

MobileViT 기반 자동 변조 인식 연구

변예린¹, 채명호², 임완수*³

¹금오공과대학교, ²국방과학연구소, ³성균관대학교

*wansu.lim@skku.edu

Novel Automatic Modulation Classification Based on MobileViT

Yerin Byeon¹, Myoungcho Chae², Wansu Lim*³

¹Kumoh National Institute of Technology, ²Agency for Defense Development, ³Sungkyunkwan University

요약

자동 변조 인식(Automatic modulation Recognition, AMR)은 무선 통신 시스템의 주요 과정 중 하나로, 변조 방식을 자동으로 인식하여 통신 시스템의 성능을 최적화하는 기술이다. 최신 AMR 연구는 CNN과 Transformer와 같은 딥러닝 모델을 사용하여 성능을 고도화시키고 있다. 그러나 딥러닝 모델은 연산 복잡도와 메모리 요구 사항이 매우 높으므로 실시간 처리와 자원이 제한된 디바이스 적용에 제약이 있다. 따라서 최근 AMR 연구는 경량화된 딥러닝 모델을 사용하여 변조 인식 성능을 유지하면서 저사양 하드웨어에서도 동작할 수 있는 연구가 활발히 진행 중이다. 이에, 본 논문은 ViT를 경량화한 MobileViT 기반 AMR 기법을 제안한다. 실험 결과 제안한 MobileViT 모델은 ViT와 CNN 모델 대비 -2dB에서 각각 약 6%와 8% 더 높은 인식 정확도를 보였으며, ViT 모델 대비 파라미터 수가 약 40배 줄어든 결과를 보였다.

I. 서론

자동 변조 인식(Automatic modulation Recognition, AMR)은 변조 방식을 자동으로 인식하고 분류하는 통신 시스템의 중요 기술이다. AMR은 통신 시스템의 성능 최적화, 간섭 최소화 그리고 데이터 전송 신뢰성을 높이는 데 매우 중요한 역할을 한다. AMR 연구는 오랫동안 진행됐으며 최근에는 CNN과 Transformer와 같은 딥러닝 기반 모델을 AMR에 적용하여 변조 인식 성능을 개선하는 연구가 진행되고 있다^{[1]-[2]}.

[1]과 [2]는 각각 CNN기반과 Transformer 기반으로 AMR 기법을 제안하여 변조 방식을 분류하였으며, [2]는 Vision Transformer를 통해 성능을 패치 단위로 나누어 학습한 후 변조를 분류하였다. 그러나 이러한 딥러닝 모델은 높은 연산 복잡도와 큰 메모리 요구 사항으로 인해 실시간 처리와 자원이 제한된 하드웨어 적용에 어려움이 있다. 따라서 최신 AMR 연구는 딥러닝 모델의 경량화가 중요한 문제로 인식되고 있다.

이러한 필요성을 바탕으로 본 논문은 ViT의 경량화 모델인 MobileViT를 이용한 자동 변조 인식 기법을 제안한다. MobileViT는 CNN을 경량화한 모델인 MobileNetV2 블록^[3]과 지역-전역 정보를 동시에 추출할 수 있는 Transformer계층으로 구성된 MobileViT 블록^[4]을 결합하여 구성하였다.

II. MobileViT 기반 자동 변조 인식 알고리즘

그림 1 (a)는 제안한 MobileViT기반 자동 변조 인식 흐름도이며, Convolution 계층, MobileNet V2 블록, MobileViT 블록, 글로벌 풀링 계층, 완전 연결 계층 그리고 분류를 위한 Softmax 계층으로 구성한다. 그림 1 (b)는 MobileNetV2 구조이고 Depth Wise Convolution을 이용하여 기존 CNN 구조를 경량화하였고, 실험 결과 기존 CNN 대비 계산량을 약 9배 줄일 수 있었다. 그림 1 (c)는 MobileViT Block 구조이고 “Local representation”, “Global representation”, “Fusion” 등 총 3단계로 구성한다. 첫 번째 단계 “Local representation”은 지역적인 정보를 추출하는 과정이다. 입력 데이터의 크기는 $H \times W \times C$ 이고, 이를 커널 크기가 $n \times n$

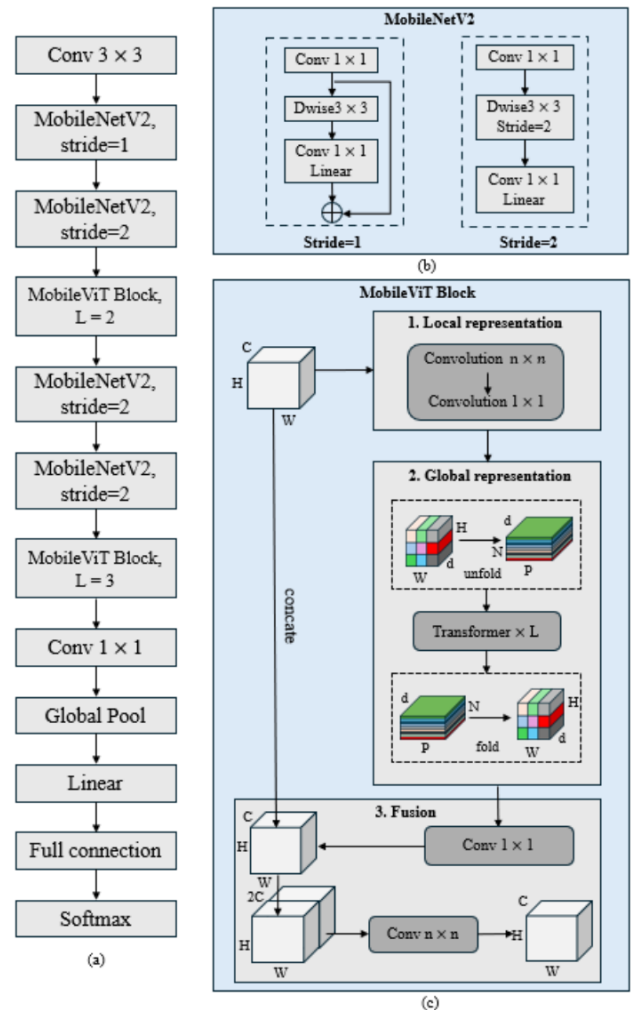


그림1. (a) 제안한 MobileViT 기반 자동 변조 인식 흐름도, (b) MobileNetV2 구조, (c) MobileViT Block 구조

Convolution 계층에 입력하여 특징을 추출한다. 추출한 특징은 커널 크기가 1 x 1인 Convolution 계층에서 선형 투영되어 HxWxd 크기의 결과를 얻는다. 두 번째 단계 “Global representaion”은 첫 번째 단계의 출력을 입력으로 받아 전역적인 정보를 추출하는 과정이다. 입력 데이터는 서로 겹치지 않는 PxNxd 크기의 패치로 펼쳐져 Transformer의 입력으로 사용한다. 이후 Transformer에서는 Attention 메커니즘을 이용하여 패치 간 관계를 학습한다. Transformer에서 전역적인 정보가 인코딩된 데이터 크기는 PxNxd이고, 이를 fold 과정을 통해 두 번째 단계 입력 크기인 HxWxd로 복원한다. 세 번째 단계 “Fusion”은 두 번째 단계의 결과 데이터를 MibileViT 블록의 초기 입력 데이터와 결합하여 정보 손실을 줄인다. 결합한 데이터의 크기는 HxWx2C이고, 커널 크기가 n x n인 Convolution 계층을 통해 Fusion하여 출력 크기가 HxWxC인 최종 결과를 도출한다.

결과적으로 제안한 MobileViT 모델은 MobileNetV2 블록과 MobileViT 블록을 여러 개 쌓아 입력 데이터의 지역적 그리고 전역적 특징을 다양한 시각으로 분석하여 학습을 진행한다. 이는 노이즈와 간섭이 높은 무선 신호 분석에 강점으로 작용하여 변조 인식 정확도를 높일 수 있었다.

III. 실험

본 논문은 RadioML2016.10a 데이터를 이용하였고 그 중 8PSK, BPSK, QPSK 등 총 3가지 변조 방식을 사용하였다. 각 변조 방식은 -20dB부터 18dB까지 2dB 증가하며, 각 dB당 1000개 샘플을 사용하였다. 또한, 각 데이터를 성상도 이미지로 변환하여 사용하였으며, 생성한 데이터의 학습, 검증, 테스트 비율은 각각 8, 1, 그리고 1이다. 학습 시 Epoch를 100으로 하였고, 배치 크기는 64를 사용했다. 옵티마이저는 Adam을 사용하였고, 손실함수는 categorical cross entropy를 사용했다.

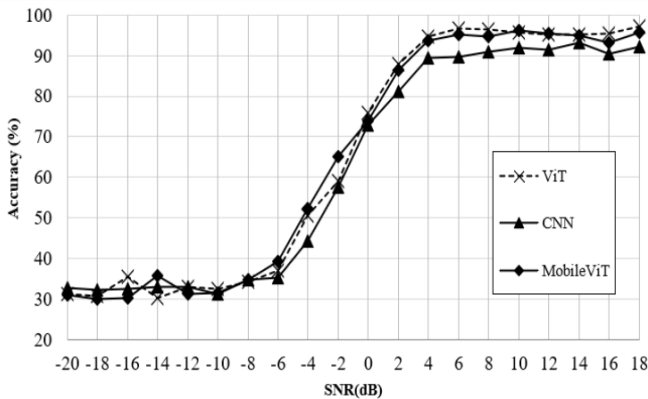


그림 2. MobileViT, ViT, CNN의 변조 방식 분류 정확도 비교

그림 2는 SNR에 따른 변조 방식 인식 정확도 그래프이다. 제안한 MobileViT 기반 변조 인식 기법은 벤치마크 모델인 CNN과 ViT와 성능을 비교했다. 실험 결과 MobileViT, ViT 그리고 CNN의 인식 정확도는 -2dB에서 각각 65.11%, 59.13% 그리고 57.56%의 정확도를 가진다. 또한, MobileViT는 4dB 이상부터 90% 이상의 정확도를 보여주며 최대 96.19%의 인식 성능을 달성했다. ViT와 MobileViT의 모델 파라미터 수는 각각 약 20M 그리고 약 0.5M으로 제안한 기법이 기존 대비 약 40배 감소하였다. 따라서 제안한 MobileViT는 기존의 ViT 대비 복잡도는 매우 낮으나 분류 성능은 높은 것으로 확인되어 자동 변조 인식을 효율적이고 효과적으로 수행할 수 있음을 증명했다.

IV. 결론

본 논문은 MobileViT 모델을 기반으로 자동 변조 인식 기법을 제안했다. MobileViT는 MobileNetV2와 MobileViT block을 결합하여 모델을 경량화하고 변조 방식 인식 정확도를 향상시켰다. 실험 결과 MobileViT는 ViT와 CNN보다 -2dB에서 각각 약 6%, 8% 더 높은 정확도를 보이며, ViT보다 약 40배 적은 모델 파라미터를 가지고 준수한 정확도를 달성했다. 따라서, 제안한 MobileViT는 자동 변조 인식에 적합한 경량 모델이라고 할 수 있다.

ACKNOWLEDGMENT

이 논문은 방위사업청의 재원으로 국방과학연구소의 지원을 받아 수행된 연구임 (311JJ5-912967201)

참고 문헌

- [1] Yeongmin Jeong, Jihwan Park, Hyunwoo Cho, Myoungcho Chae, Wansu Lim, “Research on Simultaneous Recognition of Modulation and Channel Coding using CNN,” pp.769-797, Nov.2023
- [2] Minju Lee, Youjeong Yoon, Myoungcho Chae, Wansu Lim. “Study of Vision Transformer-based Auto-Modulation Recognition,” Proceedings of Symposium of the Korean Institute of communications and Information Sciences, pp. 927-928, Oct. 2023.
- [3] SachinMehta,Mohammad Rastegari,"MobileViT:Light-weight, General-purpose, and Mobile-friendly Vision Transformer," conference paper at ICLR 2022, Oct. 2021.
- [4] Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, Liang-Chieh Chen"MobileNetV2: Inverted Residuals and Linear Bottlenecks," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Jun. 2018.