

Data-efficient image Transformers(DeiT)를 이용한 자동변조인식 기술

윤 유 정*, 정 해°, 임 완 수°°

Automatic Modulation Recognition Using Data-Efficient image Transformers(DeiT)

Youjeong Yoon*, Hae Chung°, Wansu Lim°°

요 약

자동변조인식(AMR, Automatic Modulation Recognition)은 무선통신 시스템에서 수신된 신호의 변조 방식을 자동으로 식별하는 기술로, 방해 신호를 감지하는 데 중요한 역할을 하고, 군사 및 상업 통신 시스템의 성능을 향상하는 기술이다. 최근 딥러닝의 발전으로 인해, AMR 분야에서도 딥러닝 기술들이 도입되어 활발한 연구가 진행되고 있다. 본 논문에서는 대규모 데이터 없이도 효율적으로 높은 성능을 달성할 수 있는 DeiT(Data-efficient Image Transformers) 모델을 사용하여 자동으로 변조 방식을 인식하는 새로운 알고리즘을 제안한다. DeiT 모델은 지식 증류 기법을 활용하여, CNN 아키텍처의 귀납적 편향을 유지하면서 Vision Transformer(ViT)의 구조적 장점을 결합했다. 실험 결과 DeiT기반 AMR이 ViT를 사용한 변조인식 보다 정확도가 평균 8.2% 높았다.

키워드 : 자동변조인식, 6G, Data-efficient Image Transformers(Deit), 무선통신시스템

Key Words : Automatic modulation recognition, 6G, Data-efficient Image Transformers(Deit), Wireless communication systems

ABSTRACT

Automatic Modulation Recognition (AMR) is a technology that automatically identifies the modulation of a received signal in a wireless communication system, which plays an important role in detecting jamming signals and improves the performance of military and commercial communication systems. With the recent advances in deep learning, AMR has been actively researched by introducing deep learning techniques in the field of AMR. In this paper, we propose a new algorithm to automatically recognize modulation schemes using the Data-efficient Image Transformers (DeiT) model, which can efficiently achieve high performance without the need for large datasets. The DeiT model utilizes knowledge distillation techniques to combine the advantages of each model by using the structure of Vision Transformer (ViT) while retaining the inductive bias of CNN architecture. Experimental results show that AMR based on DeiT is on average 8.2% more accurate than modulation recognition using ViT.

※ 본 연구는 금오공과대학교 대학 학술연구비로 지원되었음 (2021).

♦ First Author : Kumoh National Institute of Technology (KIT), School of Electronic Engineering, 6234sa@kumoh.ac.kr, 학생회원

° Corresponding Author : Kumoh National Institute of Technology, School of Electronic Engineering, hchung@kumoh.ac.kr, 종신회원

°° Corresponding Author : Sungkyunkwan University, School of Electronic and Electrical Engineering, wansu.lim@skku.edu, 정회원

논문번호 : 202408-183-A-RN, Received August 22, 2024; Revised October 16, 2024; Accepted October 23, 2024

I. 서 론

자동변조인식(Automatic Modulation Recognition, AMR)은 무선 통신 시스템에서 수신된 신호의 변조 방식을 자동으로 인식하고 분류하는 기술로, 6G와 같은 차세대 통신 기술의 발전과 함께 그 중요성이 더욱 부각되고 있다^[1]. 군사, 무선 통신, 레이더 등 다양한 분야에서 활용되는 AMR은 통신 시스템의 효율성을 높이고, 신호 간섭을 최소화하며, 네트워크 안정성을 향상하는데 중요한 역할을 한다^[2]. 초기 AMR 연구는 주로 통계적 방법과 기계 학습 기법을 활용했으나, 최근에는 인공지능의 발전, 특히 딥러닝 기술의 발전으로 연구의 질이 크게 향상되었다^[3]. 예를 들어, 딥러닝 모델 중 하나인 심층 신경망은 복잡한 변조 패턴을 정밀하게 분석하고 분류하여 실시간 데이터 처리와 대규모 데이터 학습 효율성을 크게 높였다. 이러한 기술의 발전은 통신 시스템의 자동화와 최적화를 가능하게 하며, 신속하고 정확한 정보 처리로 시스템 비용 절감 등에 이바지할 수 있어 AMR에 관한 연구가 활발하게 진행되고 있다^[4-5].

인공지능 기술이 발전하면서 Convolutional Neural Networks(CNN)과 Recurrent Neural Networks (RNN) 등 기법을 AMR에 적용한 연구가 활발히 진행되고 있다^[4]. [4]는 입력 신호를 스펙트로그램으로 변환한 후, CNN을 이용하여 다양한 변조 유형을 자동으로 분류하는 기법을 제안한다. 이 모델은 입력된 스펙트로그램 데이터에서 중요한 특성을 추출하기 위해 여러 합성곱 층과 풀링 층을 활용한다. 각 레이어는 특정 주파수와 시간 특성에 반응하여 변조 신호의 복잡성을 효과적으로 해석하여 분류 정확도를 높인다. [5]는 변조 인식을 두 종류의 학습 데이터를 활용한 CNN 모델들을 결합한 새로운 딥러닝 기법을 제시한다. 이 방식에서는 동상 및 직교 성분의 신호 샘플을 활용하여 학습되며, 이는 상대적으로 구분이 용이한 변조 방식을 식별하는 데 유용하다. 또한 기존 CNN과 달리, 전통적인 풀링 기법 대신 드롭아웃 기법을 적용함으로써, 16-QAM 및 64-QAM과 같은 복잡한 변조 방식도 인식할 수 있다. 하지만 CNN은 로컬 패턴에 초점을 맞춰 학습하기 때문에 변조 신호 간의 상호작용이나 전역적인 문맥 정보를 파악하는 데 어려움을 겪는다. 또한, 이러한 특성 때문에 AMR에서 다양한 변조 신호를 실시간으로 처리하는 상황에 적응력이 낮을 수 있다.

[6]는 시퀀스 데이터의 연속성을 효과적으로 처리하는 RNN을 자동 변조 인식에 활용했다. 특히, IQ 상관관계와 장기적 시간적 특징을 이용하여 변조 신호를 분류했다. 순차적인 정보 학습을 위한 계층 CNN과 직렬

로 연결하여 시퀀스 신호의 시간적 특성을 효율적으로 학습했다. 이로 인해 CNN이 단독으로 사용할 때 나타나는 시간적 의존성 부족 문제를 극복했다. 또한, IQ 상관 관계 정보를 추출한 후 신호를 2차원에서 1차원으로 축소하는 방식으로 모델을 경량화하여 높은 분류정확도를 달성했다. 하지만 AMR과 같은 복잡한 변조 신호에서 긴 시퀀스 데이터를 다룰 때 기울기 소실 문제와 병렬 처리의 어려움으로 인해 효율성이 떨어지는 단점이 있다. 또한, RNN의 순차적 처리 특성은 병렬 계산을 어렵게 만들어 AMR 환경에서 대규모 신호 데이터의 학습 시간을 증가시킨다.

이러한 문제를 해결하기 위해 Transformer 모델이 등장했다. Transformer는 주로 자연어 처리 분야에서 시작되었으며, 어텐션 메커니즘(attention mechanism)을 이용하여 입력 데이터를 동시에 처리함으로써 기존 RNN의 시퀀스 종속성 문제를 극복했다^[7]. 또한, Transformer는 CNN이 어려움을 겪는 전역적인 컨텍스트 정보를 효과적으로 통합할 수 있어, AMR과 같은 비시퀀스 데이터에도 유용하다는 잠재력을 보여주었다.

최근 연구들은 이러한 Transformer의 구조적 장점을 활용하여 AMR 기술에 적용 가능한 Vision Transformer (ViT)를 개발했다. ViT는 변조 신호 인식 분야에서도 Transformer 아키텍처의 큰 가능성을 보여주었으며, 기존 CNN과 비슷한 성능을 보이면서 변조 인식 분야에 새로운 방향을 제시한다. [3]과 [7]은 방대한 양의 데이터로 초기 학습을 수행한 후, 테스트에서 다른 모델과 비교하여 우수한 성능을 보여준다. 이 과정에서 필요한 계산 리소스는 상대적으로 적은 것으로 나타나, 효율적인 학습 과정을 가능하게 한다. 그러나 ViT 모델은 뛰어난 성능을 발휘하기 위해 방대한 양의 데이터가 필요하다는 한계가 있으며, 파라미터의 수가 많아 모델이 무겁고 계산 비용이 높아 대규모 데이터 수집이 어려운 AMR 환경에서 성능 저하로 이어지는 단점이 있다.

이에 본 논문은 대규모 데이터가 없이도 높은 성능을 달성할 수 있는 Data efficient image Transformers (DeiT) 모델을 사용하여 변조 방식을 자동으로 인식하는 알고리즘을 제안한다. DeiT는 CNN 모델을 'teacher model'로 하여 'student model'이 'teacher model'로부터 어텐션 메커니즘을 통해 학습하는 지식 증류 기법(Knowledge Distillation)을 채택하여 사용한다.

DeiT모델은 CNN 아키텍처가 가진 장점 중 하나인 귀납적 편향을 유지하면서, ViT와 같은 구조로 되어 있다^[8]. 이러한 접근 방식은 기존 ViT와 RNN모델의

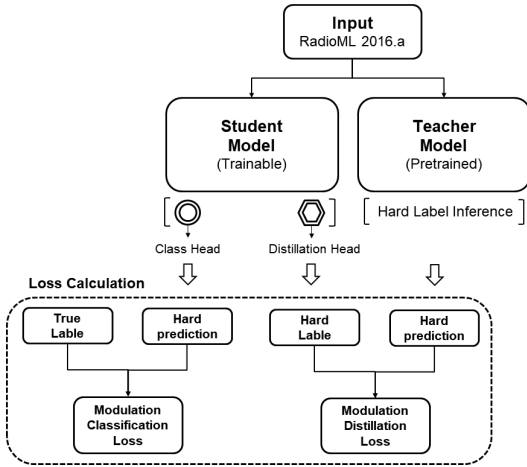


그림 1. 제안한 DeiT 모델 구조
Fig. 1. Proposed DeiT model structure

한계인 대규모 데이터와 기울기 소실 문제점을 극복하여, DeiT 모델이 기존 모델 보다 적은 양의 데이터로도 우수한 변조 인식 성능을 낼 수 있게 한다.

그림 1은 제안한 시스템 구조를 나타내며, 크게 ‘Data Input’, ‘student model & teacher model’, ‘Loss Calculation’의 세 부분으로 구성되어 있다. 입력 데이터는 AMR을 위한 벤치마크 데이터인 RadioML 2016.10a를 사용하여 다양한 통신 시스템에서의 변조 탐지와 식별을 가능하게 한다. 또한 무선 신호 처리 연구에서 다양한 변조 기법과 신호대잡음비(SNR) 조건을 포함해 실제 통신 환경을 반영한 모델 개발을 할 수 있다.

데이터 입력 부분에서는 student model과 teacher model에 같은 데이터가 공급된다. 이어지는 학습 모델 부분에서는 student model이 트랜스포머 기반 구조를 통해 입력 이미지를 처리한다. Loss Calculation 단계에서는 분류와 증류 두 가지 예측을 통해 손실을 계산하고, 이를 통해 모델을 최적화 한다. teacher model은 사전 학습된 상태로 고정되어 있으며, student model의 성능 향상을 위해 하드 레이블을 제공한다. 최종적으로 증류 헤드는 teacher model의 예측과 student model의 출력을 비교하여 손실 함수를 통해 결과를 분류하여 출력한다.

II. DeiT 증류 기법

신경망 훈련에서의 증류^[8]는 그림 2와 같이 일반적으로 크고 잘 훈련된 teacher model로부터 작은 모델인 student model로 지식을 전달하는 과정이다. 이 과정은

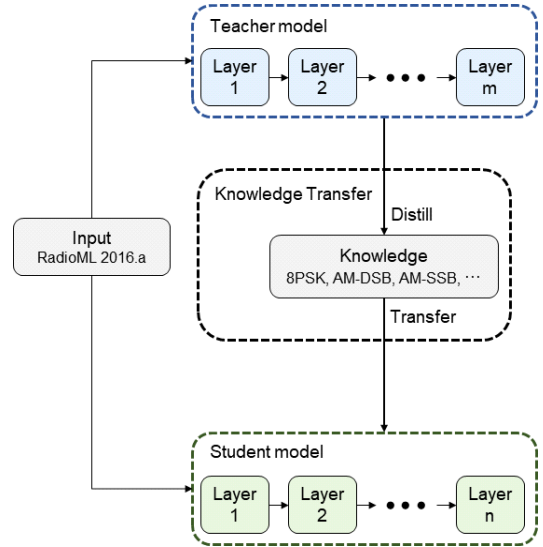


그림 2. 증류 과정
Fig. 2. Distillation process

teacher model의 출력을 활용하여 student model의 학습을 지도하며, 적은 데이터나 계산 자원으로도 성능을 향상할 수 있다. 증류는 그림 1의 ‘Loss Calculation’에서 볼 수 있듯이 주로 Soft, Hard 두 가지 형태로 구분된다. 이 증류 과정에서 teacher model의 지식이 student model에 전달되는 것을 출력된 로짓(logit) 또는 확률을 통해 이루어지며, 이 값들은 student model이 목표로 하는 출력을 따라 하도록 유도한다. 이러한 방식은 student model이 teacher model의 예측 능력을 흉내 내어 보다 정확한 예측을 할 수 있게 만들어 준다.

2.1 소프트 증류(Soft Distillation)와 레이블 스무딩(label smoothing)

소프트 증류(Soft Distillation)와 레이블 스무딩(Label Smoothing)은 모델의 학습을 안정시키고 일반화 성능을 향상하는 데 중요한 역할을 한다^[9]. 소프트 증류는 teacher model이 소프트맥스 함수를 통해 생성한 클래스 확률, 즉 소프트 레이블(soft labels)을 사용한다. 이 소프트 레이블들은 확률 분포를 더 부드럽게 만들어 하드 레이블(hard labels)보다 각 클래스에 대해 더 많은 정보를 제공한다. 이는 레이블 스무딩 기법과 유사한데, 레이블 스무딩은 교차 엔트로피 손실 계산 시 각 클래스에 대한 확신을 줄여 모델의 과적합을 방지하고 결정 경계를 더 부드럽게 만드는 데 도움을 준다. 또한 모델이 각 클래스에 대한 정보를 더 균형 있게 학습하도록 하여 일반화 성능을 높이는 데 기여한다.

2.2 하드 증류(hard Distillation)

하드 증류는 teacher model의 확실한 클래스(hard labels)를 사용하며, student model 훈련에서는 교차 엔트로피 손실 함수를 사용하여 student의 학습이 teacher의 결정을 따라 할 수 있도록 한다. 하지만 소프트 레이블은 이미지와 레이블 간의 불일치를 일으킬 수 있는 문제가 있으며, 때로는 잘못된 학습신호를 모델에 제공할 수 있다. 이러한 문제점을 완화하기 위해 DeiT는 하드 레이블을 사용하고 레이블 스무딩을 추가로 적용한다. 이를 통해 소프트 레이블의 문제점을 완화하고 레이블 스무딩을 통해 모델이 과도하게 특정 클래스에 의존하는 것을 방지하여 더욱 더 안정적이고 효과적인 모델을 구축할 수 있다.

III. DeiT를 이용한 자동변조인식

그림 3은 RadioML 2016.a 데이터를 입력으로 사용하는 DeiT 모델의 전체 구조를 시각화한 것이다. 모델은 크게 ‘패치 임베딩’, ‘위치 임베딩’, ‘트랜스포머 블록’, ‘분류 헤드와 증류 헤드’로 총 다섯 단계로 구성되어 있다. 첫 번째 단계인 패치 임베딩은 입력 데이터를 여러 개의 패치로 나누어 트랜스포머에 적합한 형태로 변환한다. 두 번째 단계인 위치 임베딩은 각 패치의 공간적 위치 정보를 모델에 제공한다. 세 번째 단계는 트랜스포머 블록으로, 여러 인코더 레이어를 통해 입력 데이터를 처리한다.

각 인코더 레이어는 레이어 정규화, 멀티 헤드 어텐

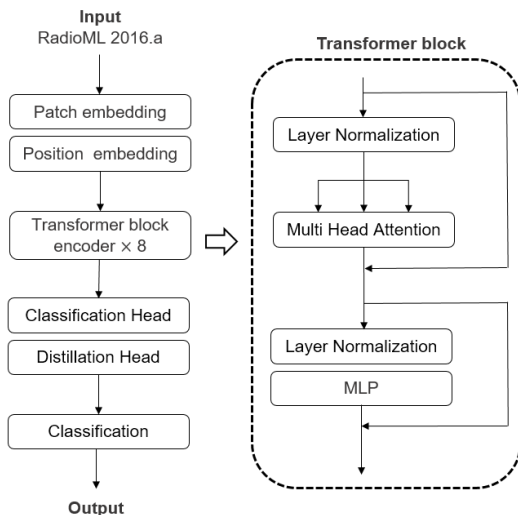


그림 3. DeiT 구조
Fig. 3. DeiT architecture

션, 다시 레이어 정규화 그리고 MLP(Multi-Layer Perceptron)로 구성되어 복잡한 특징들을 추출하고 학습한다. 마지막 단계에서는 분류 헤드와 증류 헤드가 최종적으로 클래스를 예측하여 결과를 출력한다.

3.1 Embedding and Token

그림 4는 DeiT 모델에서 클래스 토큰과 증류 토큰의 증류 과정을 보여준다. 입력 데이터는 먼저 RadioML 2016.a 데이터를 사용하여 채널의 수가 2개이고 길이가 128인 2×128 형태의 IQ 신호로 생성한다. 데이터는 패치 임베딩 과정을 통해 크기가 2×2 인 여러 개의 패치로 나누어진다. 각 패치는 원본 데이터의 연속된 부분을 포함하며, 이렇게 분할된 각 패치는 개별적인 정보 단위로 처리된다. 패치 임베딩 이후, 각 패치에는 위치 임베딩이 적용된다. 위치 임베딩은 각 패치의 공간적인 순서와 위치 정보를 모델에게 제공하여, 모델이 패치 간의 상대적인 위치를 인식할 수 있도록 한다. 임베딩된

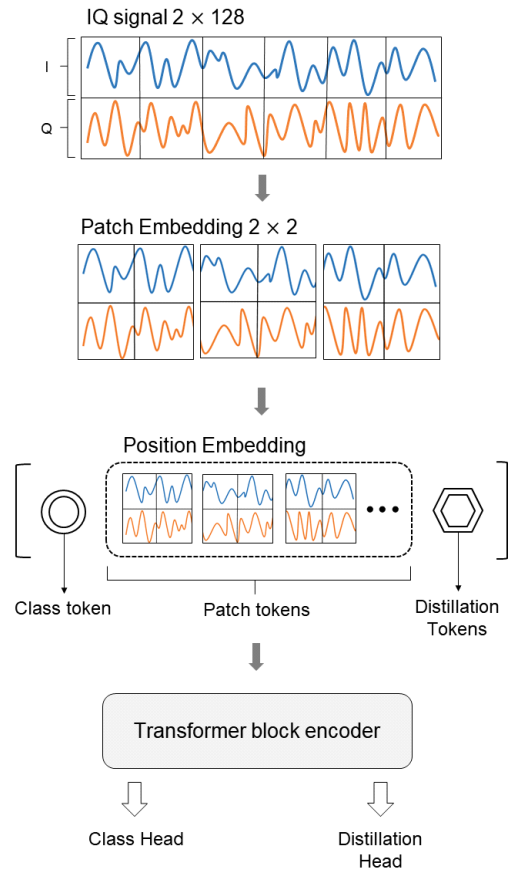


그림 4. DeiT 임베딩 과정
Fig. 4. DeiT Embedding process

데이터는 추가적으로 클래스 토큰과 증류 토큰과 함께 트랜스포머 블록으로 입력된다. 클래스 토큰은 각 입력 데이터가 속한 카테고리를 예측하는 데 사용된다. 이 토큰은 트랜스포머 블록을 통과하면서 각 패치로부터 중요한 정보를 집약하여 최종적으로 분류 헤드를 통해 클래스를 예측한다.

클래스 토큰(Class token):

AMR에서 클래스 토큰은 각 신호가 속한 클래스를 예측하기 위해 도입된다. 트랜스포머의 각 레이어를 거치면서 클래스 토큰은 입력 신호의 패치들로부터 정보를 집약하며, 최종적으로 분류 헤드를 통해 신호의 변조 종류를 결정하는 데 중요한 역할을 한다. 이 토큰은 모델이 특정 변조 클래스에 대한 학습을 강화하도록 설계되어, AMR의 분류 정확도를 높이는 데 사용된다.

증류 토큰(Distillation token):

증류 토큰은 추가적인 teacher 학습 정보를 모델에 제공한다. 이 토큰은 teacher model로부터 얻은 하드 레이블 정보를 활용하여, student model이 다양한 변조 신호를 더욱 정확하게 학습할 수 있도록 돕는다. AMR에서 증류 토큰은 증류 헤드를 통해 처리되며, teacher model의 변조 예측 간의 차이를 최소화하여 모델의 일반화 능력을 향상한다.

DeiT 모델에서 클래스 토큰과 증류 토큰을 동시에 사용하여 teacher model의 지식을 증류 받는 방식은 모델의 학습 효율성과 정확도를 동시에 향상한다. 이러한 양방향 학습 메커니즘은 DeiT 모델이 더욱 복잡하고 다양한 데이터에 대해 높은 성능을 발휘할 수 있도록 도와주며, 특히 AMR 기술을 통해 구별하기 어려운 변조 유형도 더 쉽게 식별할 수 있게 한다.

3.2 Transformer Encoder block

트랜스포머 블록은 총 세 단계의 레이어로, ‘레이어 정규화’, ‘멀티 헤드 어텐션’으로 이루어져 있다. 첫 번째 단계인 ‘레이어 정규화’는 모델 입력의 통계적 특성을 표준화하여 학습 과정의 안정성과 효율성을 증가시키는 역할을 한다. 두 번째 단계인 ‘멀티 헤드 어텐션’은 데이터의 다양한 특징을 동시에 포착하여 변조 신호의 중요한 정보를 강조하고 불필요한 정보를 억제한다. 이 단계는 특히 다양한 변조 유형을 다루는 AMR에서 신호의 패턴을 정확하게 식별하는 데 유리하다. 마지막 단계에서는 MLP를 통해 최종적인 데이터의 복잡한 패턴을 더 잘 이해하고 표현한다.

트랜스포머 블록은 전체적으로 총 8번 반복된다. 이러한 반복 구조는 AMR에서 더 깊고 복잡한 패턴을 학습할 수 있게 하며, 반복할 때마다 특징의 추출과 통합 과정을 통해 더욱 정밀한 데이터 표현을 가능하게 한다. 반복적인 어텐션 메커니즘은 AMR에서 신호의 더욱 정확한 정보를 포착하고, 최종적으로 높은 성능의 분류 결과를 도출할 수 있다. 이러한 구조적 배치와 반복적인 과정은 DeiT 모델이 복잡한 데이터에서도 높은 정확도로 작업을 수행할 수 있도록 설계되어 있으며, 특히 다양한 변조 형태와 신호 잡음을 효과적으로 구별하는 데 큰 이점을 제공한다.

3.3 Classification

트랜스포머 모델의 마지막 단계에서 데이터는 분류 헤드와 증류 헤드를 통해 데이터가 처리된다. 분류 헤드는 입력 데이터의 클래스를 예측하며, 생성된 예측은 실제 변조 신호 레이블과 비교되어 분류 로스를 계산한다. 증류 헤드는 teacher model의 하드 레이블 예측과 student model의 출력을 비교하여 증류 로스를 계산하고, 이 로스는 student model의 AMR 정확도와 일반화 능력을 향상한다. 최종적으로 모델은 로스 계산을 바탕으로 변조 신호 데이터의 클래스를 정확하게 분류한다.

IV. 시뮬레이션 및 성능평가

4.1 실험환경

이번 절은 DeiT 모델을 훈련하기 위한 파라미터를 설명한다. 입력 데이터로 RadioML 2016.10a를 사용한다. 이는 무선 신호를 인식하고 분류하기 위해 설계된 벤치마크 데이터로, 실제와 유사한 조건에서의 알고리즘 성능을 검증할 수 있다. DeiT 모델에서 데이터는 다양한 SNR(Signal-to-Noise Ratio)에서 총 11가지 변조 유형(8PSK, AM-DSB, AM-SSB, BPSK, CPFSK, GFSK, PAM4, 16QAM, 64QAM, QPSK, WBFM)을 포함하고 있으며, SNR 범위는 -20dB에서 18dB까지 2dB 간격으로 설정하였다. 데이터는 각 변조별로 20,000개로, 총 220,000개의 IQ(In-phase Quadrature) 샘플로 구성하였다. 학습 데이터와 테스트 데이터의 비율은 8:2로 설정되어, 총 176,000개의 샘플이 학습용으로, 나머지 44,000개의 샘플이 테스트용으로 사용된다.

최적화 알고리즘으로 AdamW가 사용되며, 학습률은 0.001, 가중치 감소율은 0.0001로 설정되어 있다. 배치 크기는 400으로 설정하였고 트랜스포머 인코더 레이어를 반복하는 수는 8개이며, 각 패치의 크기는 [2, 2]로 설정하였다. 활성화 함수로는 소프트맥스가 사용

되고, 드롭아웃 비율은 0.1이며, 입력 데이터의 차원은 [2, 128]이다. 논문에서 사용된 파라미터 설정값은 표 1에 자세히 정리하였다. DeiT에 가장 최적화된 에포크 수를 찾기 위해 에포크별 정확도 비교를 그림 5에 정리하였다. 실험 결과 220에포크에서 제일 좋은 성능을 보였으며 -4dB, 0dB, 16dB 모두 약 20% 이상의 성능 향상을 보였다. 과적합 현상을 방지하기 위하여 early stopping을 적용하였고, 성능이 제일 높고 과적합이 되지 않는 지점인 220 에포크에서 훈련하였다.

표 1. 제안한 DeiT 파라미터
Table 1. Proposed DeiT parameters

Optimizer	AdamW	Learning rate	0.001
		Weight Decay	0.0001
Batch Size	400	Epoch	220
Transformer Encoder	8	Patch size	[2, 2]
Activation Function	Softmax	Resolution	128
Dropout	0.1	Parameter	5.9M

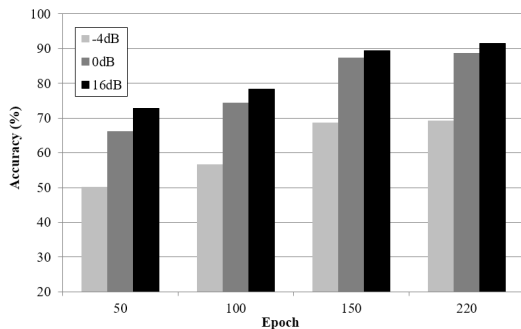


그림 5. Epoch별 정확도 비교
Fig. 5. Accuracy comparison depending on Epoch

4.2 실험 결과

그림 6은 DeiT 모델에 대한 SNR에 따른 변조 방식별 정확도를 혼돈행렬(Confusion matrix)로 나타낸 그림이다. (a)부터 (d)까지 각각 SNR이 -10dB, -6dB, 10dB, 16dB일 때의 모델 성능을 나타낸다. SNR이 -10dB일 때 8PSK와 QPSK가 SSB로 오인되는 것은 SSB 역시 동상 성분(변조할 신호)과 직교 성분(변조할 신호의 힐버트 변환)을 가지고 있기 때문이다. 또한 SSB-LC (Large Carrier)는 동상 성분의 크기가 직교 성분에 비해 매우 크기 때문에 PSK를 SSB로 오인하게 된다. 이러한 현상은 SNR이 -6dB일 때 확연히 드러난

다. 결국 SNR이 낮을 때, 디지털 변조가 백색잡음과 섞이면 아날로그 변조로 오인될 가능성이 높고, 아날로그 변조가 디지털 변조로 오인될 가능성은 낮은 것으로 보인다. 반면, SNR이 높아지면 잡음의 영향이 줄어들어 명확히 구분되는 것을 볼 수 있다.

그림 (c)와 (d)를 보면 16-QAM과 64-QAM을 혼돈하고 있는 것을 볼 수 있다. 이러한 변조별 정확도 변동 추이를 더욱 명확하게 확인하기 위해 SNR에 대한 4가지 변조 방식 정확도를 그림 7로 나타내었다. AM-DSB의 정확도는 SNR이 낮은 -12dB에서는 22.06%였으나, SNR이 4dB 이상에서는 100%로 매우 높았다. CPFSK는 -8dB에서는 15.5%로 낮은 정확도를 보이지만 2dB 이상 SNR에서는 100%로 AM-DSB와 비슷한 양상을 보이며 높은 정확도를 달성한다. 8PSK는 -6dB에서는 17.78%로 매우 낮은 정확도를 보이지만 0dB에서는 82.83%로 높아진 것을 볼 수 있고 14dB은 100%의 분류 정확도가 보이는 등 4dB 이상 SNR에서는 평균 96.97%로 매우 높은 정확도를 달성한다. 하지만 64-QAM은 -6dB에서 81.69%로 16-QAM보다 높은 정확도를 보이지만 2dB에서 제일 높은 정확도인 86.67%로, 다른 변조 방식에 비해 낮은 정확도를 보인다. 16-QAM과 64-QAM 변조 방식은 심볼 간의 거리가 매우 가까워 잡음과 감쇠 등에 의해 변조된 심볼들이 서로 구분하기 어려워진다. 또한 WBFM 변조 방식은 넓은 대역폭을 사용하고 오디오 신호의 다양성으로 인해 처리해야 할 데이터의 양이 많고 신호 패턴이 복잡해져 분류하기 어렵다. 이러한 이유로 16-QAM, 64-QAM, WBFM이 다른 변조 방식과 달리 비교적 낮은 정확도를 달성할 수 있는 것으로 예상할 수 있다.

그림 8은 SNR에 따른 DeiT와 ViT 모델의 전체 변조기법에 대한 분류 정확도의 평균을 비교한 결과이다. 특히, SNR이 -4dB에서 DeiT의 정확도는 약 69.28%로 ViT의 52.91%보다 분류 정확도가 높은 것을 볼 수 있다. 또한 SNR이 -2dB에서는 DeiT가 81.43%의 정확도를 보인 반면 ViT는 64.69%에 그쳤다. 최대 정확도를 비교해 보면 DeiT는 91.47%의 최대 정확도를 보인 반면, ViT는 80.14%의 최대 정확도를 나타냈다. 이러한 결과는 DeiT가 파라미터가 6.4M인 ViT와 비교하여 더 적은 파라미터를 가지고 있음에도 불구하고 낮은 SNR 환경뿐만 아니라 높은 SNR 환경에서도 ViT 모델보다 안정적으로 우수한 성능을 발휘함을 알 수 있다. 이를 통해 DeiT가 ViT보다 데이터가 제한적인 상황에서 높은 성능을 가지고 있음을 보여준다.

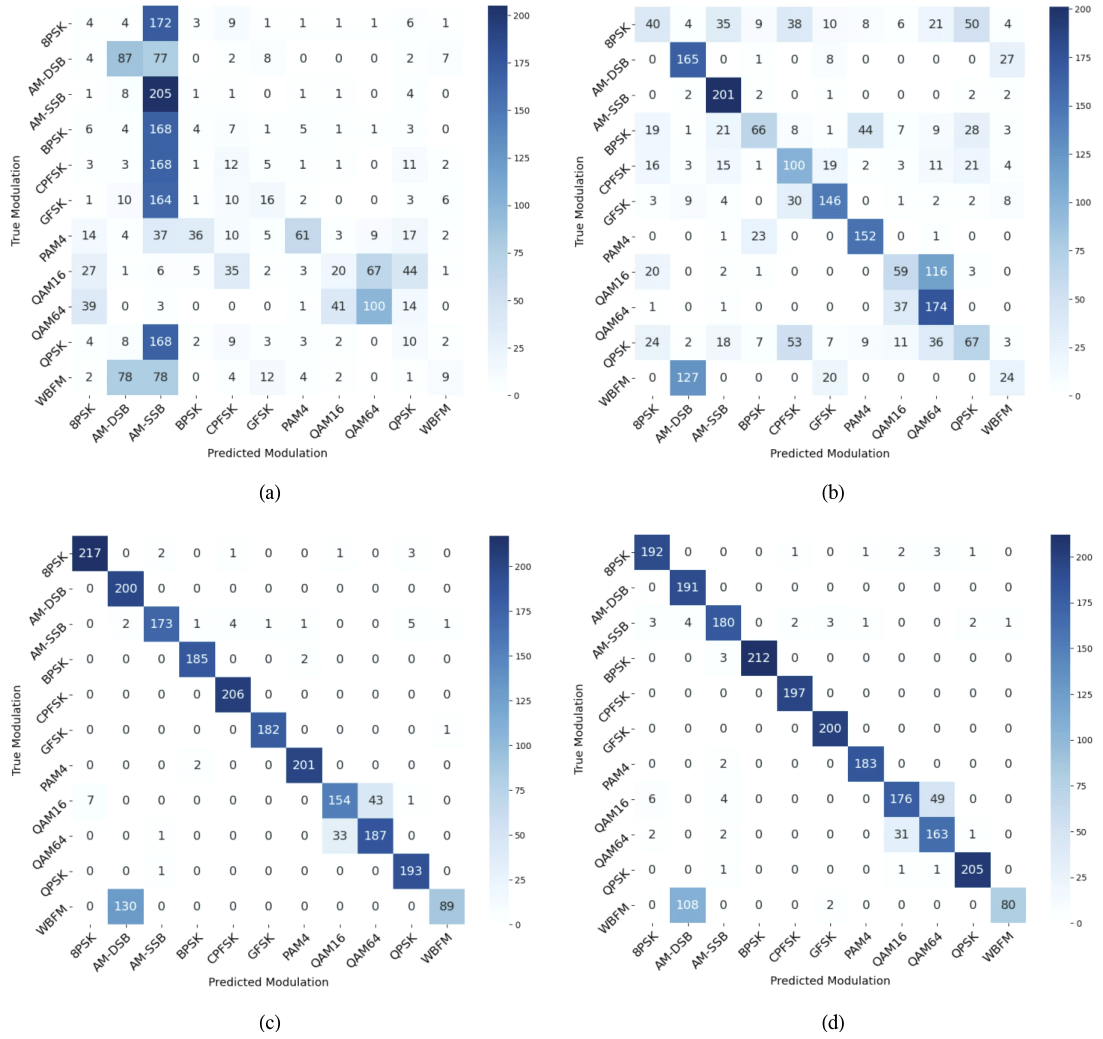


그림 6. SNR에 따른 혼동행렬 : (a) -10dB, (b) -6dB, (c) 10dB, (d) 16dB,
Fig. 6. Confusion matrix as a function of SNR : (a) -10dB, (b) -6dB, (c) 10dB, (d) 16dB

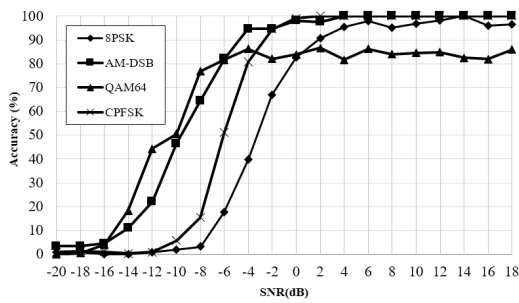


그림 7. DeiT 모델 기반 각 변조별 정확도
Fig. 7. Accuracy for each modulation based on DeiT

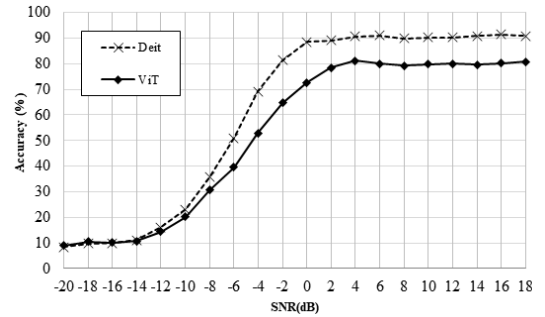


그림 8. 전체 변조 기법에 따른 DeiT, ViT의 정확도 비교
Fig. 8. Comparing the accuracy of DeiT and ViT with total modulation

V. 결 론

본 논문은 DeiT 모델을 기반으로 하여 자동변조인식 기법을 제안한다. 변조 방식은 총 11가지(8PSK, AM-DSB, AM-SSB, BPSK, CPFSK, GFSK, PAM4, 16-QAM, 64-QAM, QPSK, WBFM)로, IQ데이터를 사용하여 모델 학습 및 데이터 분류를 진행하였다. 시뮬레이션 결과 ViT보다 -6dB이상에서 약 10% 이상 높은 성능을 보였고, 0dB에서 88.59%를 달성하여 낮은 SNR에서도 높은 분류 정확도를 낼 수 있음을 보여주었다. 이를 통하여 본 연구에서 사용한 DeiT모델이 변조 방식을 효과적으로 인식할 수 있을 뿐만 아니라, ViT보다 적은 데이터로 높은 정확도를 낼 수 있음을 검증하였다.

References

- [1] R. Ding, F. Zhou, Q. Wu, C. Dong, and Z. Han, "Data and knowledge dual-driven automatic modulation classification for 6G wireless communications," *IEEE Trans. Wireless Commun.*, vol. 23, no. 5, pp. 4228-4242, 2023.
- [2] H. Cho, M. Chae, and W. Lim, "Implementation of modulation and channel coding recognition using CNN and protocol reverse engineering simulation in blind communication environment," *J. KICS*, vol. 49, no. 11, pp. 1644-1657, 2024.
- [3] M. Lee, M. Chae, and W. Lim, "Research on automatic modulation recognition using vision transformer," *J. KICS*, vol. 49, no. 8, pp. 1074-1081, Early access, 2024.
- [4] S. Jeong, U. Lee, and S. Kim, "Spectrogram-based automatic modulation recognition using convolutional neural network," *2018 Tenth ICUFN IEEE*, pp. 843-845, 2018.
- [5] Y. Wang, M. Liu, J. Yang, and G. Gui, "Data-driven deep learning for automatic modulation recognition in cognitive radios," *IEEE Trans. Veh. Tech.*, vol. 68, no. 4, pp. 4074-4077, 2019.
- [6] C. TianShu, H. Zhen, W. Dong, and L. Ruike, "A iq correlation and long-term features neural network classifier for automatic modulation classification," *2022 IEEE 10th ICICN*, pp. 396-400, 2022.
- [7] A. Dosovitskiy, A. Beyer, L. Kolesnikov, A. Weissenborn, D. Zhai, X. Unterthiner, T. Dehghani, M. Minderer, M. Heigold, and G. Gelly, "An image is worth 16x16 words: Transformers for image recognition at scale," *Inter. Conf. Mach. Learn. PMLR, arXiv preprint arXiv:2010.11929*, 2020. (<https://doi.org/10.48550/arXiv.2010.11929>)
- [8] H. Touvron, M. Cord, M. Douze, F. Massa, A. Sablayrolles, and H. Jégou, "Training data-efficient image transformers & distillation through attention," *arXiv:2012.12877v2*, 2021. (<https://doi.org/10.48550/arXiv.2012.12877>)
- [9] G. Hinton, O. Vinyals, and J. Dean, "Distilling the knowledge in a neural network," *arXiv:1503.02531v1*, 2015. (<https://doi.org/10.48550/arXiv.1503.02531>)

윤 유 정 (Youjeong Yoon)

2022년 3월~현재 : 금오공과대학교 전자공학과
<관심분야> 자동변조인식, 머신러닝, 최적화

정 해 (Hae Chung)

1998년 8월~현재 : 금오공과대학교 전자공학부 교수
<관심분야> 통신 프로토콜 및 네트워크
[ORCID:0000-0002-6920-2714]

임 완 수 (Wansu Lim)

2024년 3월~현재 : 성균관대학교 전자전기공학부 교수
<관심분야> 통신 프로토콜, 기계학습, 자동변조인식
[ORCID:0000-0003-2533-3496]