

UWB 레이더 기반 Vision Transformer를 활용한 홈 헬스케어용 객체 위치 탐지

이강현¹, 송채영¹, 조현호¹, 최상호^{1*}

광운대학교 컴퓨터정보공학부¹

Object Location Detection for Home Healthcare using UWB radar-based Vision Transformer

Kanghyun Lee¹, Chaeyoung Song¹, Hyeonho Jo¹, Sang Ho Choi^{1*}

School of Computer and Information Engineering, Kwangwoon University, Korea

*shchoi@kw.ac.kr

Abstract

Monitoring technology for the healthcare of elderly individuals and people with disabilities in a home environment is essential for effectively managing their health and safety. Although there are several devices for home healthcare, the majority of them have the limitation of requiring the user to wear them. Meanwhile, Ultra-Wideband (UWB) radar can be utilized for non-contact detection and tracking of objects. Accurate location tracking is essential for detecting falls and monitoring activity levels within a home environment. In this study, we developed an AI model based on a vision transformer (ViT) that utilizes UWB radar to detect and track objects. The data of 60 people for the study were collected, and data preprocessing was performed directly according to the experimental scenario. After preprocessing, the ViT model pretrained with Cifar10 was fine-tuned using the data of 48 people. The proposed model can detect object movements from UWB radar measurements taken at different angles with high accuracy rates of 96.9% and 98.5%.

1. 연구 배경

최근, 세계적인 고령화 추세에 따라 심장마비, 폐질환, 암, 당뇨, 뇌졸중 등 다양한 만성질환의 발병률이 증가하고 이러한 상황에서 돌봄이 필요한 약자들을 관찰할 필요성이 증가한다.[1] 이는 관찰에 필요한 많은 인력을 요구하고 따라서 지속적인 관찰 가능한 시스템의 필요성으로 이어진다. 해당 연구는 대상의 위치를 무구속적으로 관찰하면서 특정 상황을 탐지하기 위해 Ultra-wideband (UWB) 레이더를 이용한 객체 위치 탐지 시스템을 제안한다. 대상의 위치와 상태를 탐지하는데 있어서 시계, 밴드, 반지와 같은 각종 헬스케어 제품이 많이 존재한다. 해당 제품들은 모두 접촉식이며 늘 착용해야 한다는 명확한 단점이 존재한다. 하지만 UWB 레이더를 이용한다면 특정 범위내에서 대상을 비접촉식으로 언제든지 관찰 가능하다. 또한 도플러 레이더와 다르게 더 높은 대역폭과 미세한 움직임을 포착할 수 있다는 특징으로 추후 대상의 위치를 넘어서 미세한 생체 신호 또한 측정할 수 있다는 확장성을 지닌다.[1] 카메라를 이용한 측정이 아니기 때문에 대상의 개인정보 침해에 덜 민감하다는 장점 또한 존재한다.

최근 LLM에서 Generative AI 모델로 각광받는 transformer 구조를 computer vision 분야에 적용한 Vision transformer (ViT)를 학습 모델로 선정하였다.[2] UWB 레이더를 이용하여 측정된 데이터는 대상의 위치 변화를 포함하고 있다. 데이터를 이미지화하고 이를 cifar10 데이터셋으로 사전 훈련시킨 모델에 fine-tuning 했을 때의 성능을 살펴보고 transformer 모델의 레이더 데이터 훈련의 확장성을 확인한다.[2]

2. 연구 방법

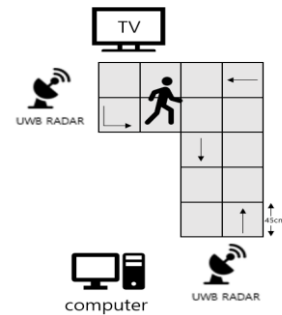


그림 1 실험 시나리오

그림 1과 같이 위치를 특정할 수 있게 한 칸당 45cm, 총 14칸 구역을 피험자가 이동하는 것을 UWB 레이더를 이용하여 측정하였다. 실제 가정환경과 같이 실험환경을 구성했고 TV와 computer와 같은 물체들이 존재한다. 피험자는 3초 단위로 한 칸씩 움직이고 총 5분 동안 측정을 진행하였다. 사용한 레이더는 novelda의 X4M300, fps는 17이다.[5]

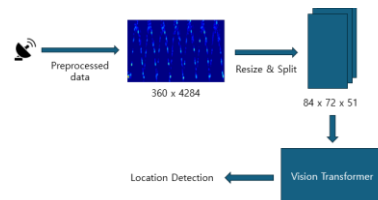


그림 2 전체 시스템 구조

그림 2는 UWB 레이더를 이용하여 측정된 움직임 데이터가 처리되는 전체 과정을 도식화한 것이다. 전처리 과정에서는 레이더의 raw data에서 baseband를 추출하고 이를 이용해 dynamic clutter를 제거한다. 이후 수식 1을 통해 threshold를 설정하고 한번 더 신호를 필터링한다. 이후 수식 2를 적용하여 신호를 정제한다. 이때 수식 2의 N의

값은 51이다.[3]

$$\text{Threshold} = \mu + 3\sigma$$

수식 1 Threshold

$$\mu = \frac{1}{N} \sum_{i=1}^N A_i$$

수식 2 Moving Everage Filter

Preprocessing 된 데이터를 학습시키기 위해 3초마다 움직인 시나리오를 기반으로 17(fps) x 3(sec) = 51의 크기로 나누고 객체가 위치하는 구역을 명확히 하기 위해 360에 해당하는 위치 인덱스를 5씩 합하여 72의 크기로 줄였다. 결과적으로 인당 84 x 72 x 51의 데이터를 이용해 모델을 학습한다.

모델은 PyTorch Image Models(timm) 라이브러리에서 제공하는 pretrained vision transformer를 사용하였다. [4] 실험 데이터는 피험자 60명중 48명의 실험 데이터로 fine-tuning하여 사용하였다.

본 연구는 광운대학교 기관윤리위원회의 승인을 받아 실시하였다 (IRB no. 7001546-202300831-HR(SB)-008-03).

3. 연구 결과

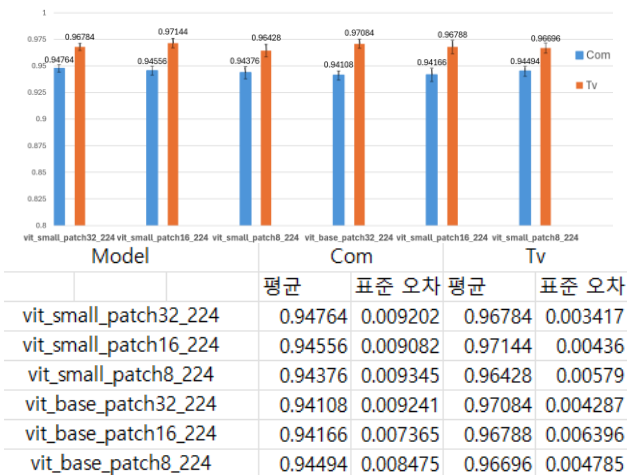


그림 3 5-fold cross validation

그림 3은 제시된 6개 모델에 각각 5-fold cross validation을 진행하고 각 fold의 validation accuracy의 평균과 표준오차를 나타낸 것이다. 총 48명 중 40명 데이터로 진행하였다. 그림 1 실험 시나리오를 기반으로 하늘색 막대는 com, 주황색 막대는 tv로 각각 레이더가 존재했던 위치를 기반으로 명명하여 범례를 지칭하였다. Com은 실험 시나리오의 객체의 위치인 14칸 중 8개, Tv는 6개를 탐지하도록 모델을 구성하였다.

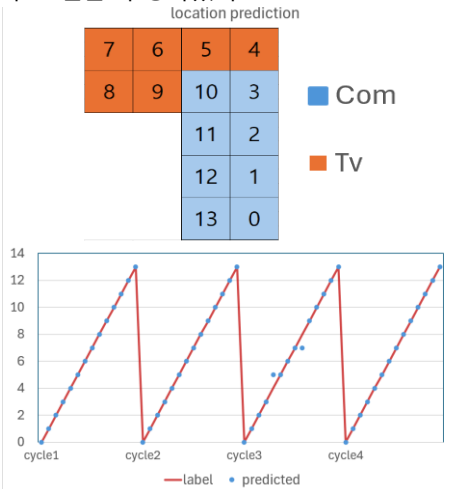


그림 4 Object Tracking 결과 예시

그림 4는 그림 1에 따른 시나리오를 따를 때 14개의 구역으로 나누어진 영역과 각각의 radar가 어느 부분을 classification 하는지 색깔로 표시한 것이다. 또한 피험자가 처음 출발한 지점으로 다시 돌아오는 것(0~13)을 한 cycle이라 할 때 test data로 평가한 결과를 이용하여 tracking하는 모습을 확인할 수 있다.

	Com	Tv
Subject1	1	0.96192
Subject2	0.99524	1
Subject3	0.99286	0.99286
Subject4	0.99048	0.98096
Subject5	0.91904	0.99762
Subject6	0.99286	0.99286
Subject7	0.99524	1
Subject8	0.94288	0.99524
평균	0.9786	0.9902
표준 오차	0.010676	0.004574

표 1 Test Accuracy

결과적으로 6개의 모델을 비교한 결과 모델의 크기가 작으면서 좋은 성능을 보인 vit_small_patch16_224 모델로 8명 데이터를 이용해 test를 진행하고 subject 각각의 Accuracy를 표 1를 통해 나타내었다.

4. Acknowledgements

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터육성지원사업과 (IITP-2024-RS-2022-001562 25) 2024년도 정부(교육부)의 재원으로 과학기술정보통신부와 정보통신기획평가원의 지원을 받아 수행된 (대학혁신지원(R&D)) 연구결과임

5.참고 문헌

[1] Marco Mercuri^{1*}, Ilde Rosa Lorato¹, Yao-Hong Liu¹, Fokko Wieringa^{1,2}, Chris Van Hoof^{1,3,4} and Tom Torfs³ “Vital-sign monitoring and spatial tracking of multiple people using a contactless radar-based sensor” Nature Electronics | VOL 2 | JUNE 2019 | 252–262

[2] Alexey Dosovitskiy^{*,†}, Lucas Beyer^{*}, Alexander Kolesnikov^{*}, Dirk Weissenborn^{*}, Xiaohua Zhai^{*}, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, Neil Houlsby^{*,†}, arXiv:2010.11929v2 [cs.CV] 3 Jun 2021

[3] MatLab, “Moving Everage Filter”, <https://kr.mathworks.com/help/matlab/ref/movmean.html>, 2024.07.06

[4] HuggingFace, “Vision Transformer”, https://huggingface.co/docs/transformers/model_doc/vit, 2024.07.06

[5] UWB radar, “XeThru-Explorer”, <https://github.com/novelai>