

WiFi CSI 기반 행동 인식 모델의 교차 환경 성능 변화 분석

김우재, 안호연*

순천향대학교

woojae.kim@sch.ac.kr, *hoyeon@sch.ac.kr

Analysis of Cross Environment Performance Variations in WiFi CSI Based Human Activity Recognition Models

Woojae Kim, Hoyeon Ahn*

Shunchunhyang University

요약

본 논문은 ESP32 기반 공개 WiFi CSI 데이터셋을 이용하여 행동 인식의 교차 환경 성능 변화를 정량적으로 분석한다. 3개 환경을 대상으로 학습 환경과 테스트 환경을 분리한 총 6개 조합에서, 원시 데이터 학습(Raw)과 EMD 기반 증강 데이터 학습(EMD)의 성능을 비교하였다. 실험 결과, EMD 증강은 교차 환경 조건에서 성능 향상을 일관되게 보장하지 못하였으며, 일부 조합에서는 정확도와 F1 지표가 오히려 감소하였다. 특히 가장 높은 성능을 보인 조합에서도 Raw 조건은 87% 정확도로 평가된 반면 EMD 조건은 83%로 낮게 나타났음을 확인하였다.

I. 서론

WiFi 채널 상태 정보인 CSI(Channel State Information)는 별도의 착용형 센서 없이 사람의 움직임에 의해 변하는 무선 채널 변화를 관측할 수 있어 비접촉식 행동 인식에 많이 활용되고 있다[1]. 그러나 CSI는 수신기 배치, 공간 구조, 반사 환경 등 환경 요인에 민감하여 학습 환경과 다른 환경에서 성능이 크게 변동하는 문제가 존재한다. 이러한 한계를 보완하기 위해 데이터 증강을 통해 학습 표본의 다양성을 확장하려는 접근이 제안되어 왔으며, 특히 EMD(Empirical Mode Decomposition)를 이용하여 CSI 신호를 성분 단위로 분해한 뒤 재합성하는 방식은 단일 환경 조건에서 분류 정확도를 크게 향상시킬 수 있다는 결과가 보고된 바 있다[2].

다만 단일 환경에서의 성능 향상이 실제 적용으로 이어지기 위해서는, 환경이 변화하더라도 동일한 이득이 유지되는지에 대한 검증이 필요하다. 본 연구는 ESP32 기반 공개 CSI 데이터셋을 대상으로 원시 CSI 데이터에 대해 EMD 기반 증강을 직접 수행하고, 교차 환경 평가를 통해 EMD 증강 적용 여부에 따른 일반화 성능 변화를 정량적으로 분석한다. 이를 통해 EMD 기반 증강이 환경 변화 조건에서도 일관되게 높은 성능을 유지하는지 여부를 확인한다.

II. 본론

1. 데이터셋

본 연구는 Mendeley Data에 공개된 ESP32 기반 WiFi CSI 행동 인식 데이터셋을 사용한다[3]. 데이터는 서로 다른 측정 조건을 반영한 Environment1~3으로 구분되어 제공되며, 본 연구는 standing, sitting_down, squatting의 3개 클래스를 대상으로 한다.

2. EMD 기반 데이터 증강 절차

본 연구는 원시 CSI 시계열의 변동성을 확장하기 위해 EMD[4]를 이용한 데이터 증강을 수행한다. 증강의 기본 아이디어는 각 서브캐리어 시계열을 EMD로 분해한 뒤, 분해된 성분(IMF, Intrinsic Mode Function)의

크기를 무작위로 조절하여 다시 합성함으로써 원시 신호와 유사하지만 서로 다른 변형 신호를 생성하는 것이다. 본 연구에서는 원본 샘플 1개로부터 증강 샘플 2개를 추가 생성하여, 학습 데이터를 총 3배로 확장한다.

먼저 각 CSI 샘플은 시간 길이 T 와 서브캐리어 수 52로 구성된 행렬 $X \in R^{T \times 52}$ 로 표현하며, 증강은 서브캐리어별 1차원 시계열 $x_s(t) = X(t, s)$ 에 대해 독립적으로 수행한다. 각 $x_s(t)$ 에 EMD를 적용하여 IMF 성분 $c_{s,k}(t)$ 를 추출하고, 계산량을 고려하여 IMF의 최대 개수는 $K \leq 3$ 으로 제한한다. 이후 증강 샘플 생성을 위해 각 IMF에 대해 무작위 가중치 α_k 를 균등분포 $U(0.8, 1.2)$ 에서 샘플링하여 곱한 뒤 재합성한다. 이때 증강된 서브캐리어 신호는 다음과 같이 정의된다.

$$\tilde{x}_s(t) = \sum_{k=1}^K \alpha_k c_{s,k}(t), \quad \alpha_k \sim U(0.8, 1.2)$$

마지막으로 52개 서브캐리어에 대해 생성된 증강 시계열을 결합하여 $\tilde{X} \in R^{T \times 52}$ 형태의 증강 CSI 샘플을 구성한다. 증강 과정의 안정성을 위해, 신호 변동이 매우 작은 경우에는 EMD 분해가 유효하지 않을 수 있으므로 해당 서브캐리어는 원시 신호를 그대로 사용한다.

3. 교차 환경 평가 설정 및 모델

3개 환경을 대상으로 총 6개 교차 환경 시나리오를 구성하여 일반화 성능을 평가하였다. 단일 환경 학습(Tr 1개)에서는 나머지 2개 환경 데이터를 결합한 테스트셋(Te)으로 평가하고, 이중 환경 학습(Tr 2개)에서는 남은 1개 환경(Te)에서 평가하였다. 모든 시나리오의

학습 및 테스트 데이터는 환경 기준으로 완전히 분리하였다. Raw는 원

표 1. 환경 및 클래스별 샘플 수

	Standing		Sitting_down		Squatting	
	Raw	EMD	Raw	EMD	Raw	EMD
Env1	81	243	799	2385	799	2397
Env2	641	1923	800	2400	800	2400
Env3	641	1923	800	2400	800	2400

표 2. 학습 환경(Tr)과 테스트 환경(Te) 조합에 따른 교차 환경 성능 비교 (%)

Metric	S1 (Tr:E1 Te:E2+E3)		S2 (Tr:E2 Te:E1+E3)		S3 (Tr:E3 Te:E1+E2)		S4 (Tr:E1+E2 Te:E3)		S5 (Tr:E1+E3 Te:E2)		S6 (Tr:E2+E3 Te:E1)	
	Raw	EMD	Raw	EMD	Raw	EMD	Raw	EMD	Raw	EMD	Raw	EMD
Accuracy	52.1	58.5	52.1	49.3	64.2	59.3	87.2	83.8	48.8	49.8	57.6	61.5
Macro F1	40.6	57.1	48.6	45.1	66.3	61.6	87.4	84.1	40.9	40.5	53.1	56.2
Weighted F1	43.4	57.2	42.7	38.9	63.9	57.9	87.1	83.5	42.7	42.6	56.8	60.5
Standing F1	0.0	56.7	74.9	72.7	76.9	77.8	92.0	93.6	15.0	10.6	44.5	46.1
Sitting F1	58.9	49.9	12.5	6.5	60.3	58.9	88.8	84.6	56.9	48.6	47.9	49.6
Squatting F1	62.8	64.8	58.3	56.0	61.7	48.0	81.4	74.2	50.7	62.3	66.9	72.8

본 데이터로 학습한 모델, EMD는 EMD 기반 증강 데이터를 포함하여 학습한 모델을 의미하며, 두 조건은 동일한 학습 절차에서 데이터 구성만 달리하여 비교하였다. 입력은 시간 길이 50으로 정규화된 52 서브캐리어 시계열로 구성하였다.

4. 실험 결과

표 2는 6개 교차 환경 시나리오(S1-S6)에서 Raw와 EMD 조건의 성능을 비교한 결과이다. 단일 환경 학습 조건(S1-S3)에서 EMD는 S1에서 정확도가 52.1%에서 58.5%로 향상되었으나, S2와 S3에서는 각각 52.1%에서 49.3%, 64.2%에서 59.3%로 감소하였다. 이중 환경 학습 조건(S4-S6)에서도 EMD는 S6에서 정확도가 57.6%에서 61.5%로 향상된 반면, S4에서는 87.2%에서 83.8%로 감소하였고 S5에서는 48.8%에서 49.8%로 유사한 수준을 보였다. Macro F1 및 클래스별 F1을 함께 고려하면, EMD 적용은 일부 조합에서 특정 클래스 성능을 크게 개선하나 다른 클래스 성능을 저하시키는 상충 관계가 관찰되었다. 즉, EMD 기반 증강은 교차 환경 조건에서 성능 향상을 일관되게 보장하지 않으며, 학습 환경 조합에 따라 효과가 달라질 수 있음을 의미한다.

그림 1에서 두 조건 모두에서 주요 오류는 squatting이 sitting_down으로 오분류되는 현상으로 나타나며, EMD 적용 시 해당 오분류가 감소하였

다. 반면 standing에서는 EMD 적용 시 정분류가 감소하고 오분류가 증가하는 경향이 관찰된다. 즉, Set4에서는 EMD 증강이 squatting 구분에는 일부 기여하였으나, standing에서 오류가 증가하여 전체 성능 지표는 표 1과 같이 Raw 대비 개선으로 이어지지 않았다.

EMD 기반 증강은 신호의 시간적 변동성을 확장하지만, 교차 환경에서 발생하는 다중경로, 센서 배치, 반사 환경 등을 충분히 대변하지 못할 수 있다. 이로 인해 모델이 특정 환경에서의 신호 형태를 더 강하게 학습하거나 클래스 경계를 왜곡하여, 교차 환경 조건에서 성능 향상이 일관되게 나타나지 않는 것으로 해석된다.

III. 결 론

본 연구는 ESP32 기반 공개 WiFi CSI 데이터셋을 이용하여 행동 인식 모델의 교차 환경 성능을 분석하였다. 원시 CSI 데이터에 대해 EMD 기반 증강을 적용하고, 증강 유무에 따른 교차 환경 일반화 성능을 비교하였다. 실험 결과, EMD 증강은 환경 조합에 따라 성능 개선과 저하가 동시에 관찰되어 교차 환경 조건에서 성능 향상을 일관되게 보장하지 못하였다. 또한 일부 조건에서 예측이 특정 클래스로 집중되는 실패 양상이 관찰되어, 증강이 환경 간 분포 차이를 완화하지 못하거나 편향을 강화할 가능성을 확인하였다. 향후에는 클래스 편향을 억제하는 학습 전략과 환경 변화에 강인한 표현 학습을 결합하여 교차 환경 조건에서의 성능 안정성을 확보할 계획이다.

ACKNOWLEDGMENT

This research was supported by Global - Learning & Academic research institution for Master's/PhD students, and Postdocs(G-LAMP) Program of the National Research Foundation of Korea(NRF) grant funded by the Ministry of Education(No. RS-2025-25441283)

참 고 문 헌

- [1] Y. Ma et al., "WiFi Sensing with Channel State Information: A Survey," ACM Computing Surveys, vol. 52, no. 3, pp. 1-36, 2019.
- [2] J. A. Armenta-Garcia et al., "Tools and Methods for Achieving Wi-Fi Sensing in Embedded Devices," Sensors, vol. 25, no. 19, 6220, 2025.
- [3] Z.-Y. Lim et al., "Radio frequency-based human activity dataset collected using ESP32 microcontroller in line-of-sight and non-line-of-sight indoor experiment setups," Data in Brief, vol. 57, 111101, 2024.
- [4] N. E. Huang et al., "The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Non-Stationary Time Series Analysis," Proceedings of the Royal Society A, vol. 454, no. 1971, pp. 903-995, 1998.

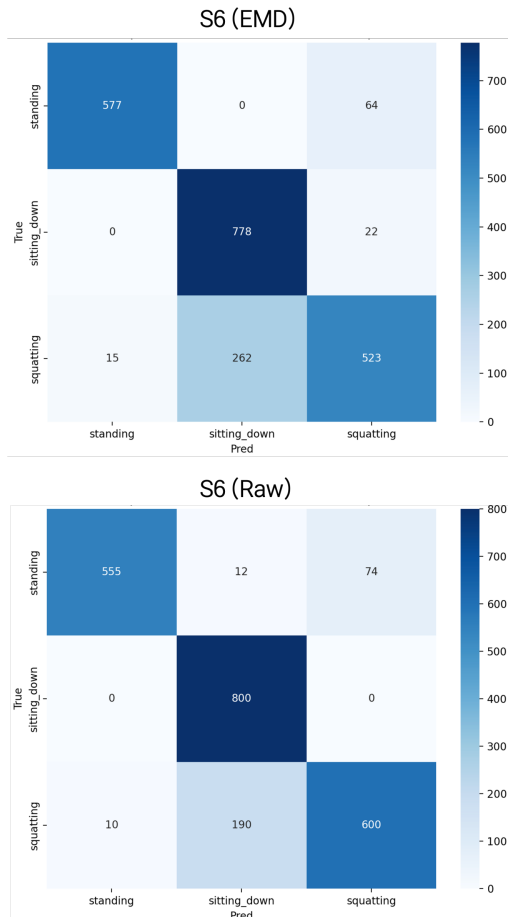


그림 1. 시나리오4(S4)에서 EMD 및 Raw 혼동행렬