

Swin Transformer 구조를 기반으로 한 Multi-Modal 자동 변조 인식 신경망

신 다 민*, 전 민 욱*, 김 형 남^o

Multi-Modal Automatic Modulation Recognition Network Based on the Swin Transformer Architecture

Da-Min Shin*, Min-Wook Jeon*, Hyung-Nam Kim^o

요 약

현대 전자전에서는 다양한 통신 신호에 대해 높은 인식 정확도를 제공하는 딥러닝 기반 자동 변조 인식에 관한 연구가 활발히 진행되고 있다. 이는 통신 기술의 발전으로 전장의 신호 환경이 점차 복잡해짐에 따라, 수신된 통신 신호의 분석에 많은 어려움이 수반되고 있기 때문이다. 본 논문에서는 1D Swin Transformer 구조를 기반으로 설계된 multi-modal 자동 변조 인식 신경망을 제안한다. 1D Swin Transformer는 계층적 특징 맵과 이동 윈도우 기반 접근을 통해 입력 데이터로부터 다양한 해상도의 특징을 효과적으로 추출한다. 제안하는 모델은 IQ 시계열 신호와 주파수 스펙트럼 정보를 동시에 활용하는 multi-modal 구조로 이루어져 있으며, 각 모달리티의 정보를 추출하여 1D Swin Transformer에 입력함으로써 통합된 특징을 이용하여 변조 기법을 인식한다. 본 논문에서 제안한 모델은 61.15%의 평균 정확도를 기록하여 기존 신경망보다 우수한 성능을 보였다. 또한 QAM 계열 신호에 대해서도 평균 57.00%의 높은 인식 정확도를 달성함으로써, 구조적으로 유사한 변조 신호 간의 구분에서도 강점을 보였다.

키워드 : 자동 변조 인식, 딥러닝, 1D Swin Transformer

Key Words : automatic modulation recognition, deep learning, 1D Swin Transformer

ABSTRACT

In modern electronic warfare, deep learning-based automatic modulation recognition (AMR) systems that provide high recognition accuracy across a wide range of communication signals have been actively studied. This trend stems from the growing complexity of signal environments in modern battlefields, resulting from the rapid evolution of communication technologies, which has significantly increased the difficulty of analyzing received signals. In this paper, we propose a multi-modal AMR neural network designed based on the 1D Swin Transformer architecture. The 1D Swin Transformer effectively extracts multi-resolution features from input sequences through a hierarchical feature representation and a shifted window-based mechanism. The proposed model employs a multi-modal structure that simultaneously utilizes IQ time-series signals in the time

※ 이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(RS-2025-00557790).

♦ First Author : Pusan National University, Department of Electrical and Electronics Engineering, ekals2020@pusan.ac.kr, 학생회원

° Corresponding Author : Pusan National University, School of Electrical and Electronics Engineering, hnkim@pusan.ac.kr, 종신회원

* Pusan National University, Department of Electrical and Electronics Engineering, gow1128@pusan.ac.kr, 학생회원

논문번호 : 202507-180-A-RU, Received July 29, 2025; Revised September 2, 2025; Accepted September 19, 2025

domain and spectrum information in the frequency domain. By extracting features from each modality and feeding them into the 1D Swin Transformer, the model performs modulation recognition by leveraging information from both the time and frequency domains. The proposed model achieved an average accuracy of 61.15% and outperformed conventional models. Furthermore, within the QAM family, the model achieved a high average recognition accuracy of 57.00% and demonstrated its ability to distinguish between structurally similar modulation schemes.

I. 서 론

현대 전자전에서는 새로운 통신 기법의 등장으로 인해 신호 환경이 복잡해지며 전장에서 정보 수집에 많은 제약이 따르고 있다^[1]. 이는 변화하는 전자전 환경에 대응하기 위한 다양한 기술 개발의 필요성을 초래하였다. 특히 자동 변조 인식은 적의 전자 장비로부터 방사된 비협조적 신호의 변조 기법을 추정하여 송신원을 식별 및 분류함으로써, 아군의 전략 수립에 도움을 줄 수 있다^[2-3]. 이에 따라 사전 정보 없이 수신된 미상 통신 신호의 변조 방식을 식별하는 자동 변조 인식 기술의 중요성이 크게 대두되고 있다.

기존의 변조 인식 기법은 우도 기반 기법과 특징 기반 기법으로 구분된다^[4]. 우도 기반 기법은 다중 가설 검정 문제를 해결함으로써 신호의 변조 기법을 추정하는 방식으로, 우도 함수를 설계하고 적절한 임계값을 설정하는 과정을 포함한다. 이 방식은 통계적으로 최적의 분류 성능을 제공할 수 있으나, 실제 사용 환경에서의 채널 변화나 간섭 등으로 인해 성능 열화가 발생할 수 있다^[5]. 특징 기반 기법은 수신된 신호로부터 대표적인 특징을 추출하고, 분류기를 통해 신호의 변조 방식을 예측하는 방법이다. 높은 SNR(signal-to-noise ratio) 환경에서는 특징 추출이 용이하다는 장점을 가지고 있으나, 시스템 구성의 복잡성과 경험적 검증에 대한 의존으로 인해 전문가의 개입이 필요하다는 한계가 있다^[6].

최근에는 딥러닝 기반의 자동 변조 인식 연구가 활발하게 진행되고 있으며, 기존 기법과 달리 전문가의 개입을 최소화하면서도 높은 인식 정확도를 보인다는 장점이 있다. 이러한 시스템에서 입력 신호의 표현 방식과 특징을 효과적으로 학습할 수 있는 신경망 구조는 성능에 핵심적인 영향을 미치기 때문에 분류 신경망의 설계는 매우 중요하다. 딥러닝 기술의 발전으로 CNN(convolutional neural network)^[6]이나 LSTM(long short-term memory)^[7]을 넘어 다양한 구조의 신경망이 제안되고 있으며, 그중에서도 Transformer 계열의 백본 구조가 여러 분야에서 우수한 성능을 보이면서 주목받고 있다. 그러나 Transformer는 구조적으로 국소적 특

징을 효과적으로 포착하기 어렵다는 한계를 지닌다. 이를 보완하기 위해 Transformer의 multi-head self attention을 shifted window attention으로 변형한 Swin Transformer^[8]가 제안되어 비전 분야의 백본 신경망으로 널리 사용되고 있다. 또한 이를 시계열 데이터에 적합하게 변형한 1D Swin Transformer가 자율 주행 및 신호 품질 평가를 위해 연구되었으며^[9-10], 시계열 데이터에도 Swin Transformer의 계층적 특징 맵, shifted-window partitioning과 같은 구조적 장점과 고정된 윈도우 크기로 인한 연산량 절감이 적용됨을 증명하였다. 그러나 Swin Transformer를 통신 신호처리 분야의 변조 인식 문제에 적용한 사례는 아직 제한적이다.

통신 신호는 진폭 및 위상, 주파수 스펙트럼 등 다양한 표현으로 변환할 수 있기 때문에 다양한 입력 데이터를 활용한 자동 변조 인식 신경망이 연구되고 있다. 입력 데이터로는 시계열 형태의 IQ(in-phase, quadrature) 신호가 주로 사용되며, 주파수 스펙트럼이나 큐몰런트와 같이 신호의 특징을 강조할 수 있는 1차원 데이터를 활용하기도 한다^[11]. 또한 성상도(constellation diagram)처럼 신호를 시각적으로 표현한 2차원 이미지 형태의 데이터를 입력으로 사용하는 경우도 있다^[12]. 이처럼 기존의 변조 인식 연구는 대부분 하나의 입력 모드에만 의존하는 단일 채널 입력 기반으로 이루어져 왔다. 그러나 최근에는 여러 분야에서 multi-modal 학습을 이용하여 다른 데이터 유형의 정보를 융합하고 있다^[13]. 이와 같이 다양한 형태의 입력 데이터를 통합적으로 활용하는 multi-modal 접근법은 복잡한 환경에서도 견고한 변조 인식 성능을 달성할 수 있는 방법으로 주목받고 있다.

이에 본 논문에서는 1D Swin Transformer 구조를 이용하여 특징을 추출하는 multi-modal 변조 인식 신경망을 제안하고, 자동 변조 인식 문제에 대한 1D Swin Transformer의 성능을 분석한다. 제안한 모델은 IQ 신호와 주파수 스펙트럼의 크기를 입력으로 하며, 시간 영역과 주파수 영역의 정보를 모두 활용함으로써 다양한 변조 방식의 특성을 효과적으로 포착한다. 입력된 신호는 우선 VGGNet^[14] 구조 기반 모듈을 통해 처리되

며, 이 과정에서 국소적인 시공간적 특징을 추출하여 데이터의 표현력을 향상시킨다. 이후 시간 영역과 주파수 도메인의 특징을 결합한 후, Swin Transformer 구조로 입력함으로써, 제안된 모델은 통신 신호의 국소적, 전역적 패턴을 동시에 학습하게 된다. 이때 신호의 국소적인 패턴에는 PSK(phase shift keying) 계열의 위상 천이, FSK(frequency shift keying) 계열의 주파수 도약, QAM(quadrature amplitude modulation) 신호의 심볼 변화 등이 있다. 또한 신호 전체에 적용되는 잡음, 채널 효과와 같은 전역적 특징이 존재한다.

본 논문의 구성은 다음과 같다. 2장에서는 Swin Transformer 구조와 이를 기반으로 설계한 변조 인식 신경망에 대해 기술한다. 3장에서는 통신 신호 데이터셋을 이용하여 제안하는 신경망을 학습시킨 결과에 대해 여러 관점에서 분석하며, 4장에서 본 논문의 결론을 맺는다.

II. 본 론

2.1 1D Swin Transformer

Swin Transformer는 CNN과 Transformer의 장점을 통합함으로써 이미지 분류, 의미론적 분할 등과 같은 비전 영역에서 널리 사용되고 있는 백본 구조이다^[8]. 이 구조는 데이터에 윈도우를 적용하고, 그 내부의 패치 간 attention 연산을 수행하며, 계층적 특징 맵을 이용하여 단계에 따라 연산 수행 범위를 조절한다. 이를 통해 연산 효율성을 유지하면서도 다양한 해상도의 특징을 추출할 수 있다. Swin Transformer의 이러한 구조적 장점으로 인해 시계열 데이터에 대한 확장 가능성이 제시되면서, 시계열 데이터에 적합하도록 변형한 1D Swin Transformer가 제안되었다^[9]. 여기에서는 일반적인 Swin Transformer와 동일하게 윈도우 범위 내에서 self-attention 연산을 단계적으로 수행하여 해상도를 점진적으로 축소한다. 또한 shifted-window partitioning을 사용하여 윈도우 간 연결성을 유도할 수 있도록 하였다.

그림 1은 4개의 계층을 가지는 Swin Transformer에 대해 윈도우와 패치 간의 관계를 나타낸 예시 그림이다. 윈도우 크기는 2로 고정되며, 블록이 깊어질수록 패치 크기가 두 배로 확장되어 해상도가 점차 축소된다. 또한 각 단계에서는 윈도우를 한 패치만큼 이동시켜, 인접 패치 간의 정보를 효과적으로 통합한다. 또한 블록에서 패치 간 크기 불일치가 발생하면 패딩을 적용하여 패치의 크기를 적절하게 조절한다. 이와 같은 계층적 구조를 통해 국소적인 영역에서부터 전역적인 영역까지 점진

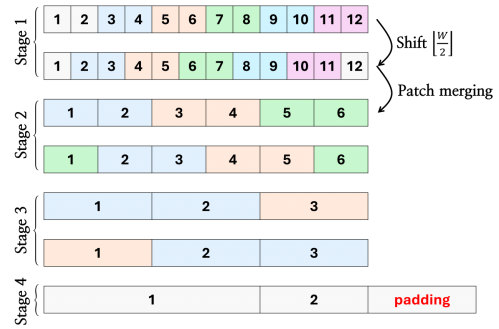


그림 1. 1D Swin Transformer 구조의 윈도우 및 패치 분할
Fig. 1. Windows and patch partitioning of Swin Transformer architecture.

적으로 추출 영역을 확장한다.

2.2 자동 변조 인식 신경망 구조

본 논문에서는 1D Swin Transformer를 기반으로 변조 인식을 위한 신경망 구조를 설계하였으며, 제안하는 모델의 전체 구조는 그림 2에 제시되어 있다. 제안한 모델은 시간 영역과 주파수 영역 정보를 동시에 활용하는 multi-modal 접근 방식으로, 두 입력은 각각 개별적으로 특징을 추출한 뒤 이후 동일한 1D Swin Transformer 기반 백본을 통해 통합적으로 처리된다.

시간 영역 입력은 I, Q 성분으로 구성된 2×128 크

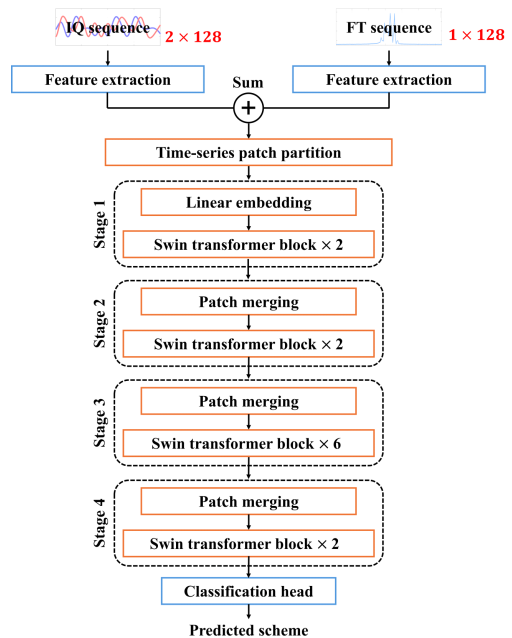


그림 2. 제안하는 신경망의 구조
Fig. 2. Structure of the proposed neural network.

기의 시계열 데이터이며, 주파수 영역 입력은 1×128 길이의 FFT(fast Fourier transform) 시퀀스의 크기를 사용하였다. 각각의 입력은 VGGNet 구조를 활용한 특징 추출 모듈을 통해 저차원 특징을 학습하며, 3×3 크기의 컨볼루션 필터를 13번 사용한다. 이를 통해 입력 내의 지역적 특성을 학습하며, 4×512 크기의 특징 벡터가 출력된다. 이렇게 추출된 두 입력의 특징 벡터는 합산되어 1D Swin Transformer 기반 백본에 입력된다. 이후 패치 분할 과정을 통해 시계열 데이터를 일정한 길이의 패치로 나누고, 각 패치를 저차원 임베딩 벡터로 변환한다. 이후 변환된 임베딩 벡터는 총 4개의 단계로 구성된 백본 네트워크를 통과하며, 각 단계는 패치 병합 모듈과 1D Swin Transformer 블록으로 이루어져 있다. 이러한 구조를 기반으로 통신 변조 신호의 지역적, 전역적 특징을 동시에 학습하고자 하였다. 이후 출력 특징은 평탄화(flatten)된 후, 완전 연결층(fully connected layer)을 통해 변조 방식에 대한 최종 분류를 수행한다. 본 모델은 CNN 기반의 로컬 패턴 학습과 Transformer 기반의 전역 문맥 학습의 장점을 결합함으로써, 다양한 변조 방식에 대해 강인하고 정밀한 분류를 가능하게 한다.

III. 모의실험 및 성능평가

3.1 데이터셋

본 연구에서는 신경망 학습을 위해 통신 변조 신호 데이터셋인 RadioML2016.10A^[15]를 사용하였다. 해당 데이터셋은 3종의 아날로그 변조 방식(AM-DSB, AM-SSB, WBFM)과 8종의 디지털 변조 방식(8PSK, AM-SSB, WBFM)과 8종의 디지털 변조 방식(8PSK, BPSK, CPFSK, GFSK, PAM4, 16-QAM, 64-QAM, QPSK)을 포함하고 있으며, 총 220,000개의 신호 샘플

로 구성되어 있다. 신호는 SNR이 -20 dB부터 18 dB 까지 2 dB 간격으로 분포되어 다양한 통신 환경을 모사하고 있다. 각 샘플은 길이 128의 시계열 IQ 데이터이며, 실제 통신 환경에서 발생할 수 있는 주파수 오프셋 및 다중 경로 페이딩과 같은 열화 요소가 포함되어 있어 신경망의 일반화 성능을 평가하기에 적합하다. 그림 3은 각 변조 방식에 대한 시계열 신호의 예시를 시각적으로 나타낸 것이다. 데이터는 학습, 검증, 테스트 용도로 각각 176,000개, 22,000개, 22,000개로 분할하여 사용하였으며, 학습 안정성을 확보하기 위해 학습 데이터의 통계값을 기반으로 평균은 0, 표준 편차를 1로 정규화를 수행하였다.

3.2 신경망 및 학습 파라미터

제안하는 모델은 데이터의 특성 및 하드웨어 성능에 따라 신경망의 파라미터를 조절할 수 있으며, 본 논문에서 수행한 모의실험에서는 표 1에 제시된 파라미터를 사용하였다. 신경망은 표 2의 하이퍼 파라미터를 사용하여 학습하였으며, 검증 손실이 10 에포크 연속으로 감소하지 않으면 학습을 조기 종료하였다. 테스트 시에는 학습 과정에서 가장 낮은 검증 손실을 기록한 시점의 모델 파라미터를 사용하였다.

표 1. 모의실험에 사용된 신경망의 파라미터
Table 1. Parameters of the Swin Transformer used in the simulation.

Number of stages	4
Number of Swin Transformer blocks per stage	2, 2, 6, 2
Initial patch size	16
Window size	31

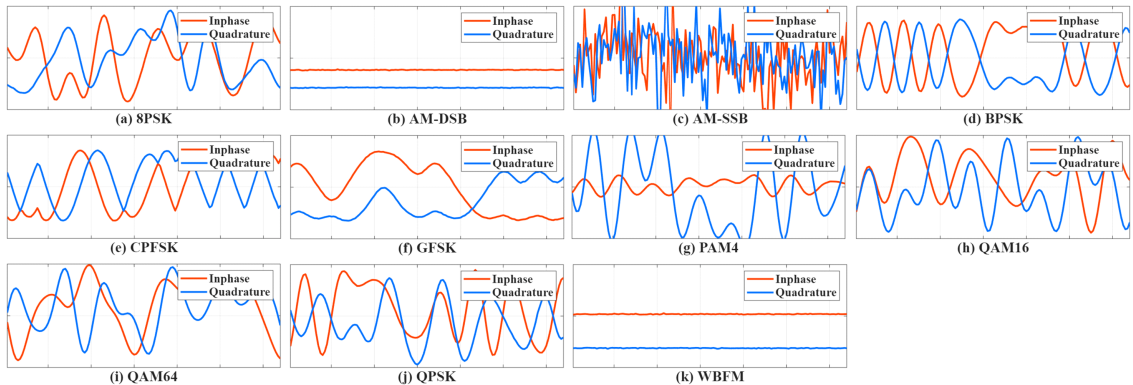


그림 3. RadioML2016.10A 데이터 예시 (SNR = 18 dB)
Fig. 3. Examples of RadioML2016.10A (SNR = 18 dB).

표 2. 신경망 학습에 사용된 하이퍼 파라미터
Table 2. Hyperparameter used in neural network training.

Loss function	Cross-entropy
Optimizer	Adam ^[16]
Batch size	128
Initial learning rate	10^{-3}
Learning rate scheduling	Decreased by 20% every 5 epochs

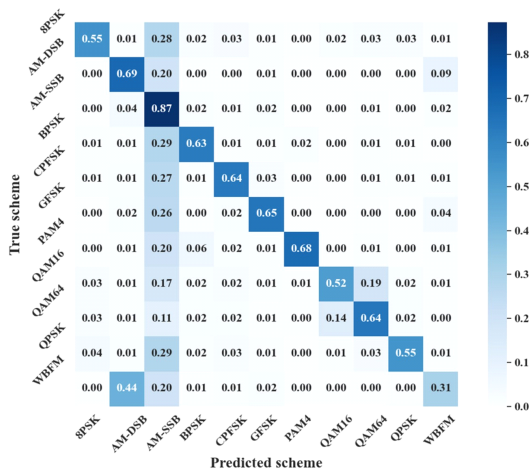
3.3 성능 분석

본 논문에서는 신경망에 대한 성능 평가 척도로 SNR에 따른 변조 인식 정확도와 혼동행렬(confusion matrix)을 사용하였다. 그림 4는 SNR 범위에 따른 혼동행렬을 나타낸 것으로, 그림 4 (a)는 전체 SNR 범위에 대한 결과이고, 그림 4 (b)는 높은 SNR 환경에서의 결과를 보여준다. 그림 4 (a)에서는 잡음의 영향으로 인해 타 변조 기법을 AM-SSB로 오분류하거나, 16-QAM과 64-QAM 간의 혼동이 발생하는 것을 확인할 수 있다. 반면 그림 4 (b)에서는 대부분의 변조 방식에서 매우 우수한 분류 성능을 보이며, AM-SSB에 대한 오분류가 줄어들고 QAM 계열 신호의 인식 성능이 향상된 것을 알 수 있다.

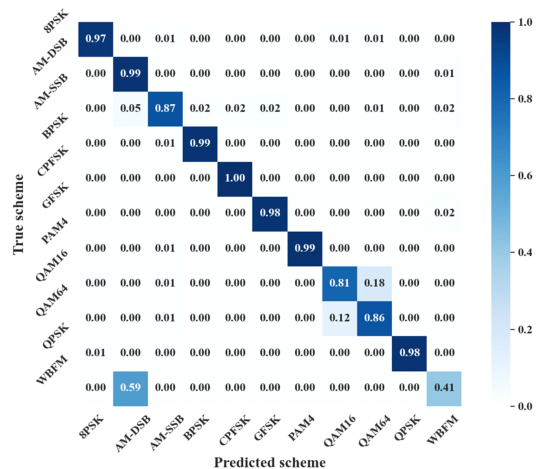
낮은 SNR 환경에서는 잡음 신호의 I, Q 성분이 AM-SSB 신호와 유사한 패턴을 보이기 때문에, 다양한 신호가 AM-SSB로 잘못 분류되는 현상이 발생하는 것으로 분석된다. 또한 16-QAM과 64-QAM은 진폭과 위상을 함께 변조하는 방식으로, 변조 기법의 특성상 유사

한 성상도를 가진다. 특히 변조 차수가 증가할수록 심볼 간의 간격이 좁아지기 때문에 잡음이나 채널 왜곡에 매우 민감해진다. 이러한 이유로 자동 변조 인식 분야에서 QAM 계열의 인식 성능을 향상시키는 것이 중요한 과제로 여겨진다.

이에 따라 본 연구에서는 QAM 계열 변조 방식(QPSK, 16-QAM, 64-QAM)에 대한 다양한 모델들의 분류 성능을 비교하여 제안 모델의 성능을 보다 구체적으로 분석하였다. 표 3에서 SNR -20 dB~18 dB 구간에서 세 가지 디지털 변조 신호에 대한 모델별 분류 정확도 및 평균 정확도를 제시하며, 비교 대상으로는 CNN, VGGNet, LSTM, CGDNet, Transformer를 사용하였다. 본 논문에서 제안하는 모델(Proposed)은 평균 57.00%를 기록하여 다른 모델 대비 우수한 인식 성능을 입증하였다. 특히 각 변조 기법에 대해 높은 분류 성능을 가지면서도, QPSK, 16-QAM, 64-QAM 모두에서 55%, 52%, 64%로 고른 성능을 보였다. 반면 국소적 특징 추출에 강점을 가지는 CNN과 VGGNet의 경우 각각 16-QAM과 64-QAM으로 인식 결과가 치우치는 것을 볼 수 있다. 이는 전역적 특징 추출의 미비함으로 인해 신호 전체에 걸쳐있는 고차 QAM 신호의 심볼의 분포를 효과적으로 학습하지 못했기 때문으로 해석할 수 있다. 이러한 결과를 통해 Swin Transformer 기반 모델은 shifted window를 기반으로 데이터의 국소적 패턴을 정밀하게 인식하고, 계층적인 구조를 통해 전역적인 특징과의 관계를 함께 학습함으로써, 고차 변조 방식 간의 미세한 차이를 효과적으로 구분할 수 있음을 확인하였다.



(a)



(b)

그림 4. 혼동행렬 (a) 전체 SNR(-20~18 dB) (b) 높은 SNR(0~18 dB)

Fig. 4. Confusion matrix : overall SNR region (-20~18 dB) (b) high SNR region (0~18 dB).

그림 5는 모든 변조 기법에 대해 다양한 SNR에 따른 분류 정확도를 나타낸 것으로 각각의 변조 방식의 잡음에 대한 강건도를 보여주고 있다. 대부분의 변조 기법은 SNR이 증가함에 따라 정확도가 증가하여 99.00% 이상의 정확도로 수렴하지만, AM-SSB, 16-QAM, 64-QAM은 약 80~90%, WBFM은 40%로 수렴한다. WBFM이 다른 변조 기법에 비해 낮은 정확도로 수렴하는 것은 대역폭이 넓고, 복잡한 신호 패턴을 가지고 있기 때문이다^[19].

그림 6은 SNR 변화에 따른 제안한 신경망과 기존 신경망의 변조 인식의 정확도를 보여준다. 여기에서는 신경망의 성능을 비교하기 위해서 SNR 별로 모든 변조 기법에 대한 정확도의 평균값을 인식 정확도로 사용하였으며, 다양한 잡음 환경에서의 성능 차이를 종합적으로 확인할 수 있다. SNR이 증가함에 따라 모든 모델에서 분류 정확도가 향상되며, 제안한 모델은 -5 dB 이상부터 기존 기법보다 확연히 높은 정확도를 가진다. 또한

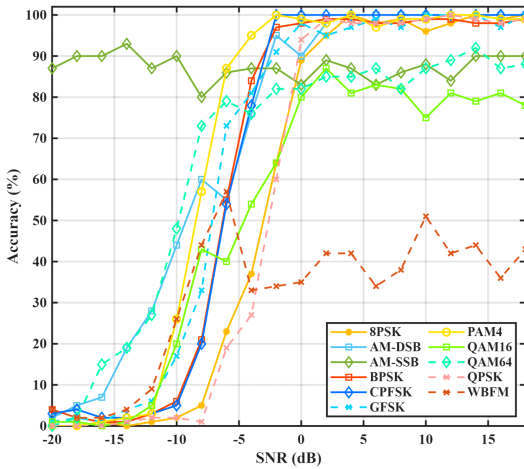


그림 5. 변조 기법별 SNR에 따른 분류 정확도
Fig. 5. Classification accuracy by modulation scheme across different SNR levels.

표 3. QAM 계열에 대한 모델별 정확도 (-20 dB ~ 18 dB)

Table. 3. Model-wise accuracy (-20 dB ~ 18 dB) for the QAM family of modulation schemes.

Modulation scheme Model	QPSK	16-QAM	64-QAM	Average recognition accuracy
CNN ^[6]	30 %	46 %	25 %	33.67 %
VGG ^[14]	55 %	21 %	61 %	45.67 %
LSTM ^[6]	53 %	51 %	50 %	51.33 %
CGDNet ^[17]	51 %	48 %	40 %	46.33 %
Transformer ^[18]	56 %	53 %	49 %	52.67 %
Proposed	55 %	52 %	64 %	57.00 %

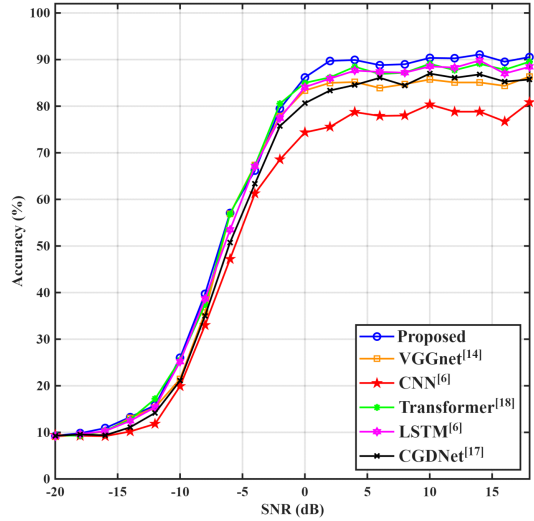


그림 6. 신경망 모델별 변조 방식 인식 정확도 비교
Fig. 6. Comparison of modulation recognition accuracy across neural network models.

SNR 5 dB 이상의 높은 SNR 구간에서도 제안 모델은 90% 내외의 정확도를 안정적으로 유지하는 반면, 비교 모델들은 상대적으로 낮은 정확도로 수렴하는 경향을 보인다. 이는 본 모델이 다양한 환경과 변조 방식에 강건함을 의미한다.

IV. 결 론

본 논문에서는 통신 신호 변조 인식을 위해 1D Swin Transformer 기반의 multi-modal 신경망을 제안하였다. 제안된 모델은 시계열 형태의 IQ 신호와 주파수 영역의 주파수 스펙트럼에서 추출한 특징을 결합하여 1D Swin Transformer의 입력으로 활용한다. RadioML2016.10A 데이터셋을 사용하여 모델을 학습하였으며, SNR에 따른 정확도와 혼동행렬을 사용하여

성능 분석을 수행하였다. 제안한 모델은 평균 정확도 61.15%, 최고 정확도 91.09%로 기존 모델과 비교하여 우수한 성능을 기록하였다. 이를 통해 Swin Transformer 구조가 비전 영역에 국한되지 않고, 시계열 데이터 기반의 자동 변조 인식 영역에서도 뛰어난 성능을 보임을 확인하였다. 또한 혼동행렬 분석 결과, 구조적으로 유사한 QAM 계열에 대해서도 57.00%의 인식 정확도를 달성하여 기존 신경망에 비해 혼동 확률이 줄어드는 양상을 보였다. 이러한 분석 결과는 제안된 모델이 잡음이나 채널 효과가 존재하는 환경에서도 변조 기법의 특징을 효과적으로 학습하고 정확하게 구분할 수 있음을 명확히 보여준다. 향후 연구에서는 각 입력 모달리티로부터 추출한 특징을 더욱 정교하게 융합할 수 있는 도메인 간 통합 전략 및 attention 기반 결합 기법에 대해 추가적인 분석을 수행할 예정이다.

References

- [1] Y. Peng, et al., "Automatic modulation classification using deep residual neural network with masked modeling for wireless communications," *Drones*, vol. 7, no. 6, p. 390, 2023.
(<https://doi.org/10.3390/drones7060390>)
- [2] W. Xiao, Z. Luo, and Q. Hu, "A review of research on signal modulation recognition based on deep learning," *Electr.*, vol. 11, no. 17, p. 2764, 2022.
(<https://doi.org/10.3390/electronics11172764>)
- [3] S.-Z. Hsue and S. S. Soliman, "Automatic modulation recognition of digitally modulated signals," in *Proc. IEEE MILCOM*, pp. 645-649, 1989.
(<https://doi.org/10.1109/MILCOM.1989.104004>)
- [4] O. A. Dobre, A. Abdi, Y. Bar-Ness, and W. Su, "Survey of automatic modulation classification techniques: Classical approaches and new trends," *IET Commun.*, vol. 1, no. 2, pp. 137-156, Apr. 2007.
(<https://doi.org/10.1049/iet-com:20050176>)
- [5] F. Meng, P. Chen, L. Wu, and X. Wang, "Automatic modulation classification: A deep learning enabled approach," *IEEE Trans. Veh. Technol.*, vol. 67, no. 11, pp. 10760-10772, Nov. 2018.
(<https://doi.org/10.1109/TVT.2018.2868698>)
- [6] M. Zhang, Y. Zeng, Z. Han, and Y. Gong, "Automatic modulation recognition using deep learning architectures," *Int. Wkshp SPAWC*, pp. 1-5, 2018.
(<https://doi.org/10.1109/SPAWC.2018.8446021>)
- [7] S. Hochreiter and J. Schmidhuber, "Long short-term memory," in *Neural Computation*, vol. 9, no. 8, pp. 1735-1780, Nov. 1997.
(<https://doi.org/10.1162/neco.1997.9.8.1735>)
- [8] Z. Liu, et al., "Swin transformer: Hierarchical vision transformer using shifted windows," *ICCV*, pp. 10012-10022, 2021.
(<https://doi.org/10.48550/arXiv.2103.14030>)
- [9] L. Shi, Y. Chen, M. Liu, and F. Guo, "DuST: Dual swin transformer for multi-modal video and time-series modeling," *IEEE/CVF Conf. CVPRW*, pp. 4537-4546, Seattle, WA, USA, 2024.
(<https://doi.org/10.1109/CVPRW63382.2024.00456>)
- [10] G. Chen, et al., "SwinDAE: Electrocardiogram quality assessment using 1D swin transformer and denoising autoEncoder," *IEEE J. Biomedical and Health Inf.*, vol. 27, no. 12, pp. 5779-5790, Dec. 2023.
(<https://doi.org/10.1109/JBHI.2023.3314698>)
- [11] S. Peng, S. Sun, and Y.-D. Yao, "A survey of modulation classification using deep learning: Signal representation and data preprocessing," in *IEEE Trans. Neural Netw. and Learn. Syst.*, vol. 33, no. 12, pp. 7020-7038, Dec. 2022.
(<https://doi.org/10.1109/TNNLS.2021.3085433>)
- [12] S. Peng, et al., "Modulation classification based on signal constellation diagrams and deep learning," in *IEEE Trans. Neural Netw. and Learn. Syst.*, vol. 30, no. 3, pp. 718-727, Mar. 2019.
(<https://doi.org/10.1109/TNNLS.2018.2850703>)
- [13] P. Xu, X. Zhu, and D. A. Clifton, "Multimodal learning with transformers: A survey," *arXiv preprint arXiv:2206.06488*, 2022.
(<https://doi.org/10.48550/arXiv.2206.06488>)
- [14] K. Simonyan and A. Zisserman, "Very deep

convolutional networks for large-scale image recognition,” *arXiv preprint arXiv:1490.1556*, 2014.

(<https://doi.org/10.48550/arXiv.1409.1556>)

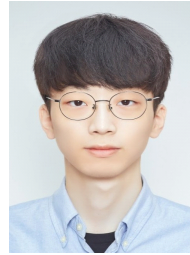
- [15] T. J. O’shea and N. West, “Radio machine learning dataset generation with GNU radio,” in *Proc. GNU Radio Conf.*, vol. 1, pp. 1-6, 2016.
- [16] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” in *Proc. 3rd ICLR*, 2014.
(<https://doi.org/10.48550/arXiv.1412.6980>)
- [17] J. N. Njoku, M. E. Moroch-Cayamcela, and W. Lim, “CGDNet: Efficient hybrid deep learning model for robust automatic modulation recognition,” *IEEE Netw. Lett.*, vol. 3, no. 2, pp. 47-51, 2021.
(<https://doi.org/10.1109/LNET.2021.3057637>)
- [18] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, et al., “Attention is all you need,” *Advances in NIPS*, vol. 30, 2017.
(<https://doi.org/10.48550/arXiv.1706.03762>)
- [19] Y. Yoon, H. Chung, and W. Lim, “Automatic modulation recognition using data-efficient image transformers (DeiT),” *J. KICS*, vol. 50, no. 2, pp. 245-252, 2025.
(<https://doi.org/10.7840/kics.2025.50.2.245>)

신 다 민 (Da-Min Shin)



2024년 2월 : 부산대학교 전자공학과 졸업
 2024년 3월~현재 : 부산대학교 전기전자공학과 석사과정
 <관심분야> 딥러닝 신호처리, 레이더 신호처리
 [ORCID:0009-0004-0489-5072]

전 민 욱 (Min-Wook Jeon)



2023년 2월 : 부산대학교 전자공학과 졸업
 2023년 3월~현재 : 부산대학교 전기전자공학과 석박통합과정
 <관심분야> 전자전·레이더 신호처리, 딥러닝
 [ORCID:0009-0001-7005-9062]

김 형 남 (Hyoung-Nam Kim)



1993년 : 포항공과대학교 전자전기공학과 학사 졸업
 1995년 : 포항공과대학교 전자전기공학과 석사 졸업
 2000년 : 포항공과대학교 전자전기공학과 박사 졸업.
 2000년 : 포항공과대학교 전자컴퓨터공학부 박사후 연구원
 2000년~2003년 : 한국전자통신연구원 무선방송연구소 선임연구원
 2003년~2007년 : 부산대학교 전자전기통신공학부 조교수
 2007년~2012년 : 부산대학교 전자전기통신공학부 부교수
 2009년~2010년 : Johns Hopkins Univ. Visiting Scholar
 2015년~2016년 : Univ. of Southampton Visiting Professor
 2012년~현재 : 부산대학교 전기전자공학부 전자공학전공 교수
 <관심분야> 통신신호처리, 레이더 및 소나 신호처리, 머신러닝, 생체 신호처리
 [ORCID:0000-0003-3841-448X]