

심층학습을 적용한 자동 변조 분류기의 입력 형태에 따른 성능 비교

Performance Comparison of Deep Learning Based Automatic Modulation Classifiers According to Input Types

강 솔 송 · 서 보 석

Solsong Kang · Bo-Seok Seo

요 약

이 논문에서는 수신 신호의 자동 변조 분류를 위해 다양한 심층학습 모델을 적용하고 입력 형태에 따른 분류기의 성능을 비교하고자 한다. 성능을 비교하기 위해 아날로그와 디지털 변조 신호 26종으로 구성된 데이터셋을 생성한다. 분류기 입력 형태로는 2차원 벡터의 신호열(sequence)로 구성된 시간 영역의 실수부와 허수부, 진폭과 위상, 주파수 영역의 실수부와 허수부, 크기와 위상, 스펙트로그램(spectrogram) 등 5가지를 고려한다. 분류기로는 CNN(convolutional neural network), LSTM(long short-term memory), ResNet(residual network), 트랜스포머(transformer)와 CNN과 LSTM을 결합한 CNN-LSTM 등 5개의 기본적인 심층학습 모델을 고려하고 SNR(signal-to-noise power ratio)에 따라 분류 정확도를 비교·분석한다. 모의실험 결과 LSTM 모델에서 입력형태가 시간 영역 실수부와 허수부인 경우 다른 모델 또는 다른 입력 형태에 비해 큰 차이로 가장 높은 분류 정확도를 나타낸다. 그 외 모델에서는 시간 영역 입력 형태가 주파수 영역에 비해서는 높지만, 시간 영역 두 입력 형태에 대해서는 큰 차이를 보이지 않는다. 이 결과를 통해 모델에 따라 또 입력 형태에 따라 심층학습 기반 변조 분류기의 성능에 큰 영향을 미치는 것을 알 수 있다.

Abstract

In this paper, we use various deep learning models for automatic modulation classification (AMC) of received signals and compare the classification performance according to different input types. To enable a fair comparison, a dataset consisting of 26 analog and digital modulation schemes is generated. Five input types are considered: time-domain inphase and quadrature components, time-domain magnitude and phase, frequency-domain real and imaginary components, frequency-domain magnitude and phase, and time-frequency spectrogram. The classification performance is evaluated using five representative deep learning models, including convolutional neural network (CNN), long short-term memory (LSTM), residual network (ResNet), Transformer, and a hybrid CNN-LSTM model, under various signal-to-noise ratio (SNR) conditions. Simulation results show that the LSTM model with time-domain inphase and quadrature inputs achieves the best classification accuracy with a significant margin compared to the other models and input formats. For the remaining models, the time-domain inputs generally outperform the frequency-domain inputs, while no significant performance difference is observed between the two time-domain types. These results demonstrate that both the choice of deep learning model and input type have a substantial impact on the performance of deep learning-based modulation classifiers.

Key words: AMC, Deep Learning-Based Classification, Modulation Signal Dataset, CNN, LSTM

「이 논문은 2025 GVC30 프로젝트의 지원을 받아 수행된 결과임.」

충북대학교 전자공학과(Department of Electronics Engineering, Chungbuk National University)

· Manuscript received December 19, 2025 ; Revised December 27, 2025 ; Accepted February 3, 2026. (ID No. 20251219-156)

· Corresponding Author: Bo-Seok Seo (e-mail: boseok@cbnu.ac.kr)

I. 서론

무선 통신 환경에서 수신 신호의 변조 방식을 자동으로 판별하는 자동 변조 분류(AMC, automatic modulation classification)는 전파 감시, 간섭 탐지, 스펙트럼 모니터링 등 군사용뿐만 아니라, 인지 무선통신(cognitive radio), 적응 통신 등 민간 통신 시스템에도 적용하기 위해 많이 연구되고 있는 기술이다^[1].

AMC 방법으로 기존에는 주로 신호의 고차 통계적 특성(high-order statistics), 고차 모멘트(high-order moment), 순환 정상성(cyclostationarity) 등의 신호 특징을 기반(feature-based)으로 하는 것이 대부분이었다^[2]. 그러나 이러한 방식은 특정 환경에 최적화되어 SNR(signal-to-noise power ratio)이 낮은 경우 성능이 급격하게 저하되는 한계를 가진다. 이로 인해 최근에는 데이터 기반의 심층학습(deep learning)을 적용한 AMC 방법이 많이 연구되고 있다^[3].

심층학습을 기반으로 하는 AMC는 CNN(convolutional neural network), LSTM(long short-term memory), ResNet(residual network), 트랜스포머(transformer) 등 다양한 신경망 구조가 제안되었다^[4-7]. 이들 대부분의 분류기는 변조 신호의 시간 영역 I/Q(inphase/quadrature phase) 또는 진폭/위상(magnitude/phase)을 입력으로 사용한다. 그러나 동일한 신호라도 입력 형태에 따라 신경망의 성능에 어떤 영향을 미치는가에 대해서는 아직 체계적인 분석이 충분히 이루어지지 않았다. 변조 신호는 시간 영역의 I/Q 또는 진폭/위상 형태뿐만 아니라, FFT(fast Fourier transform)한 후 주파수 영역의 실수부/허수부, 크기/위상 스펙트럼, 시간-주파수 기반의 스펙트로그램(spectrogram) 등 다양한 형태로 표현할 수 있다. 이러한 입력 형태는 신경망이 추출할 수 있는 특징 공간을 변화시키기 때문에 모델별 특성과 결합되어 분류 성능이 달라질 수 있다. 즉 신경망 구조뿐만 아니라 입력 형태도 AMC의 성능을 좌우하는 중요한 요인이 될 수 있다.

이 연구에서는 CNN, LSTM, CNN-LSTM, ResNet, 트랜스포머 등 5가지 신경망 구조에 대해 다섯 가지의 입력 형태, 즉 (1) 시간 영역 실수부/허수부(Tiq로 표기함), (2) 시간 영역 진폭/위상(Tmp), (3) 주파수 영역 실수부/허수부(Fri), (4) 주파수 영역 크기/위상(Fmp), (5) 스펙트로그램(Fsg) 등의 입력을 적용하여 모델과 입력 형태에 따른 성능을 비교한다. 성능을 비교하기 위해 아날로그 및 디지털 변조를 모두 포함하는 26개 변조 신호의 생성기를 직접 개발하여 다양한 환경에서 대규모 학습용 데이터셋을 생성한다. 여기에는 AWGN(additive white gaussian noise), 다중경로 페이딩, 대역제한 필터의 롤오프율(roll-off factor) 등의 채널 환경을 반영한다.

II. 신호 모델

2-1 신호 생성

이 논문에서 훈련과 테스트를 위해 사용하는 변조 신호는 전파 감시 환경에서 미지 신호를 기저대역으로 천이한 후 아날로그-디지털 변환하여 수신하는 상황을 가정하여 생성한다. 신호 생성기는 Matlab을 사용하여 작성하였으며, 그림 1과 같이 송신기, 채널, 수신기로 구성하고 생성한 데이터는 HDF5 형식으로 저장한다. 송신기는 변조 방식에 따라 기저대역 등가신호 또는 통과대역 변조 신호를 생성한다. 아날로그 변조인 경우 메시지 신호는 Youtube에서 취득한 음성 및 음악 신호이며, 디지털 변조인 경우 불규칙하게 생성한 비트열을 사용한다. 채널에는 AWGN과 다중경로, 도플러(Doppler) 주파수, 비선형 증폭기 등의 채널 영향과, 대역제한 필터의 롤오프율, 주파수 및 위상 오프셋 등 송수신기의 영향을 반영한다. 수신기에서는 미지의 신호를 기저대역으로 옮겨 수신하는 상황을 가정하여 대역의 중심이 0 Hz가 되도록 신호 대역을 천이하고, 과표본화율(over-sampling ratio)에 맞춰 재표본화(resampling)를 수행한다. 생성한 데이터는 변조 방식별 파일로 저장하여 새로운 변조 방식의 추가나 일부 파일의 수정이 용이하다.

생성한 8개 아날로그 변조 신호와 18개 디지털 변조 신호의 종류는 표 1과 같다. 대부분의 신호는 실제 많이 사용되고 있는 시스템의 규격을 기반으로 생성하며, 디지털

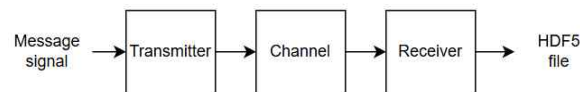


그림 1. 신호 생성기의 구조
Fig. 1. Structure of the signal generator.

표 1. 생성한 변조 신호

Table 1. Generated modulation signals.

Analog modulations	AM	DSB-WC, DSB-SC, LSB-WC, LSB-SC, USB-WC, USB-SC
	FM	WBFM, NBFM
Digital modulations	ASK	2ASK, 8ASK
	APSK	16APSK, 32APSK, 64APSK, 128APSK
	PSK	BPSK, QPSK, OQPSK, 8PSK
	QAM	16QAM, 32QAM, 64QAM, 128QAM, 256QAM
	FSK	GMSK, 2FSK, 4FSK

변조 신호 중에서 APSK(amplitude and phase shift keying)와 8PSK(phase shift keying)는 주로 위성 통신에서 사용되므로 TWTA(travelling wave tube amplifier) 비선형 증폭기 모델^[8]을 적용한다.

다중경로 채널은 Watterson, Rayleigh, Rician 모델을 사용한다^{[9]~[11]}. Watterson 모델은 아날로그 변조인 AM(amplitude modulation)과 FM(frequency modulation)에 적용하고, Rician 모델은 LOS(line-of-sight) 성분이 존재하는 환경을 고려해 위성통신 방식을 모의한 APSK와 8PSK에 적용한다. 그 외 변조 방식은 일반적인 비가시선(non-line-of-sight) 채널을 가정하는 Rayleigh 모델을 적용한다.

모든 변조 신호는 재표본화를 통해 표현되는 전체 대역의 폭과 신호 대역폭 비율인 과표본화율이 일정하도록 맞춘다. 생성된 신호는 I 및 Q 신호에 대해 길이 1024인 2x1,024 행렬로 HDF5 형식의 파일에 저장한다. 이때 채널 특성은 각 신호열마다 불규칙하게 선택된다. 각 변조 신호 파일은 -20~30 dB SNR에 대해 2 dB 간격으로, 매 SNR에 대해 4,096개, 총 106,496개의 신호열로 구성된다.

2-2 분류기 입력 형태

기저대역 등가 복소신호열은 식 (1)과 같이 나타낼 수 있다.

$$r[n] = [m_I(nT_s) - jm_Q(nT_s)] \exp\{j2\pi\Delta f nT_s + \phi(nT_s) + \phi_0\} + A_c \exp\{j2\pi\Delta f nT_s + \phi_0\} + (\eta_I(nT_s) + j\eta_Q(nT_s)) \quad (1)$$

여기서 m_I, m_Q 는 각각 메시지 신호의 I 및 Q 성분을 나

타내고, T_s 는 표본 간격, Δf 는 주파수 오프셋, ϕ 는 시변 위상, ϕ_0 는 초기 위상을 나타낸다. 또 A_c 는 반송파의 진폭을 나타내고(반송파 억압 변조인 경우는 $A_c=0$), $n_I + jn_Q$ 는 복소 가우스 잡음(gaussian noise)을 나타낸다. 주파수 오프셋과 위상은 매 신호열마다 불규칙적으로 달라지지만 사용한 신호열 길이 1,024는 비교적 짧은 시간에 해당하므로 하나의 신호열에 대해서는 고정한다. 주파수 오프셋 또는 스펙트럼의 중심주파수 편이는 신호 생성 시에는 0으로 놓고 학습 시에 표본화 주파수를 기준으로 반영한다. 학습에 사용되는 신호열은 크기와 표본화 주파수를 1로 정규화하면 식 (2)와 같이 표현할 수 있다.

$$r_n[n] = r[n] \exp\{j2\pi f_{dev} n\} \quad (2)$$

여기서 f_{dev} 는 표본율에 대해 정규화한 스펙트럼 편이를 나타내며, 최대 주파수 오프셋을 $f_{dev,max}$ 라 할 때, 학습열은 $[-f_{dev,max} \ f_{dev,max}]$ 사이에서 균일한 분포를 가진 불규칙 변수이다. 이것은 표본율의 $\pm f_{dev,max}$ 범위까지 스펙트럼 편이가 있음을 의미한다.

이 논문에서는 시간 영역 신호의 T_{iq} 와 T_{mp} , 주파수 영역의 F_{ri} 와 F_{mp} , 그리고 시간-주파수 영역의 F_{sg} 등 5가지 형태의 분류기 입력(신호는 전처리 과정에서 형태를 바꾸고 신경망에 입력되지만, 편의상 신경망 입력을 분류기 입력이라 부른다)을 고려한다. 시간 영역 T_{iq} 신호는 정규화시킨 신호열 (2)의 실수부(I)와 허수부(Q)로서 식 (3) 및 식 (4)와 같이 표현된다.

$$x_I[n] = \text{real}\{r_n[n]\} \quad (3)$$

$$x_Q[n] = \text{imag}\{r_n[n]\} \quad (4)$$

신호열의 길이를 N 이라 할 때, 이 경우 분류기 입력은 $2 \times N$ 행렬이 된다. 시간 영역 신호의 진폭과 위상을 나타내는 T_{mp} 신호는 식 (5) 및 식 (6)과 같이 나타낼 수 있다.

$$x_M[n] = \sqrt{x_I^2[n] + x_Q^2[n]} \quad (5)$$

$$x_P[n] = \frac{1}{\pi} \text{atan} \left(\frac{x_Q[n]}{x_I[n]} \right) \quad (6)$$

주파수 영역 신호의 실수부와 허수부를 나타내는 F_{ri}

신호는 k 번째 주파수에 대해 식 (7) 및 식 (8)과 같이 나타낼 수 있다.

$$X_R[k] = \frac{\text{real}\{X[k]\}}{|X|_{\max}} \quad (7)$$

$$X_I[k] = \frac{\text{imag}\{X[k]\}}{|X|_{\max}} \quad (8)$$

여기서 $X[k] = FFT\{x_n[n]\}$ 로 기저대역 신호열을 FFT 한 것이고, $|X|_{\max}$ 는 전체 주파수 성분 중에서 최대 진폭을 나타낸다. 주파수 영역 스펙트럼의 크기와 위상을 나타내는 Fmp 신호는 식 (9) 및 식 (10)과 같이 나타낼 수 있다.

$$X_M[k] = \frac{|X[k]|}{|X|_{\max}} \quad (9)$$

$$X_P[k] = \frac{1}{\pi} \text{atan}\left(\frac{X_R[k]}{X_I[k]}\right) \quad (10)$$

스펙트로그램은 길이 N 인 신호를 길이 M 인 B 개의 블록으로 나눈 다음, 각 블록을 M 점 FFT하여 구한다.

$$X_b[k] = \frac{1}{|X|_{\max}} |FFT\{x_b[n]\}| \quad (11)$$

여기서 $X_b[k]$ 는 b 번째 블록을 FFT한 것의 k 번째 성분 값으로, Fsg 신호는 $M \times B$ 2차원 행렬이 된다. 스펙트로그램은 각 블록이 서로 겹치게 구하는 방법도 있지만 이 논문에서는 $M \times B = N$ 을 만족하도록 겹치지 않게 블록을 나누었다^[12].

분류기의 입력 형태는 신경망에 입력되기 전에 전처리 과정에서 변환하며, 그 결과 Fsg 신호열은 0과 1 사이로, 나머지는 -1과 1 사이로 정규화된다.

III. 분류기 구조

입력 형태가 분류 성능에 미치는 영향을 분석하기 위해, AMC 분야에서 널리 사용되는 대표적인 심층 학습 구조인 CNN, LSTM, ResNet, 트랜스포머 4가지 모델과 CNN, LSTM을 결합한 하이브리드 모델(CNN-LSTM)을 비교 대상으로 선정하였다. 이 논문의 목적은 모델 자체

의 성능을 최적화하기보다, 모델의 구조가 동일한 조건에서 입력 형태에 따라 어떤 성능 차이를 보이는지 비교하고자 하는 것이다. 따라서 대표적인 모델의 기본적인 구조를 가정하였다.

CNN은 국소 영역의 패턴을 효과적으로 추출하고 학습할 수 있으며, LSTM은 신호열의 장기 의존성을 활용하는 구조이다. CNN-LSTM 모델은 CNN의 국소 특징 추출 능력과 LSTM의 시간적 상관관계 학습 능력을 결합한 구조로 두 모델의 장점을 동시에 활용할 수 있다는 특징이 있다. ResNet은 깊은 합성곱(convolution) 구조를 통해 복잡한 변조 패턴을 표현할 수 있으며, 고차 변조 방식의 세밀한 특징을 학습하는 데 유리한 특성을 가진다^[13]. 트랜스포머는 self-attention 기반 구조로 전체 신호열을 동시에 고려할 수 있으나, 신호열 길이나 잡음 특성 등에 따라 기존 모델과는 다른 학습 양상이 나타날 수 있다.

분류기에 사용한 심층학습 신경망 모델 및 구조는 표 2와 같다. 각 신경망 구조는 참고문헌의 구조를 기반으로 이 논문에서 적용하는 학습신호에 대해 다수의 시행착오를 바탕으로 최고의 분류 정확도를 나타내는 하이퍼파라미터(hyperparameter)를 선택하였다. 그 결과로 학습 파라미터의 개수는 모델별로 차이가 있음을 볼 수 있다. 표에서 CNN 기반 모델의 입력(input) 차원은 영상을 기반으로 하여 (높이×가로폭×채널 수)를 의미한다. 또 c 는 채널 수, k 는 필터 길이, FC는 완전연결층(fully connected layer)을 의미하고 괄호 안 숫자는 출력 개수를 의미한다. LSTM의 h 는 은닉상태 벡터의 차원을 나타내고, Transformer의 d_k 는 head의 차원을 의미한다.

IV. 모의실험 결과

모의실험에서는 생성한 26가지 변조 신호의 입력 형태에 따른 모델별 성능 차이를 확인하여 각 입력 형태가 학습 및 분류 성능에 미치는 영향을 분석한다. 모의실험에 사용한 데이터셋 구성 및 모의실험 환경은 표 3과 같다. 5개 신경망 모델의 학습 시 최적화 기법은 Adam을 사용하고, 에폭(epoch) 수는 20, 배치(batch) 크기는 1,000으로 설정하였다. 학습률은 초기 0.001에서 최종적으로 약 0.0001이 되도록 줄여간다. 모의실험은 Matlab R2025a를

표 2. 분류기에 사용한 심층학습 신경망 모델
Table 2. Deep learning neural network models for classifiers.

Model	Structure	Number of trainable parameters
CNN ^[2]	Input (1×1024×2), Conv1 (c:128, k:3×3), Conv2 (c:128, k:3×3), Conv3 (c:64, k:3×3), Conv4 (c:64, k:3×3), Flatten, FC1 (260), FC2 (26)	1.3×10 ⁶
CNN-LSTM ^{[5],[6],[14]}	Input (1×1024×2), Conv1 (c:128, k:3×3), Conv2 (c:128, k:3×3), Flatten, LSTM1 (h:256), LSTM2 (h:256), LSTM3 (h:12), Flatten, FC1 (260), FC2 (26)	3.5×10 ⁷
ResNet ^{[4],[15]}	Input (1×1024×2), Conv1 (c:128, k:1×1), [ResBlock (c:128, k:1×3)] ×4, Flatten, FC1 (260), FC2 (26)	2.9×10 ⁶
LSTM ^{[16],[17]}	Input (dim=8), LSTM1 (h:256), LSTM2 (h:25), LSTM3 (h:128), Flatten, FC1 (260), FC2 (26)	1.0×10 ⁶
Transformer ^[7]	Input (dim=8), PositionalEmbedding(Seq_len=1024), [Heads (4, d_k:128), FF (128), FC (16)]×2, FC1 (260), FC2(26)	3.1×10 ⁴

표 3. 데이터셋 구성 및 모의실험 환경
Table 3. Dataset configuration and simulation environment.

Parameter	Value
Number of training sequences	300,000
Complex sequence length	1,024
Samples per symbol	4
Number of modulation classes	26
Training:validation:test	7:1:2
SNR (dB)	-20:2:30
Number of blocks in Fsg input	4

사용하여 수행하였다.

그림 2는 분류기 모델별 입력 형태에 따른 분류 정확도를 SNR에 따라 나타낸 것이다. 트랜스포머를 제외하고 나머지 모델에서는 모두 4~8 dB를 기준으로 그보다 높은 SNR 영역에서는 시간 영역 입력 형태인 T_{iq} , T_{mp} 가 주파수 영역 입력 형태인 F_{ri} , F_{mp} , F_{sg} 에 비해 상대적으로 우수하고, 낮은 SNR 영역에서는 반대의 경향을 나타낸다. 특히 LSTM의 경우 높은 SNR 영역에서 그 성능 차이가 매우 크다. 한편 다른 모델은 모두 T_{mp} - T_{iq} - F_{sg} - F_{mp} - F_{ri} 입력 순으로 성능을 나타내지만, LSTM은 그 차이는 크지는 않지만 시간 영역에서 T_{iq} - T_{mp} 순으로 뒤바뀐 점이다. 이것은 신호열의 시간적인 상관성을 이용하는 LSTM에서는 T_{iq} 입력이 더 효과적임을 나타낸다. LSTM이 시간적 상관성이 있는 입력에 효과적이라는 결과는, LSTM의 입력 차원을 정하는 과정에서 과표본화율 4(표의 차원에서는 I, Q를 합한 크기로 표현하므로 8) 부근에서 가장 높은 정확도를 나타냈는데, 이로부터도 그 타당성을 입증할 수 있다.

그림에서 보면 트랜스포머가 상대적으로 다른 모델에 비해 성능이 낮은 것은 알 수 있다. 트랜스포머는 입력 신호열 전체 구간 내에서 상호 간에 상관성이 있을 때 효과적이다. 그러나 변조 신호, 특히 디지털 변조 신호는 한 심볼 내에서는 상관성이 크지만 심볼 간에는 불규칙(random)하므로 변조 신호를 구별하는 목적으로는 트랜스포머가 덜 효과적인 것으로 판단된다.

특이한 것은 전체 SNR 영역에서 주파수 영역의 진폭 차원만을 이용하는 F_{sg} 입력일 때 모델 모두 다른 주파수 영역 입력 형태인 F_{ri} , F_{mp} 에 비해 비슷하거나(낮은 SNR 영역에서) 훨씬 높은 성능(높은 SNR 영역에서)을 나타내는 점이다. 이 이유는 두 가지 관점에서 해석할 수 있다. 먼저 스펙트럼의 진폭에 대부분의 정보가 있으며, 오히려 위상 정보는 성능을 저하시키는 것으로 작용한다고 볼 수 있다. 두 번째는 신호 발생 시 시변 채널을 적용하였는데, F_{sg} 입력은 시변 특성을 반영하므로 그 특성이 성능에 큰 영향을 준 것으로도 해석할 수 있다.

한편 모델 모두 높은 SNR에서 시간 영역 입력 형태가 상대적으로 우수한 이유는 디지털 변조 신호(전체 26개 변조 신호 중에서 18개가 디지털 변조 신호임)의 경우 변

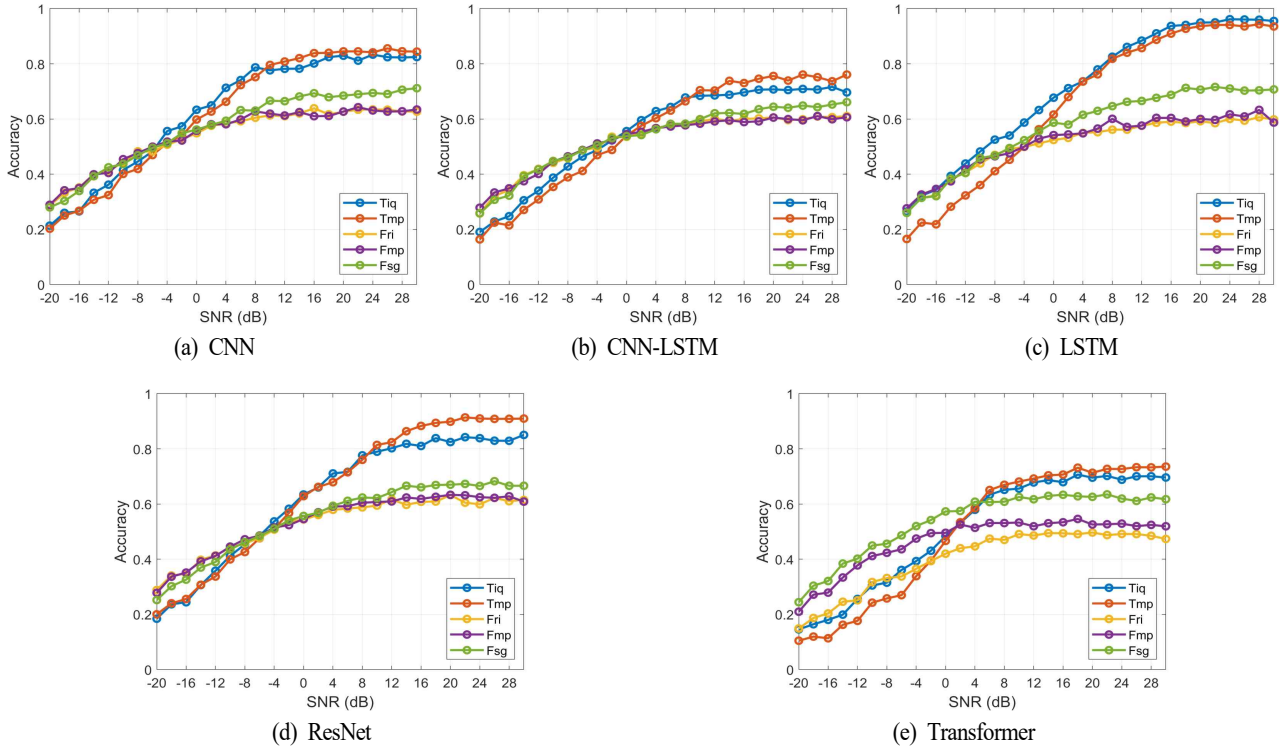


그림 2. 신경망의 입력 형태에 따른 분류기의 분류 정확도

Fig. 2. Classification accuracy of the classifiers according to the input format of the neural network.

조 방식마다 파형의 차이가 크지만, 주파수 영역 스펙트럼은 대역 내에서 거의 평탄하여 유사한 형태이기 때문에 판단된다. 또 낮은 SNR에서는 오히려 주파수 영역 입력 형태가 상대적으로 우수한 이유는 잡음의 주파수 특성이 백색(white)이므로 일반 정보신호와 구별되기 때문으로 판단된다.

그림 3은 아날로그 변조 신호와 디지털 변조 신호를 분리하여 분류 정확도를 나타낸 것이다. 분류기 모델 모두 5개 입력 형태 모두에서 아날로그 변조 신호가 디지털 변조 신호에 비해 전반적으로 높은 분류 정확도를 나타내며, 낮은 SNR 영역에서는 그 차이가 더욱 두드러진다. 이는 디지털 변조 방식의 경우 고차 변조 방식이 다수 포함되어 있어 상대적으로 신호 분류가 어렵기 때문으로 해석할 수 있다.

디지털 변조 방식에서는 시간 영역 파형에서 변조 방식 간 특징 차이가 뚜렷하게 나타난다. 따라서 높은 SNR 영역에서 시간 영역 입력(Tiq, Tmp)인 경우의 분류 정확

도가 상대적으로 높게 나타나며, 특히 LSTM은 차이가 가장 크다.

아날로그 변조 신호는 CNN이 포함된 모델의 경우 대부분의 SNR 영역에서 주파수 영역 입력 형태가 시간 영역 입력 형태와 비슷하거나 더 나은 성능을 나타낸다. 이것은 시간 영역 신호 패턴 사이에는 서로 불규칙(random)하여 상관성이 크지 않은 반면 스펙트럼은 변조 방식별로 다르게 나타나 구별되기 때문이다. 이러한 성능 차이는 저자의 이전 연구 참고문헌 [18]과 동일한 결과를 나타내고 있다. LSTM은 낮은 SNR 영역에서 Tmp 입력인 경우 다른 입력에 비해 성능이 열악하다. 또 디지털 변조 신호에서도 유사한 결과를 나타내는데 이것은 시간 영역 위상 특성이 잡음에 의해 크게 영향을 받기 때문인 것으로 판단된다. 이런 결과는 상관성을 이용하는 트랜스포머에서도 유사하게 나타난다. 한편 트랜스포머는 SNR이 증가해도 Fsg 입력을 제외한 입력 형태 모두에서 분류 정확도가 비교적 큰 것을 볼 수 있다. 이것은 아날로그 변조

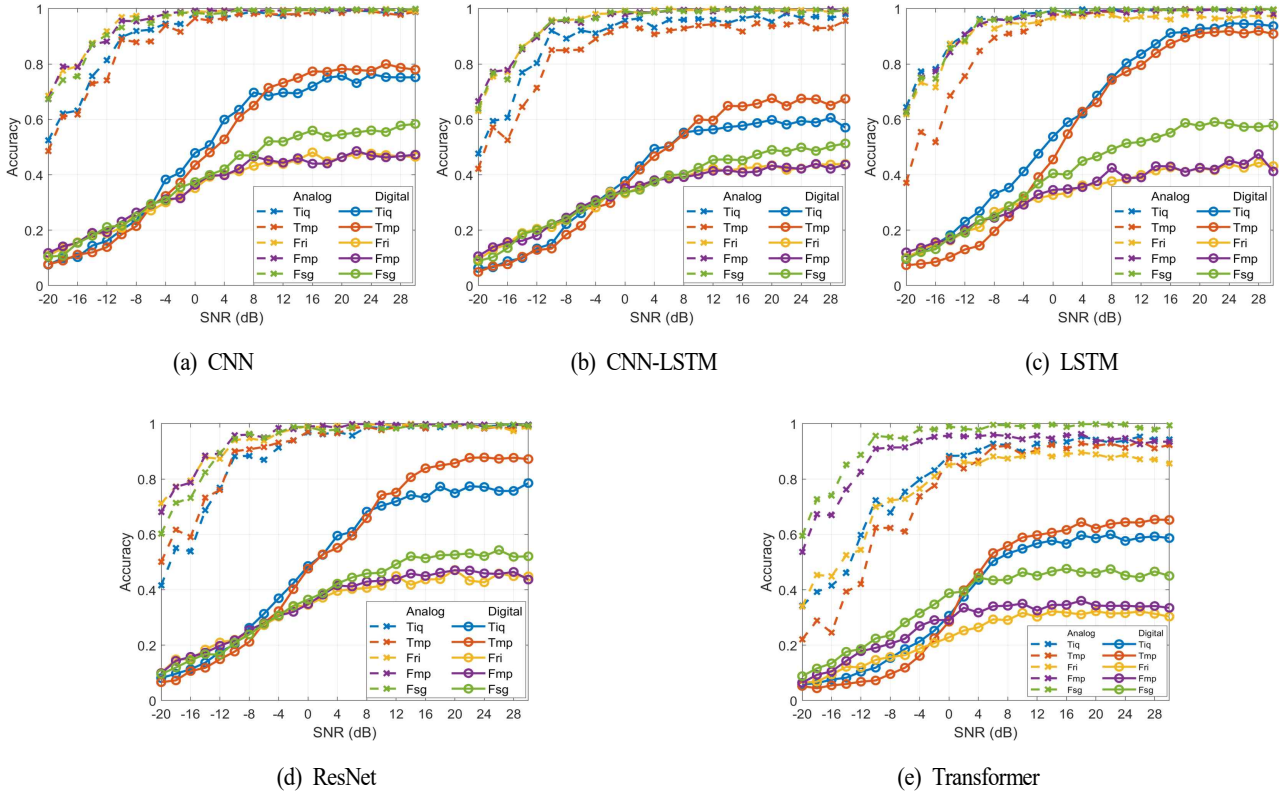


그림 3. 입력 형태에 따른 아날로그 및 디지털 변조 신호의 분류 정확도

Fig. 3. Classification accuracy of the classifiers for analog and digital modulation signals according to the input format.

신호의 경우 입력 신호 전체 구간에서 상호간의 상관성은 크지 않기 때문인 것으로 판단된다.

그림 4는 5가지 입력 형태에 대해 심층학습 모델별 분류 정확도를 비교한 결과를 나타낸다. 10 dB 이상의 높은 SNR 구간에서는 시간 영역 입력 형태 T_{iq} , T_{mp} 의 경우 LSTM - ResNet - CNN - CNN-LSTM - 트랜스포머 순서의 분류율을 나타내며, 특히 T_{iq} 입력에서는 LSTM이 다른 모델 대비 뚜렷한 성능 우위를 나타낸다. 주파수 영역 입력 형태 F_{ri} , F_{mp} 인 경우는 트랜스포머를 제외하고 모델에 따라 성능 차이가 크지 않다. 반면에 트랜스포머는 대부분의 입력 형태에서 다른 모델에 비해 낮은 성능을 보이고 있으며, 특히 F_{mp} 입력인 경우 그 차이가 크다. 이것은 트랜스포머가 입력 시퀀스 상호 간 또는 입출력 시퀀스 간의 상관성을 이용한다는 점에서 변조 신호에는 크게 효과적이지 않음을 나타낸다.

그림 5는 심층학습 모델별로 4~8 dB 이상에서 가장

우수한 성능을 보인 입력 형태에 대해 SNR에 따른 분류 정확도를 비교한 결과이다. LSTM을 제외하고는 모두 T_{mp} 입력일 때 성능이 가장 높다. LSTM 모델의 경우 T_{iq} 입력을 사용했을 때 전체 SNR 구간에서 다른 모델에 비해 뚜렷하게 높은 분류 정확도를 나타낸다. 전체적으로 LSTM - ResNet - CNN - CNN-LSTM - 트랜스포머 순의 성능을 보이며, 특히 트랜스포머는 4 dB 이하의 SNR에서 다른 모델에 비해 훨씬 낮은 성능을 나타낸다.

V. 결 론

이 논문에서는 26종의 변조 신호를 생성하고, 이를 사용하여 입력 형태가 신경망 기반 자동 변조 분류기의 성능에 미치는 영향을 비교 분석하였다. 입력 형태는 시간 영역 실수부/허수부(T_{iq}), 시간 영역 진폭/위상(T_{mp}), 주파수 영역 실수부/허수부(F_{ri}), 주파수 영역 크기/위상

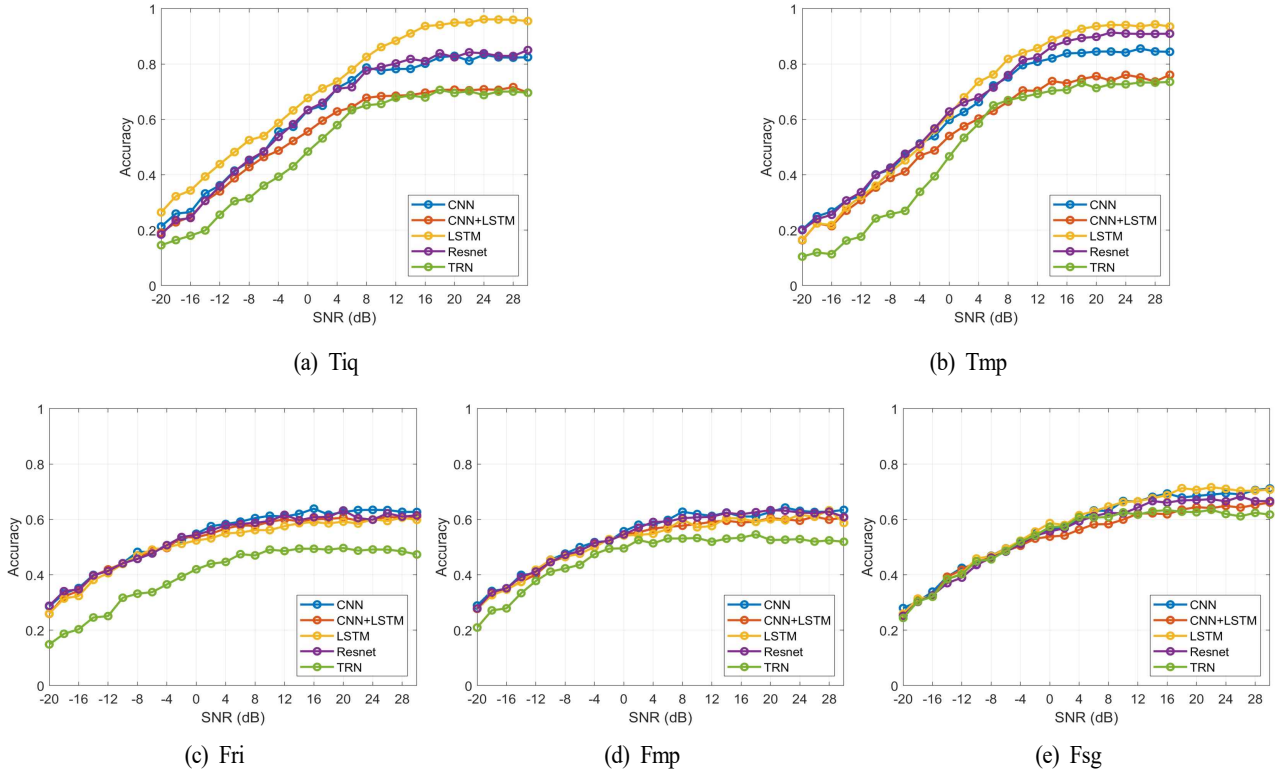


그림 4. 특정 입력 형태에 대한 여러 분류기 모델의 분류 정확도
 Fig. 4. Classification accuracy of various classifier models for a specific input format.

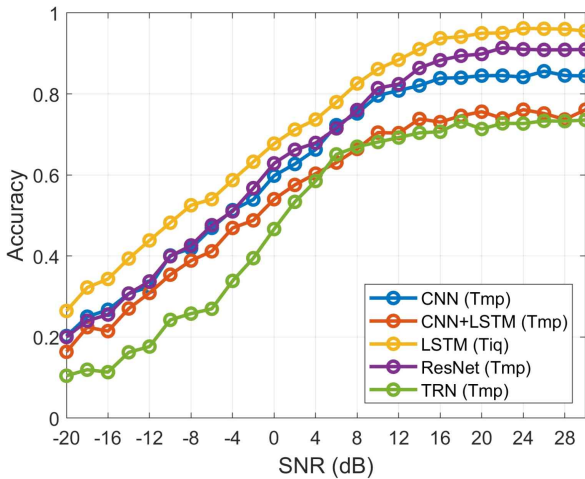


그림 5. 모델별 최고 성능의 입력 형태에 대한 분류 정확도
 Fig. 5. Classification accuracy for the best performance input type of each model.

(Fmp), 스펙트로그램(Fsg) 등 5가지를 사용하였고, 신경망은 CNN, LSTM, CNN-LSTM, ResNet, 트랜스포머 등 5개 모델을 고려하였다.

모의실험 결과 실제 통신이 이루어지는 4~5 dB 이상의 SNR 영역에서 모든 모델에 대해 시간 영역 입력 형태가 주파수 입력 형태보다 높은 정확도를 나타내었다. 이 경우 LSTM 모델에서는 Tmq 입력이 가장 우수한 성능을 나타냈으며, 나머지 모델에서는 Tmp 입력이 근소한 차이로 나은 성능을 나타냈다.

아날로그 변조 신호와 디지털 변조 신호에 대한 비교에서는 모델 모두 전체 SNR 영역에서 아날로그 변조 신호가 입력 형태와 무관하게 디지털 변조 신호보다 높은 분류 정확도를 나타냈으며, 높은 SNR에서는 입력 형태에 따라 성능 차이가 디지털 변조에 비해 상대적으로 적었다.

입력 형태별로 모델 간 성능을 비교한 결과, 시간 영역 Tmq, Tmp 입력의 경우 LSTM - ResNet - CNN - CNN-LSTM

- 트랜스포머 순으로 비교적 뚜렷한 성능 차이를 보였다. 그러나 주파수 영역 입력에서는 트랜스포머를 제외하고는 모델들 간의 성능 차이가 크지 않았다.

전체적으로는 LSTM-Tiq 조합이 전체 SNR 구간에서 가장 우수한 성능을 나타냈으며, 트랜스포머는 전반적으로 가장 낮은 분류 정확도를 보였다.

이 연구의 결과는 AMC 설계 시 신경망 구조뿐만 아니라 입력 형태도 성능에 크게 영향을 미친다는 것을 의미한다. 한편 모의 생성한 데이터셋을 사용하여 학습한 분류기의 한계성이 존재하므로 실측 데이터 적용성에 대해서는 향후 더 연구가 필요할 것으로 판단된다.

References

- [1] Z. Zhu, A. K. Nandi. *Automatic Modulation Classification: Principles, Algorithms and Applications*, Hoboken, NJ, John Wiley & Sons, 2015.
- [2] O. A. Dobre, A. Abdi, Y. Bar-Ness, and W. Su. "Survey of automatic modulation classification techniques: Classical approaches and new trends," *IET Communications*, vol. 1, no. 2, pp. 137-156, Apr. 2007.
- [3] T. Huynh-The, Q. V. Pham, T. V. Nguyen, T. T. Nguyen, R. Ruby, and M. Zeng, et al., "Automatic modulation classification: A deep architecture survey," *IEEE Access*, vol. 9, pp. 142950-142971, Oct. 2021.
- [4] T. J. O'Shea, T. Roy, and T. Charles Clancy, "Over-the-air deep learning based radio signal classification," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 1, pp. 168-179, Feb. 2018.
- [5] Y. Wu, X. Li, and J. Fang, "A deep learning approach for modulation recognition via exploiting temporal correlations," in *2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Kalamata, Jun. 2018.
- [6] N. E. West, T. O'Shea, "Deep architectures for modulation recognition," in *2017 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*, Baltimore, MD, Mar. 2017.
- [7] S. Ying, S. Huang, S. Chang, Z. Yang, Z. Feng, and N. Guo, "A convolutional and transformer based deep neural network for automatic modulation classification," *China Communications*, vol. 20, no. 5, pp. 135-147, May 2023.
- [8] A. A. M. Saleh, "Frequency-independent and frequency-dependent nonlinear models of TWT amplifiers," *IEEE Transactions on Communications*, vol. 29, no. 11, pp. 1715-1720, Nov. 1981.
- [9] C. Watterson, J. Juroshek, and W. Benserna, "Experimental confirmation of an HF channel model," *IEEE Transactions on Communication Technology*, vol. 18, no. 6, pp. 792-803, Dec. 1970.
- [10] J. G. Proakis, *Digital Communications*, 4th ed. New York, NY, McGraw-Hill, 2001.
- [11] E. C. Van Der Meulen, "Three-terminal communication channels," *Advances in Applied Probability*, vol. 3, no. 1, pp. 120-154, 1971.
- [12] A. V. Oppenheim, R. W. Schaffer, *Discrete-Time Signal Processing*, Englewood Cliffs, NJ, Prentice-Hall, 2009.
- [13] L. J. Ba, R. Caruana. "Do deep nets really need to be deep?" *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 27, pp. 2654-2662, 2014.
- [14] T. N. Sainath, O. Vinyals, A. Senior, and H. Sak, "Convolutional, long short-term memory, fully connected deep neural networks," in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, South Brisbane, QLD, Apr. 2015, pp. 4580-4584.
- [15] S. Ramjee, S. Ju, D. Yang, X. Liu, A. E. Gamal, and Y. C. Eldar, "Fast deep learning for automatic modulation classification," *arXiv preprint*, arXiv:1901.05850, Jan. 2019.
- [16] V. Clerico, J. González-López, G. Agam, and J. Grajal, "LSTM framework for classification of radar and communications signals," in *2023 IEEE Radar Conference (RadarConf23)*, San Antonio, TX, May 2023.
- [17] S. Rajendran, W. Meert, D. Giustiniano, V. Lenders, and S. Pollin, "Deep learning models for wireless signal classification with distributed low-cost spectrum sensors," *IEEE Transactions on Cognitive Communications and Networking*, vol. 4, no. 3, pp. 433-445, Sep. 2018.
- [18] B. S. Seo, S. Kang, and Y. G. Yi, "Performance comparison of CNN classifiers for analog modulated signals according to input types," *The Journal of Korean Institute of Electromagnetic Engineering and Science*, vol. 36, no. 2, pp. 212-220, Feb. 2025.

강 솔 송 [충북대학교/석사과정]

<https://orcid.org/0009-0003-3136-9476>



2021년 8월: 한밭대학교 전자제어공학과 (공학사)

2026년 2월: 충북대학교 전자공학과 석사과정

[주 관심분야] 통신신호처리, 변조인식, 신경망

서 보 석 [충북대학교/교수]

<https://orcid.org/0000-0002-1610-3667>



1987년 2월: 서울대학교 전자공학과 (공학사)

1989년 2월: 서울대학교 전자공학과 (공학석사)

1997년 2월: 서울대학교 전자공학과 (공학박사)

2004년 9월~현재: 충북대학교 전자공학과 교수

[주 관심분야] 통신신호처리, 레이더 신호처리, 기계학습