

# ViT, VLM 기반 LPI 레이더 변조 분류 기술 연구 동향

장민석, 남해운

한양대학교

areess306@hanyang.ac.kr, hnam@hanyang.ac.kr

## Research Trends on LPI Radar Modulation Recognition Using Vision Transformer and Vision Language Model

Minseok Kang, Haewoon Nam

Hanyang Univ.

### 요약

전자전 환경에서 LPI 레이더는 낮은 피크 전력과 펄스 내 변조 기반의 광대역 과형 설계로 ESM 수신기의 신호 탐지를 어렵게 만든다. 이러한 특성으로 인해 수신 신호의 변조 방식을 식별하여 레이더 신호의 유형을 분류하고 위협 수준을 판단하는 것이 중요하며, 이를 위해 딥러닝 기반의 변조 분류 기법들이 지속적으로 연구되어 왔다. 그러나 기존 CNN 기반의 접근 방식은 낮은 SNR 환경에서 배경 잡음에 취약하여 분류 성능이 급격히 저하되는 한계를 보인다. 이를 보완하기 위해 최근에는 Vision Transformer(ViT)와 Vision Language Model(VLM)을 활용한 기법들이 활발히 연구되고 있다.

본 논문에서는 LPI 레이더 변조 분류 분야에서 ViT, VLM을 활용한 최신 연구 동향을 분석하고, 향후 발전 방향을 제시하고자 한다.

### I. 서론

저페탐(Low Probability of Intercept, LPI) 레이더는 현대 전자전(Electronic Warfare, EW) 환경에서 핵심 기술로 평가받는다. 낮은 피크 전력과 펄스 내 변조(intra-pulse modulation)를 통한 광대역 과형 설계로 전자 지원(Electronic Support Measure, ESM) 수신기가 신호를 탐지하기 어렵게 만들기 때문이다.

이러한 특성으로 인해 ESM 수신기 관점에서는 단순 에너지 검출 방식만으로 LPI 레이더 신호를 탐지 및 식별하는 데 한계가 있으며, 수신 신호에 포함된 변조 특성을 분석하는 것이 중요하다. 특히 펄스 내 변조 방식을 정확히 파악할 수 있다면, 레이더 신호의 유형 식별뿐만 아니라 레이더 운용 방식 및 위협 수준을 추정하는 데 중요한 정보를 얻을 수 있다 [1].

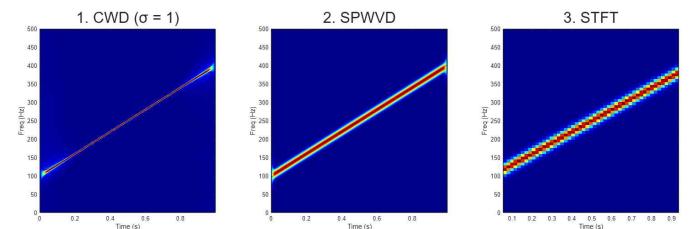
현재 레이더 신호 변조 분류를 위해 시간 - 주파수 이미지(Time - Frequency Image, TFI)를 입력으로 하는 딥러닝 기반 기법이 활발히 연구되고 있으며, 합성곱 신경망(Convolutional Neural Network, CNN)을 적용하는 방법이 주로 활용되어 왔다. 그러나 LPI 레이더 신호의 TFI는 일반적으로 이미지 중앙 영역에 신호의 주요 특징이 집중되는 특성을 보이는데 [2], 낮은 SNR(Signal-to-Noise Ratio) 환경에서 배경 전역의 잡음을 강해질 경우 CNN은 분류 성능이 현저히 저하된다.

이러한 문제를 해결하기 위해 자기주의(Self-Attention) 메커니즘을 활용한 Vision Transformer(ViT) 모델이 대안으로 제시되고 있으며, 최근에는 시각 정보와 텍스트를 통합하여 학습하는 멀티모달 기반의 Vision Language Model(VLM) 접근법 또한 주목받고 있다. 본 논문에서는 LPI 레이더 변조 분류 분야에서 ViT 및 VLM을 활용한 최신 연구 동향을 분석하고, 향후 발전 방향을 제시하고자 한다.

### II. 본론

#### 1. Time-Frequency Image

딥러닝 기반의 레이더 변조 분류 기법에서는 수신 신호를 시간 - 주파수 영역에서 표현한 시간 - 주파수 이미지(Time - Frequency Image, TFI)를 모델의 입력으로 사용한다. ViT, VLM 기반 연구에서는 CWD(Choi-Williams Distribution), SPWVD(Smoothed Pseudo-Wigner-Ville Distribution), STFT(Short-Time Fourier Transform)를 통해 생성된 TFI가 주로 활용되고 있으며, [그림 1]은 CWD, SPWVD, STFT 방식을 비교하기 위해 생성된 LFM(Linear Frequency Modulation) 변조 신호의 TFI이다.



[그림 1] LPI 레이더 LFM 신호의 CWD, SPWVD, STFT 이미지

CWD는 Wigner-Ville 분포(WVD)의 교차항(Cross-Term)을 효과적으로 억제하기 위해 지수함수 커널을 사용하는 시간-주파수 분석 기법으로, 다음과 같이 정의된다.

$$\int \int_{-\infty}^{\infty} \sqrt{\frac{\sigma}{4\pi\tau^2}} \exp\left(-\frac{\sigma(\mu-t)^2}{4\tau^2}\right) x(\mu + \frac{\tau}{2}) x^*(\mu - \frac{\tau}{2}) e^{-j2\pi f\tau} d\mu d\tau \quad (1)$$

여기서  $\sigma$ 는 교차항의 감쇠 정도를 조절하는 스케일 파라미터(Scale Parameter)이며,  $\sigma$  값이 커질수록 해상도는 좋아지지만 교차항 억제력이 감소하게 된다 [3]. CWD는 WVD 대비 교차항을 효과적으로 줄이면서도 높은 시간-주파수 해상도를 제공하나, 낮은 SNR 환경에서는 SPWVD에 비해 배경 잡음 억제 성능이 상대적으로 부족할 수 있다는 한계가 있다.

SPWVD는 시간 영역 윈도우 함수  $h(\tau)$ 와 주파수 영역 윈도우 함수  $g(u)$

를 동시에 적용하는 이중 윈도우 기반의 시간 - 주파수 분석 기법이다.

$$\int_{-\infty}^{\infty} h(\tau) \left[ \int_{-\infty}^{\infty} g(s-t) x(s + \frac{\tau}{2}) x^*(s - \frac{\tau}{2}) ds \right] e^{-j2\pi f\tau} d\tau \quad (2)$$

이러한 이중 윈도우 구조는 WVD의 높은 해상도를 최대한 유지하면서도 교차항으로 인한 거짓 신호를 최소화할 수 있어, 낮은 SNR 환경과 복잡한 다중 성분 신호 분류에 매우 강건한 특성을 보인다. 그러나 2차원 적분 연산에 따른 높은 계산 복잡도로 인해 비교 기법 중 처리 시간이 가장 길다는 단점이 존재한다.

STFT는 단일 윈도우 함수  $w(\tau-t)$ 를 슬라이딩하며 각 시간 위치에서 신호의 푸리에 변환을 수행하는 기법이다.

$$\int_{-\infty}^{\infty} x(t) w(t - \tau) e^{-j2\pi f\tau} d\tau \quad (3)$$

STFT는 계산 복잡도가 가장 낮고 처리 속도가 매우 빠르다는 장점이 있어 실시간 레이더 신호 처리에 적합하다. 다만 윈도우 크기에 따라 시간 해상도와 주파수 해상도가 반비례하는 트레이드오프가 존재하며, 낮은 SNR 환경에서는 윈도우 밖의 강한 잡음이 신호를 마스킹하여 분류 성능이 급격히 저하되는 문제가 있다.

## 2. 연구 동향

기존 CNN 기반의 LPI 레이더 변조 분류 연구에서는 잡음이 지배적인 낮은 SNR 환경에서 분류 성능이 현저히 저하되는 한계가 있었다. 이를 극복하기 위해 [4]은 자기주의(Self-Attention) 메커니즘을 기반으로 하는 Vision Transformer(ViT)를 LPI 레이더 변조 분류에 처음 도입하였다. LPI 레이더 신호의 TFI는 신호의 특징이 이미지 중앙에 집중되어 있는데, ViT는 중요한 신호 영역의 패치에는 높은 가중치를, 무관한 잡음 성분에는 낮은 가중치를 할당한다. 이는 CNN의 국소 필터와 달리 이미지의 모든 영역을 동시에 고려하면서 신호와 잡음을 효과적으로 구분할 수 있음을 의미한다. 해당 연구에서는 LFM, Costas, BPSK, Frank, P1-P4, T1-T4 총 12가지 변조 신호를 대상으로 실험을 진행하였다. 입력 TFI로는 CWD를 사용하였으며, -10dB의 낮은 SNR 환경에서 CNN 대비 12.8% 향상된 분류 정확도를 달성하였다.

[5] 논문은 극도로 낮은 SNR 환경에서 LPI 레이더 변조 신호를 분류하기 위해 VViT(Vicinity Vision Transformer) 모델을 적용하였다. VViT는 기존 ViT의 이차 계산 복잡도 문제를 개선하기 위해 Vicinity Attention 메커니즘을 도입하여 연산 복잡도를 선형 수준으로 낮추도록 설계된 모델이다. 해당 연구에서는 STFT를 통해 얻은 Phase Spectrum을 입력으로 사용하였으며, Barker, Frank, P1, P2, P3, P4 총 6가지 변조 신호를 극도로 낮은 SNR=-16 dB 환경에서 93%의 정확도로 분류하는 데 성공하였다. 이는 동일한 조건에서 ViT 92.7%, DCNN 89% 대비 높은 분류 성능이다.

[6]은 Vision Language Model(VLM)을 기반으로 중첩(Overlapping)된 LPI 레이더 변조 신호를 분류하는 기법을 제안하였다. VLM은 멀티모달 대조학습(Contrastive Learning) 메커니즘을 통해 이미지와 텍스트 임베딩을 공유 공간으로 정렬하여 제로샷(Zero-Shot) 학습이 가능한 모델이다. 여기서 제로샷 학습이란, 훈련 단계에서 보지 못한 중첩 신호의 조합을 단일 변조 신호의 학습만으로도 분류할 수 있는 능력을 의미한다. 해당 연구에서는 LFM, Barker, Frank, Costas 4개의 변조 신호와 각 중복 조합을 포함한 총 15개 신호를 SPWVD-TFI로 변환하여 입력으로 사용하였으며, 단일 변조 신호 4,000개 이미지로만 훈련하여 11,000개의 중첩 신호를 분류하는 데 성공하였다. -10dB ~ 10dB SNR 환경에서 분류 성능

을 평가한 결과, -2 dB 이상의 SNR 환경에서 100%에 가까운 분류 정확도와 1 ms/image의 추론시간을 달성하며 복잡한 전자기 환경에서 실시간 제로샷 분류의 가능성을 보여주었다.

## III. 결론

본 논문에서는 ViT 및 VLM 기반 LPI 레이더 변조 분류의 최신 연구 동향을 종합적으로 분석하였다. 기존 CNN 기반의 변조 분류 기법은 낮은 SNR 환경에서 배경 잡음의 영향으로 분류 성능이 급격히 저하되는 문제 가 있었으며 [2], 이를 극복하기 위해 ViT, VViT 기반의 기법들이 제안되었다 [4] [5]. 또한 VLM의 멀티모달 대조학습으로 이미지와 텍스트 임베딩이 공유 공간에 정렬되면서 제로샷 학습이 가능해졌으며, 훈련 데이터에 미포함된 중첩된 신호의 조합도 분류가 가능함을 보여주었다 [6]. 향후 연구에서는 이러한 모델들의 높은 연산 부잡도를 개선하여 실제 임베이드 ESM 시스템에 탑재할 수 있는 모델 경량화 기술과, 극저(Ultra-Low) SNR 환경에서도 신호 특징을 선명하게 추출할 수 있는 고도화된 신호 강조(Signal Enhancement) 기법에 대한 연구가 병행되어야 할 것이다. 이러한 연구들은 복잡한 전자기 환경에서 LPI 레이더 신호 분류의 신뢰성과 실시간 운용성을 향상시키는 데 기여할 것으로 기대된다.

## ACKNOWLEDGMENT

이 논문은 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No.2022R1A2C1011862)

## 참 고 문 헌

- [1] Yuan, Shibo, Bin Wu, and Peng Li. "Intra-pulse modulation classification of radar emitter signals based on a 1-d selective kernel convolutional neural network." *Remote Sensing* 13.14 (2021): 2799.
- [2] Kim, Junseob, et al. "Enhancing LPI radar signal classification through patch-based noise reduction." *IEEE Signal Processing Letters* 31 (2024): 716-720.
- [3] Jeong, Jechang, and William J. Williams. "Kernel design for reduced interference distributions." *IEEE Transactions on Signal Processing* 40.2 (1992): 402-412.
- [4] Kim, Junseob, et al. "Automatic lpi radar waveform recognition using vision transformer." *2023 IEEE International Radar Conference (RADAR)*. IEEE, 2023.
- [5] Bhatti, Sidra Ghayour, et al. "Transformer-based models for intrapulse modulation recognition of radar waveforms." *Engineering Applications of Artificial Intelligence* 136 (2024): 108989.
- [6] Yang, Pengkun, et al. "Automatic LPI radar waveform recognition of overlapping signals based on vision language model." *Scientific Reports* 15.1 (2025): 29757.