

# IoT 환경에서 실내 사람 수 추정을 위한 DNN 모델

김연주, 김승구\*

충북대학교

jj27kk@cbnu.ac.kr, \*kimsk@cbnu.ac.kr

## The DNN Model for Indoor People Counting in IoT Environments

Kim Yeon Ju, Kim Seungku\*

Chungbuk National Univ.

### 요약

IoT 환경에서 실내 사람 수의 추정은 스마트 시스템 실현을 위해 필요한 기술이다. 사람 수 추정을 위해 카메라, 센서, RFID 등을 활용한 다양한 방식이 있다. WiFi 신호의 CSI를 활용한 실내 인원 수를 추정에 관한 연구는 다른 방식 대비 경제성이며 사생활 침해의 우려가 없다는 장점이 있다. 본 논문은 WiFi 신호의 CSI를 활용하여 사람 수 추정의 정확도를 높이기 위한 새로운 DNN 모델을 제안한다. 4단의 은닉 계층을 갖는 DNN 모델을 설계하였고, 실험 결과 최대 92%의 정확도를 얻었다.

### I. 서론

실내 사람 수 추정은 스마트빌딩의 보안 및 에너지 관리와 같은 사람 중심의 사물인터넷(IoT) 시스템을 구축하기 위해 필요한 기술이다. 실내 인원 수를 추정하는 기법에 대한 연구는 몇 년 전부터 진행되어왔다. 카메라를 이용한 이미지 기반 방식, 센서나 RFID를 이용하는 비이미지 기반 방식 등 다양한 방안이 제안되었다. 하지만 카메라를 이용하는 방식은 사생활 침해의 우려가 있고, 조명, 장애물 등 환경의 영향을 크게 받는 문제가 있다. 또한 카메라, 센서, RFID를 이용하는 방식은 시스템을 설치하기 위한 추가 비용이 발생한다.

본 논문에서는 WiFi AP를 활용하여 추출한 Channel State Information (CSI)로부터 실내 인원 수를 추정한다. WiFi 신호를 이용하는 방식은 비이미지 기반으로 사생활 침해의 우려가 없으며, 새로 장치를 설치할 필요가 없어 경제적이다. [1]은 WiFi AP에서 수집한 CSI 데이터로부터 기계학습(ML)을 통해 원하는 정보를 얻기 위한 기법을 설명하며, CSI 데이터 세트[2]를 제공한다. [3]은 DNN 모델을 기반으로 CSI를 이용해 사람 수를 추정하는 방식을 제안한다. 알려진 DNN 모델을 사용하여 사람 수를 추정하였을 때 어떠한 시나리오에서는 낮은 정확도를 보인다. 본 논문에서는 IoT 환경의 다양한 실내공간 시나리오에서 사람 수 추정의 정확도를 높이는 새로운 DNN 모델을 제안한다.

섹션II에서는 연구 배경과 선행연구들에 대해 소개한다. 섹션III에서는 설계한 DNN 모델에 대해 자세히 기술한다. 섹션IV에서 제안하는 DNN 모델과 다른 모델의 성능을 비교하고 섹션V에서 마무리 짓는다.

### II. 배경

#### A. CSI(Channel State Information)

실내 환경에서 WiFi 신호는 반사, 산란, 회절을 거쳐 다중 경로 신호로 수신기에 전송된다.

$$h(t) = \sum_{i=1}^N \alpha_i(t) e^{-j\Phi_i(t)} \delta(\tau - \tau_i(t)) \quad (1)$$

$N$ 은 다중경로 신호 성분의 수,  $\alpha_i(t)$ ,  $\Phi_i(t)$ ,  $\tau_i(t)$ 는 각각  $i$ 번째 다중 경로 성분의 진폭, 위상, 지연시간이다. 대부분의 WiFi 장치는 OFDM 시스템을 사용하며, 수식(1)에서  $h(t)$ 의 고속 푸리에 변환(FFT)은 OFDM 시스템의 채널 주파수 응답 또는 물리계층 CSI에 해당한다. OFDM의 CSI는 아래와 같다.

$$H(n) = [H_1(n), \dots, H_k(n)] \quad (2)$$

$k$ 번째 서브캐리어에 대한 CSI는  $H_k(n) = |H_k(n)| e^{j\angle H_k(n)}$ 로 나타내며, 여기서  $|H_k(n)|$ 는 진폭,  $\angle H_k(n)$ 는 위상이다.

#### B. 관련 연구

어떠한 장치도 가지고 다니지 않는 사람들의 활동을 WiFi 신호를 이용해 추적하는 것에 대한 연구가 몇 년 전부터 진행되어왔다. [1]은 WiFi 신호의 CSI로부터 원하는 활동 정보를 추출해내기 위한 피쳐 엔지니어링 기법과 ML 모델에 대해 자세히 기술하며, 이들을 종합적으로 검토한다. 또한 여러 실내환경에서 WiFi 측정값과 사람 수를 포함하는 개방형 데이터 [2]를 제공한다. [2](이하 EHUCOUNT)는 OFDM 신호의 53개의 서브캐리어에 대해 진폭과 위상 정보를 포함하는 채널상태정보를 제공한다. 본 논문에서는 EHUCOUNT를 이용해 사람 수 추정에 유용한 ML 모델을 설계한다.

[3]은 DNN 모델을 기반으로 하는 실내 인원 수의 추정 방식을 제안한다. sigmoid를 활성화함수로 사용하여 3단의 은닉층 레이어를 갖는 DNN 모델을 설계하고, 실제 실내에서 측정된 테스트 데이터로 그 성능을 검증한다. 실내 시나리오에서 최대 9명의 모집단의 수를 추정할 때 96.90%, 88.66%까지의 높은 정확도를 보인다. [3]에서 제안하는 모델로 EHUCOUNT dataset을 사용하여 사람 수를 추정했을 때 대부분의 시나리오에서 정확도가 낮은 경향을 보였다. 본 논문의 섹션III에서 이보다 높은 정확도를 갖는 새로운 모델을 설계하였으며, [3] 모델과 정확도를 비교하여 성능을 평가한다.

Scenario	Model	apoch					
		1	20	40	60	80	100
A	Legacy model [3]	0.247	0.567	0.652	0.721	0.747	0.767
	Proposed Model	0.284	0.684	0.748	0.773	0.807	0.826
B	Legacy model [3]	0.171	0.382	0.491	0.526	0.558	0.577
	Proposed Model	0.181	0.511	0.601	0.66	0.674	0.725
C	Legacy model [3]	0.194	0.309	0.646	0.71	0.755	0.784
	Proposed Model	0.213	0.693	0.758	0.818	0.847	0.874
D	Legacy model [3]	0.17	0.263	0.478	0.556	0.622	0.672
	Proposed Model	0.171	0.534	0.67	0.721	0.747	0.764
E	Legacy model [3]	0.2	0.297	0.538	0.703	0.796	0.84
	Proposed Model	0.271	0.748	0.824	0.89	0.909	0.932
F	Legacy model [3]	0.167	0.41	0.514	0.588	0.639	0.684
	Proposed Model	0.166	0.505	0.645	0.724	0.799	0.833

표 1. apoch 횟수에 따른 교차검증 정확도

### III. 사람 수 추정 DNN 모델 설계

제안하는 모델은 입력층과 은닉층, 출력층으로 구성된다. EHUCOUNT dataset은 53개 서브캐리어의 진폭과 위상 정보를 제공한다. 따라서 DNN 모델의 입력층은 그에 따라 106개의 노드로 구성하였다. 은닉층은 4단의 레이어를 갖는다. 활성화 함수로 3개의 레이어에는 swish를, 네 번째 레이어에는 relu를 사용했다. 각 레이어의 노드 수는 각각 100, 100, 50, 20개로 구성된다. 출력층에서는 softmax를 사용하여 분류한다. Learning rate는 초기값 0.004로 설정하여 Adam Optimizer를 사용한다. 손실 함수는 softmax cross entropy를 이용하였다.

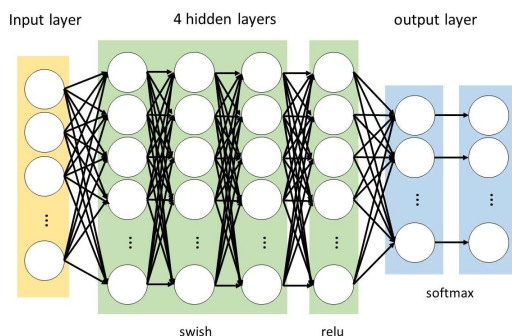


그림 1. DNN 모델 구조

### IV. 성능 평가

성능을 비교하기 위한 모델로 [3]을 사용하였다. EHUCOUNT는 106개의 CSI 정보를 제공하므로 그에 맞추어 [3]의 모델도 입력층의 노드를 106개로 설정하였다. 또한 learning rate는 초기값 0.004로 Adam Optimizer를 사용하였다. KFold를 이용해 K=5로 설정하여 학습했고, 정확도를 높이기 위해 반복 학습이 이루어졌다. 마지막으로 학습에 사용되지 않은 검증 데이터를 사용하여 최종적으로 accuracy를 구하고 평가하였다.

표1은 apoch에 따른 KFold 교차검증 결과이다. apoch가 늘어날수록 두 모델에서 정확도가 높아지고 새로운 모델에서는 더 나은 성능을 보인다. 그러나 같은 데이터 세트와 여러 번 학습하면 과대적합이 발생할 가능성이 있다. 표2는 EHUCOUNT dataset을 apoch를 100으로 설정하여 학습을 마친 후 최종 검증데이터로 예측하여 얻은 정확도이다. Sigmoid 활성화 함수를 사용한 DNN 모델과 새로 설계한 DNN 모델을 비교하였을 때 제안하는 모델을 사용한 경우 정확도가 4~14% 상승하는 결과를 얻었다.

Scenario	final test accuracy	
	Legacy model [3]	Proposed DNN
A	0.760.01	0.800.02
B	0.590.02	0.700.02
C	0.800.02	0.860.01
D	0.670.02	0.770.25
E	0.850.02	0.920.02
F	0.690.01	0.840.02

표 2. 최종 검증 정확도

sklearn에서 제공하는 KNN 라이브러리 함수를 이용하여 (k=1)로 학습시켰을 때 Scenario A에서 89%, Scenario B에서 80%의 높은 정확도를 보이나, 220~660sec의 긴 학습시간이 소요된다. 제안하는 DNN 모델은 KNN으로 학습했을 때보다 정확도는 떨어지지만 학습시간이 50~77sec로 개선되었다.

### V. 결 론

본 논문에서는 WiFi AP로 추출한 CSI로부터 사람 수를 추정하는 DNN 모델을 설계하였다. EHUCOUNT 데이터셋을 사용하여 평가하였을 때, [3]의 모델과 비교 결과 모든 시나리오에서 정확도가 상승하였다. KNN과 비교하였을 때는 학습시간이 빨라지지만 정확도가 다소 떨어진다.

### ACKNOWLEDGMENT

이 논문은 산업통상자원부 ‘산업전문인력역량강화사업’의 재원으로 한국 산업기술진흥원(KIAT)의 지원을 받아 수행된 연구임. (2020년 임베디드 SW 전문인력 양성사업, 과제번호 : N0001884)

### 참 고 문 헌

- [1] Sobron, I.; Ser, J.D.; Eizmendi, I.; Velez, M. "Device-Free People Counting in IoT Environments: New Insights, Results, and Open Challenges" IEEE Internet Things J. 2018
- [2] I. Sobron, J. Del Ser, I. Eizmendi, and M. Velez. (2017). EHUCOUNT Dataset. Accessed: Nov. 30, 2017. [Online]. Available: [www.ehu.eus/tsr\\_radio/index.php/research-areas/data-analytics-in-wireless-networks](http://www.ehu.eus/tsr_radio/index.php/research-areas/data-analytics-in-wireless-networks)
- [3] YK Cheng, RY Chang. "Device-Free Indoor People Counting Using Wi-Fi Channel State Information for Internet of Things" IEEE Global Communications Conference, 2017