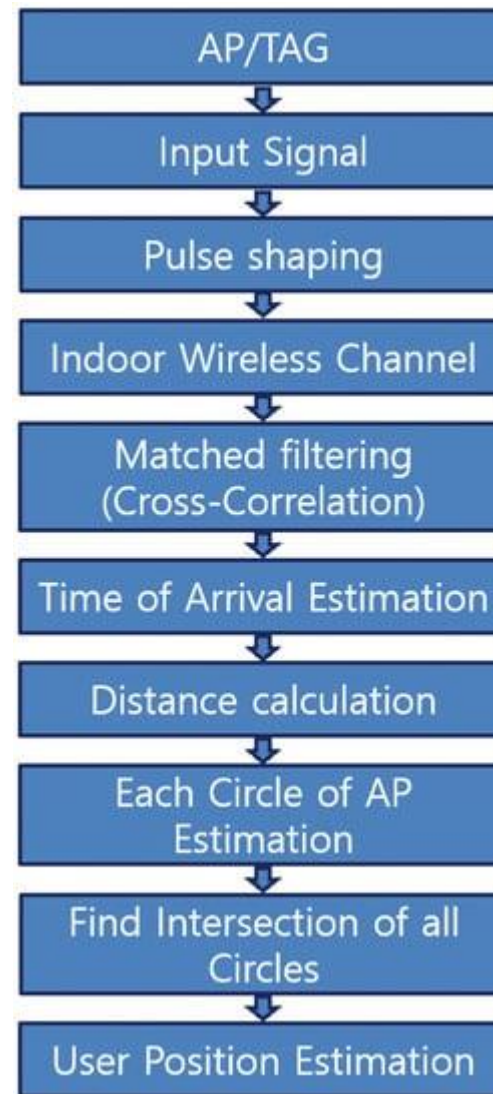
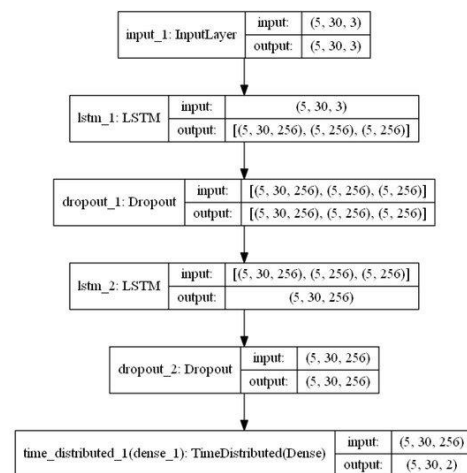
- 모델은 앞서 구한 데이터 인코딩을 거쳐 BP 모델에 들어가 x좌표 y좌표를 도출한다
- 비교군으로 실험한 CNN 모델의 convolution과 pooling 과정에서 중요한 정보를 손실하는 경우가 있다.
- 즉 인코딩 디코딩 과정을 통해 안정적으로 특성을 뽑는 결과를 보였다.





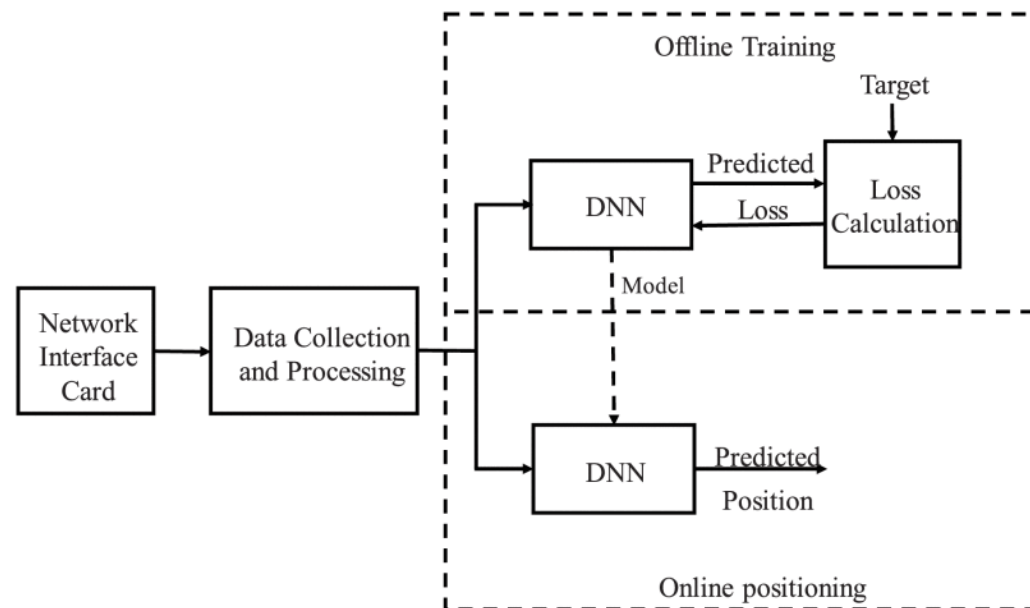
- 본 실험은 관측한 UWB 신호에서 Cross correlation을 이용한 신호 간 분리 작업을 통해 ToA 데이터를 구한다.
- 즉, 수신기(앵커)로부터 얻은 신호를 여러 신호들로 분리한 뒤, 이러한 신호들이 도착한 시간을 구한다.
- 이를 앵커당 분리한 신호들의 도착 시간을 다음과 같은 행렬로 나타낼 수 있다.
- 본 논문에서는 이렇게 구한 D 행렬을 인풋으로 하는 LSTM 모델을 구축(윈도의 크기는 5)
- 하이퍼파라미터 튜닝 과정을 거치는 과정을 실험으로 진술했으며 그 결과 LOS 문제를 해결하고 기존보다 더 나은 측위 정확도를 제공

$$D_{x,y} = \begin{bmatrix} D_{1-1}, D_{1-2}, \dots, D_{1-M_1}, \\ D_{2-1}, D_{2-2}, \dots, D_{2-M_2}, \\ \dots, \\ D_{N-1}, D_{N-2}, \dots, D_{N-M_N} \end{bmatrix}$$



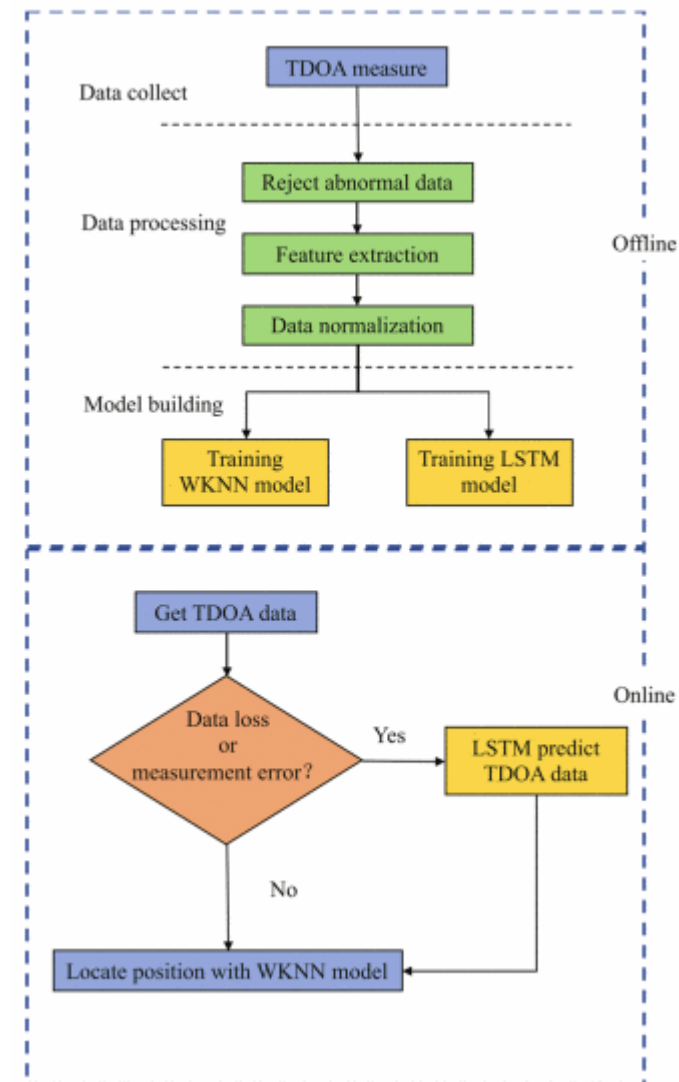


- CSI : 채널 상태 정보
- RSS는 수신기에서 수신되는 실제 신호 전력 강도
- 본 논문은 CSI와 RSS를 사용
- 특정 신호 주기를 기준으로 잘라 한 신호에 대해 채널 생성
- 앞서 구한 데이터를 바탕으로 1D-CNN을 구축
- 해당 방식을 통해 대상의 위치 추적
- CSI 정보를 사용하는 1D-CNN이 간단한 모델로도 훌륭하게 사용 가능한 모델을 만들 수 있었다.





- 본 모델은 측정된 TDOA 값을 이용해서 위치 좌표 추정
- 실험은 실제 실험실에서 앵커 4개를 이용해서 태그를 움직이면서 실험 데이터를 구축
- 이때 아예 너무 다른 데이터는 제거해서 사용
- LSTM은 측정된 TDOA 데이터를 수정하여 오류를 줄이는 데 사용
- 수정된 TDOA 데이터는 위치 좌표를 얻기 위한 Weighted KNN 모델로 전송
- 사전에 학습된 Weighted KNN 모델을 통해서 타겟의 위치 좌표 추정치를 리턴







- 본 논문은 실내 공간에서 타겟이 행동한 동작을 예측하는 것을 목표로 하고 있다.
- 예를 들어 요리를 진행하면 그 순간 미세먼지 농도가 급격하게 증가한다. 이러한 특성에 주목해 여러 센서들을 이용해 습도, 온도, 미세먼지 농도이며 데이터를 수집하고 행렬로 표현한다.
- 본 논문에서는 윈도우를 40으로 세팅하여 차원 데이터셋으로 변환
- 변환한 데이터를 Convolutional-LSTM 신경망에 넣어 동작 예측(7종류)
- 그 결과 다음과 같이 conv-lstm 모델이 가장 좋은 성능을 보였다.

**print** a shape of  $S_{3D}$

```
S3D = {[[[H11, T11, PM2511, PM1011],  
[H12, T12, PM2512, PM1012],  
....  
[H140, T140, PM25140, PM10140]],  
[[[H21, T21, PM2521, PM1021],  
....  
[H22, T22, PM2522, PM1022],  
[H240, T240, PM25240, PM10240]],  
...  
...  
[[[Hn1, Tn1, PM25n1, PM10n1],  
[Hn2, Tn2, PM25n2, PM10n2],  
....  
[Hn40, Tn40, PM25n40, PM10n40]]]}
```

표 7 학습 모델별 행동 인식 정확도

Model	Accuracy (Epochs=25)	Accuracy (Epochs=50)	Accuracy (Epochs=75)	Accuracy (Epochs=100)
CNN	50.127%	53.531%	59.201%	66.092%
LSTM	75.359%	82.211%	86.554%	90.199%
FC-LSTM	63.237%	77.620%	84.049%	88.484%
Convolutional LSTM	77.261%	82.012%	88.247%	92.151%



- Niitsoo, A., Edelhäußer, T., & Mutschler, C. (2018, September). Convolutional neural networks for position estimation in tdoa-based locating systems. In *2018 International Conference on Indoor Positioning and Indoor Navigation (IPIN)* (pp. 1-8). IEEE.
- Wang, Z. Q., Zhang, X., & Wang, D. (2018, August). Robust TDOA Estimation Based on Time-Frequency Masking and Deep Neural Networks. In *Interspeech* (pp. 322-326).
- Adavanne, S., Pertilä, P., & Virtanen, T. (2017, March). Sound event detection using spatial features and convolutional recurrent neural network. In *2017 IEEE international conference on acoustics, speech and signal processing (ICASSP)* (pp. 771-775). IEEE.
- Huang, L., Li, J., Hao, H., & Li, X. (2018). Micro-seismic event detection and location in underground mines by using Convolutional Neural Networks (CNN) and deep learning. *Tunnelling and Underground Space Technology*, 81, 265-276.
- Ye, X., & Zhang, Y. (2020, December). Research on UWB positioning method based on deep learning. In *2020 5th International Conference on Mechanical, Control and Computer Engineering (ICMCCE)* (pp. 1505-1508). IEEE.
- Poulose, A., & Han, D. S. (2020). UWB indoor localization using deep learning LSTM networks. *Applied Sciences*, 10(18), 6290.
- Zandian, R., & Witkowski, U. (2018, October). Differential NLOS Error Detection in UWB-based Localization Systems using Logistic Regression. In *2018 15th Workshop on Positioning, Navigation and Communications (WPNC)* (pp. 1-6). IEEE.
- Hsieh, C. H., Chen, J. Y., & Nien, B. H. (2019). Deep learning-based indoor localization using received signal strength and channel state information. *IEEE access*, 7, 33256-33267.
- Wang, H., Wang, X., Xue, Y., & Jiang, Y. (2020, June). UWB-based Indoor Localization Using a Hybrid WKNN-LSTM Algorithm. In *2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)* (Vol. 1, pp. 1720-1725). IEEE.
- Xue, Y., Su, W., Wang, H., Yang, D., & Jiang, Y. (2019). DeepTAL: deep learning for TDOA-based asynchronous localization security with measurement error and missing data. *IEEE Access*, 7, 122492-122502.
- 조찬호. "Convolutional LSTM을 이용한 실내 행동 인식." 국내석사학위논문 高麗大學校, 2019. 서울