

<http://dx.doi.org/10.17703/JCCT.2022.8.6.899>

JCCT 2022-11-111

UWB 레이더와 실내 환경 측정 센서를 이용한 랜덤 포레스트 모델의 재실활동 유형 감지

Activity Type Detection Of Random Forest Model Using UWB Radar And Indoor Environmental Measurement Sensor

박진수*, 정지성**, 양철승**, 이정기**

Jin Su Park*, Ji Seong Jeong**, Chul Seung Yang**, Jeong Gi Lee**

요약 본 세계적으로 출생률이 줄고 기대 수명이 늘어나 고령화 사회가 되어감에 따라 고령 인구의 건강 관리를 위한 시스템이 필요하다. 그 중 실내 건강 관리를 위한 스마트 홈 케어 서비스를 위해서는 재실 여부 및 활동 유형에 관한 다양한 연구가 수행되고 있다. 본 논문에서는 스마트 홈 케어 서비스를 위해 실내 온도, CO₂, 미세먼지 값과 UWB 레이더 측위를 통해 재실 여부 뿐만 아니라 활동 유형을 분류하는 랜덤 포레스트 모델을 제안한다. 실험은 실내 온도, CO₂, 미세먼지를 측정하는 센서 3개와 UWB Radar 2개를 사용하여 2초 간격으로 실내 환경 및 재실자의 측위 데이터를 측정한다. 측정된 데이터는 이상치와 결측치를 보정 처리한 후 80%의 트레이닝 셋 데이터와 20%의 테스트 셋 데이터로 이분하며 랜덤 포레스트 모델을 적용하여 중요도 상위 변수 목록, 정확도, 민감도, 특이도, T1 스코어를 평가한다.

주요어 : 사람 활동 인식, 랜덤 포레스트, UWB, 실내 환경 측정 데이터

Abstract As the world becomes an aging society due to a decrease in the birth rate and an increase in life expectancy, a system for health management of the elderly population is needed. Among them, various studies on occupancy and activity types are being conducted for smart home care services for indoor health management. In this paper, we propose a random forest model that classifies activity type as well as occupancy status through indoor temperature and humidity, CO₂, fine dust values and UWB radar positioning for smart home care service. The experiment measures indoor environment and occupant positioning data at 2-second intervals using three sensors that measure indoor temperature and humidity, CO₂, and fine dust and two UWB radars. The measured data is divided into 80% training set data and 20% test set data after correcting outliers and missing values, and the random forest model is applied to evaluate the list of important variables, accuracy, sensitivity, and specificity.

Key words : Human Activity Recognition, Random Forest, UWB, Indoor Environment Measurement Data

*정회원, 한국전자기술연구원 스마트가전혁신지원센터 (제1저자) Received: October 21, 2022 / Revised: November 3, 2022

**정회원, 한국전자기술연구원 스마트가전혁신지원센터 (참여저자)

Accepted: November 8, 2022

*Corresponding Author: hmfl1436@naver.com

접수일: 2022년 10월 21일, 수정완료일: 2022년 11월 3일
게재확정일: 2022년 11월 8일

Smart Appliances Innovation Support Center, Korea Electronics
Technology Institute, Korea

I. 서 론

세계적으로 식생활 향상, 의료기술 예방과학의 발달 등의 다양한 요인으로 인간의 기대수명은 점점 높아지고 있다. 기대수명이 늘어나고 출생률은 점점 감소하는 상황으로 고령인구의 비율이 급격히 증가하고 있다. 한국의 경우에도 2025년이 되면 인구의 25% 이상이 65세 이상의 고령 인구가 될 것으로 전망된다[1]. 이에 따라 고령인구의 건강 관리를 지원하기 위한 의료 및 케어 산업은 각 나라에서 국가차원의 전략적 사업으로 추진하고 있다[2]. 재실 중인 거주자의 케어를 위한 스마트 홈 기술은 세계적으로 점점 중요해지고 있으며[3] 이를 위해 재실 감지 및 활동 유형 분석이 필요하다. 재실감지는 일반적으로 카메라를 사용하여 재실 상태를 감지하는 경우[4]와 RIP 센서를 활용해 신체의 열을 감지하여 열 변화량을 계산해 움직임을 측정하는 경우가 있다.[5]. 카메라를 통한 재실감지는 데이터 처리를 위해 많은 계산량이 요구되며 사생활과 인권침해 문제에 대한 단점이 있다[6]. RIP 센서 기반의 재실 탐지는 정지 상태 일 때 감지 성능이 낮으며 빛 반사에 민감하다[7].

본 논문에서는 온습도, CO₂, 미세먼지, UWB 레이더 측위 데이터로 재실 여부 그리고 활동 유형을 분류하는 랜덤 포레스트 모델을 제안한다. 실험은 실내 온습도, CO₂, 미세먼지를 측정하는 센서 3개와 UWB Radar 2개를 사용하여 2초 간격으로 실내 환경 및 재실자의 측위 데이터를 측정한다. 측정된 데이터는 이상치와 결측치 보정 후 정적인 활동과 동적인 활동으로 나누어 점양분 상관계수(Point-biserial correlation coefficient)를 통해 각 변수들의 관련성을 파악한 후 전체 데이터를 80%의 트레이닝 데이터셋과 20%의 테스트 데이터셋으로 구분한다. 이후, 분류를 위한 랜덤 포레스트 모형과 엑스트라 트리 모형을 구축한다. 구축된 모형들을 트레이닝 데이터셋을 사용하여 학습시키며 테스트 데이터셋을 사용하여 성능을 평가한다. 평가는 AUC(Area under the ROC curve), 정확도(Accuracy), 민감도(Sensitivity), 특이도(Specificity)를 계산하여 평가한다.

II. 실험 환경 및 방법

1. 실험 환경

1) IR-UWB 레이더

IR-UWB 파형의 송수신 역할로 Novelda의 Xethru X4를 사용하며 그림 1과 같다. X4는 세팅된 값의 파형을 송신(Tx)부를 통해 송신하며, 사물에 부딪힌 후 반사된 신호를 수신(Rx)부를 통해 수신한다. 수신된 1536개 빈 프레임의 총 길이의 합은 약 10m에 해당하며 각 프레임의 신호 크기를 통해 재실자의 거리를 파악하며 총 두 개의 레이더에서 측정된 거리 데이터를 바탕으로 위치를 측정한다.

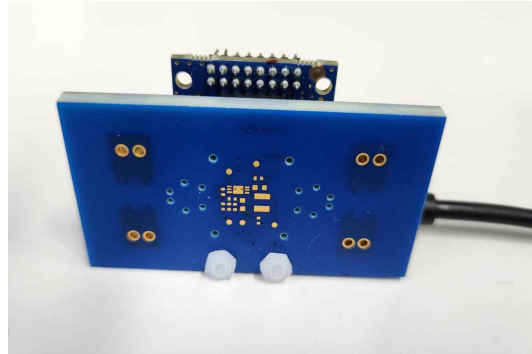


그림 1. IR-UWB 레이더
Figure 1. IR-UWB Radar

2) 실내 환경 센서

그림 2는 실내 환경 센서를 나타내며 사양은 표 1과 같다. 실내 환경 정보는 온습도, CO₂, 미세먼지(PM1.0, PM2.5, PM10)를 측정하며 센서는 Innociple의 InnoAir-615IoT 제품을 사용한다. 표는 사용된 센서의 사양을 나타낸다. 온도는 -30 °C에서 100 °C까지 측정되며 습도는 0%에서 100%까지 측정된다. 미세먼지는 PM1.0 $\mu\text{g}/\text{m}^3$, PM2.5 $\mu\text{g}/\text{m}^3$, PM10 $\mu\text{g}/\text{m}^3$ 의 입자 크기를 검출하며 농도의 범위는 0.01 $\mu\text{g}/\text{m}^3$ 에서 1,000 $\mu\text{g}/\text{m}^3$ 까지 측정된다.



그림 2. 실내 환경 센서
Figure 2. Indoor environment sensor

표 1. 실내 환경 센서 사양

Table 1. Indoor environment sensor

| Model | Value | Range | Resolution |
|--------------------|--------------------------------------|---|-------------------------------|
| InnoAir-61 5LoT | Temperature(°C) | -30 °C ~ 100 °C | 0.1°C |
| | Humidity(%) | 0 % ~ 100% | 0.1 % |
| | CO2(ppm) | 0 ppm ~ 5000 ppm | 1 ppm |
| | Finedust($\mu\text{g}/\text{m}^3$) | 0.01 $\mu\text{g}/\text{m}^3$ ~ 1,000 $\mu\text{g}/\text{m}^3$ | 0.01 $\mu\text{g}/\text{m}^3$ |

3) 실험 환경

그림 3은 실험 환경 및 장비 설치 위치를 나타내며 전체 공간은 현관, 거실, 주방, 침실, 화장실로 구분된다. IR-UWB 레이더는 거실의 전열 교환기에 각각 1대씩 총 2대가 위치하며 영역 내의 재실자의 위치를 추적한다. 실내 환경 센서는 총 3개가 있으며 각각 침실, 거실, 주방에 위치하여 각 영역의 실내 환경 변화를 측정할 수 있다.



그림 3. 실험 환경

Figure 3. Experiment environment

2. 실험 방법

1) 실험 방법

본 실험의 재실활동 유형은 공부, 요리, 수면, 청소, 운동으로 구성된다. 공부 활동의 경우 재실자는 책상에 위치하며 독서, 필기 등의 행동을 실시했다. 요리 활동의 경우 재실자는 주방에 위치하며 육류, 라면 등의 음식을 조리하였다. 수면 활동의 경우 재실자는 침실의 침대에 누운 상태로 위치하며 수면을 취했다. 청소의

경우 침실, 거실, 주방에서 자유롭게 이동하며 진공 청소기를 사용하여 청소를 실시했다. 운동의 경우 거실의 내에서 진행하였으며 재실자는 아령을 이용한 무산소 운동 및 제자리에서 유산소 운동을 진행하였다. 각 활동별로 발생하는 실내 활동은 30분 진행 후 15분 동안 환기하는 식으로 이전의 활동에 따라 이후 활동의 데이터에 영향을 주지 않도록 진행하였다.

III. 측정 결과 및 분석

1. 측정 결과

1) 측정 결과

표2는 재실활동 유형별 온습도, CO₂, 미세먼지 측정값의 최소, 최대, 평균 및 표준편차를 나타낸다. 측정된 데이터에서 센서 오작동으로 판단되는 이상치는 제거하였으며, 온도는 최소 20.4 °C에서 최대 28 °C까지, 습도는 최소 22.5%에서 최대 100%까지 측정됐다. CO₂ 농도는 최소 415 ppm에서 최대 988ppm 까지 측정됐고, 미세먼지는 최소 3.9 $\mu\text{g}/\text{m}^3$ 에서 최대 470.4 $\mu\text{g}/\text{m}^3$ 까지 측정됐다. 총 데이터의 수는 84,323set 이며 공부 23,406 set 요리 10,008set 수면 4,778set 운동 33,299set 청소 12,830set 으로 구성된다.

2) 측정 결과 분석

각 유형의 특징을 살펴보면 공부 활동과 수면 활동 유형의 경우 온습도, CO₂, 미세먼지의 데이터는 비슷하게 측정되었지만 공부의 측위 데이터는 거실의 책상으로 측정되고 수면의 측위 데이터는 침실의 침대의 위치로 측정된다. 요리 활동 유형의 경우 그림 4와 같이 활동을 시작 할 때 미세먼지 데이터의 측정 값이 매우 높게 나타나며 이후 시간이 지남에 따라 점점 낮아지는 형태를 보이며 측위 데이터는 주방에 위치하는 것으로 측정된다.

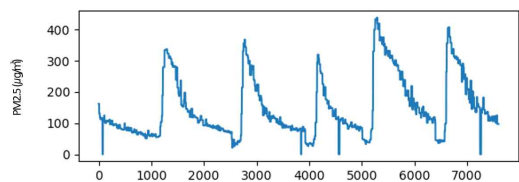


그림 4. 요리 활동 유형 실내 미세먼지 데이터

Figure 4. Indoor finedust data of cook activity

표 2. 재실활동 유형별 실내 환경 센서 측정 값
Table 2. Indoor environment sensor measurement value by type of indoor activity

| Activity type | Value | Min | Max | Average | SD |
|---------------|------------------|------------|-------------|-------------|-------|
| Study | Temperature (°C) | 21.6 °C | 27.6 °C | 25.3 °C | 1.19 |
| | Humidity (%) | 23 % | 43 % | 31.3% | 4.7 |
| | CO2 (ppm) | 433 ppm | 775 ppm | 569 ppm | 68.6 |
| | Finedust (µg/m³) | 5.8 µg/m³ | 64.4 µg/m³ | 28.6 µg/m³ | 9.2 |
| Cook | Temperature (°C) | 21.0 °C | 23.2 °C | 22.0 °C | 1.06 |
| | Humidity (%) | 22.5 % | 68 % | 56.8 % | 19.7 |
| | CO2 (ppm) | 415 ppm | 768 ppm | 566 ppm | 95 |
| | Finedust (µg/m³) | 5.2 µg/m³ | 470.4 µg/m³ | 151.1 µg/m³ | 105 |
| Sleep | Temperature (°C) | 21.2 °C | 25.2 °C | 22.6 °C | 1.4 |
| | Humidity (%) | 23 % | 65.2 % | 43.5 % | 22.1 |
| | CO2 (ppm) | 511 ppm | 907 ppm | 737.6 ppm | 97 |
| | Finedust (µg/m³) | 12.2 µg/m³ | 66.7 µg/m³ | 23.8 µg/m³ | 6.4 |
| Exercise | Temperature (°C) | 20.4 °C | 27.8 °C | 23.1 °C | 1.7 |
| | Humidity (%) | 24 % | 100 % | 56.9 % | 18.1 |
| | CO2 (ppm) | 423 ppm | 988 ppm | 669.3 ppm | 120.3 |
| | Finedust (µg/m³) | 3.9 µg/m³ | 63.2 µg/m³ | 24 µg/m³ | 10.9 |
| Clean | Temperature (°C) | 21.4 °C | 28 °C | 25.1 °C | 2 |
| | Humidity (%) | 25 % | 71.2 % | 41.1 % | 17.1 |
| | CO2 (ppm) | 439 ppm | 952 ppm | 577 ppm | 94.8 |
| | Finedust (µg/m³) | 7.1 µg/m³ | 152.2 µg/m³ | 37.1 µg/m³ | 37.3 |

운동 활동 유형의 경우 그림 5와 같이 각 활동을 시작 한 이후 꾸준하게 CO2가 증가하는 추세를 보인다.

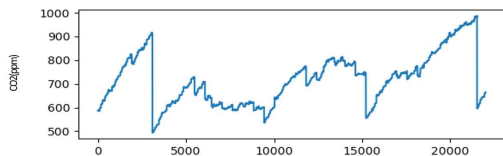


그림 5. 운동 활동 유형 CO2 데이터
Figure 5. CO2 data of exercise activity

IV. 학습모델 구축 및 성능 평가

1. 학습모델 구축

1) 랜덤 포레스트 모델 구축

랜덤 포레스트 모델은 분류, 회귀 분석 등에 사용되는 학습 방법으로 의사결정트리에 배깅(Bagging)을 적용한 결과를 종합한 방법이다[8]. 랜덤 포레스트 모델은 특히 다수의 변수를 활용할 때 성능이 뛰어나며 각 변수의 변수중요도(variable importance)가 산출 가능하다[9]. 또한, 생성된 모델의 내부를 관찰 할 수는 없고 [10] 생성할 트리의 수를 지정하면 그 수만큼의 의사결정 나무를 구성한다. 전체 데이터는 무작위하게 선택 및 배정되며 각 결과를 종합하여 배깅을 하여 과적합을 최소화한다. 표 3은 재실활동 유형에 따른 트레이닝 셋 및 테스트 셋의 수를 나타낸다.

표 3. 재실활동 유형에 따른 트레이닝 셋 및 테스트 셋의 수
Table 3 Number of training sets and test sets according to the type of indoor activity

| Group | Training set(n) | Test set(n) | Total (n) |
|----------|-----------------|-------------|-----------|
| Study | 18,724 | 4,682 | 23,406 |
| Cook | 8,006 | 2,002 | 10,008 |
| Sleep | 3,822 | 956 | 4,778 |
| Exercise | 26,639 | 6,660 | 33,299 |
| Clean | 10,265 | 2565 | 12,830 |
| Total | 67,456 | 16,865 | 84,323 |

2. 성능 평가

1) 랜덤 포레스트 모델 성능 평가

구축한 모델의 성능은 테스트 셋을 대상으로 각각 정확도, 재현율, 정밀도, f1 score를 계산해 평가했다. 아래의 식 (1),(2),(3),(4)는 각각 정확도, 민감도, 특이도, F1 Score를 나타낸다. 식의 TP(True Positive)는 참을 참으로 예측한 경우의 수를 TN(True negative)는 거짓의 경우를 거짓으로 올바르게 예측한 경우의 수를 나타낸다. FP(False positive)는 거짓을 참으로 잘못 인식한 경우의 수를 FN(False negative)는 거짓을 참으로 인식한 사례를 나타낸다. 각 민감도와 특이도 F1 Score는 재실활동 유형별 결과값의 평균을 도출해낸 값으로 표기한다.

$$Accuracy = \frac{TP + TN}{TP + TN + EP + FN} \quad (1)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (2)$$

$$Specificity = \frac{TN}{TN + EP} \quad (3)$$

$$F1\ Score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (4)$$

표 4는 테스트셋 데이터를 사용하여 랜덤 포레스트 모델의 성능을 평가한 결과표로 정확도는 98.5%, 민감도는 99.115%, 특이도는 98.796%, F1 Score는 98.118%로 나타났다.

표 4. 테스트 셋 데이터를 사용한 랜덤 포레스트 모델의 성능 평가 결과

Table 4 Performance evaluation result of random forest model using test set data

| Model | Accuracy (%) | Sensitivity (%) | Specificity (%) | F1 Score |
|---------------|--------------|-----------------|-----------------|----------|
| Random forest | 98.5 | 99.115 | 98.796 | 98.118 |

V. 결 론

본 논문에서는 IR-UWB 레이더 측위 시스템과 실내 환경 센서 데이터를 이용한 재실활동 유형 분류하는 랜덤 포레스트 모델을 구축 및 평가했다. 주거 공간을 대상으로 한 본 연구에서는 랜덤 포레스트 모델을 활용하여 재실활동 유형 감지에 대한 가능성을 확인했다. 운동 유형 및 요리 유형의 경우 시간의 경과에 따라 각각 CO2와 미세먼지의 값이 유의미하게 변화하기 때문에 시계열 예측에 적합한 LSTM이나[11] LSTM보다 더 간단한 구조를 가진 GRU[12] 모델을 구축하여 추가적인 데이터 획득 후 학습을 진행 할 경우 더 좋은 성능의 재실 활동 유형 감지 모델을 구축 할 수 있다. 또한 실시간 성능을 강화하기 위해 본 논문에서 구축된 랜덤 포레스트 모델과 앙상블(Ensemble)하여 중,단기적으로 강화된 성능의 모델을 얻을 수 있을 것이다[13]. 또한 다양한 측정 데이터를 결합한 재실 활동 유형 감지에 관한 연구는 거의 조사되지 않기 때문에 이와 관련된 추가적인 연구가 필요할 것으로 보인다.

References

- [1] Hwang, Jung-Hyun..An Analysis of the Research and Policy on the Smart Home Design for the Elderly based on the U-Healthcare.JOURNAL OF THE ARCHITECTURAL INSTITUTE OF KOREA Planning & Design,31(4),53-60, 2015. DOI:10.5659/JAIK_PD.2015.31.4.53
- [2] Weon, Dal Soo, Lee, Sang San, and Jung, Yong Gyu, "Global Convergence for Healthcare ICT Services," The journal of the convergence on culture technology, vol. 2, no. 2, pp. 45 - 49, May 2016. DOI:10.17703/JCCT.2016.2.2.45
- [3] MARTIN, Suzanne, et al. Smart home technologies for health and social care support. Cochrane database of systematic reviews, 2008, 4. DOI: 10.1002/14651858.CD006412.pub2
- [4] Sung Hoon Yoon and Kil Soo Lee "Indoor Surveillance Camera based Human Centric Lighting Control for Smart Building Lighting Management," The International Journal of Advanced Culture Technology (IJACT), Vol.8 No.1 pp. 207-212, 2020. DOI:10.17703/IJACT.2020.8.1.207
- [5] F. Wahl, M. Milenkovic, and O. Amft, "Adistributed PIR-based approach for estimating people count in office environments", IEEE 15th International Conference on Computational Science and Engineering, pp. 640-647, 2012.DOI: 10.1109/ICCSE.2012.92
- [6] SLOBOGIN, Christopher. Public privacy: camera surveillance of public places and the right to anonymity. Miss. IJ, 2002, 72: 213. DOI:10.2139/ssrn.364600
- [7] S. S. Leea, C. Choia, and M. Kim, "CNN-based People Recognition for Vision Occupancy Sensors". Journal of Broadcast Engineering, Vol. 23, no. 2, pp. 274-282, 2018. DOI:10.5909/JBE.2018.23.2.274
- [8] BIAU, Gérard; SCORNET, Erwan. A random forest guided tour. Test, 2016, 25.2: 197-227. DOI:10.48550/arXiv.1511.05741
- [9] J. D. Suh, "Foreign Exchange Rate Forecasting Using the GARCH extended Random Forest Model", Journal of Industrial Economics and Business, Vol. 29, no. 5, pp. 1607-1628, 2016.
- [10]A. M. Prasad, L. R. Iverson, and A. Liaw, "Newer classification and regression tree techniques: bagging and random forests for ecological prediction," Ecosystems, Vol. 9, pp.

- 181-199, 2006. DOI:10.1007/s10021-005-0054-1
- [11]YU, Yong, et al. A review of recurrent neural networks: LSTM cells and network architectures. *Neural computation*, 2019, 31.7: 1235-1270. DOI:10.1162/neco_a_01199
- [12]DEY, Rahul; SALEM, Fathi M. Gate-variants of gated recurrent unit (GRU) neural networks. In: 2017 IEEE 60th international midwest symposium on circuits and systems (MWSCAS). IEEE, 2017. p. 1597-1600. DOI:10.1109/MWSCAS.2017.8053243
- [13]TEBALDI, Claudia; KNUTTI, Reto. The use of the multi-model ensemble in probabilistic climate projections. *Philosophical transactions of the royal society A: mathematical, physical and engineering sciences*, 2007, 365.1857: 2053-2075. DOI:10.1098/rsta.2007.2076

※ 본 연구는 국토교통부의 국토교통기술사업화
지원사업의 연구비지원에 의해 수행되었음
[과제명 : 비접촉 생체정보 측정기능이 포함된
스마트 디퓨저 기반 거주자 맞춤형 Home-
HAS(Health, Air, Safety) 서비스 개발]
[과제번호 : 21TBIP-C161696-01].