

거리-도플러 맵 기반 레이더 표적 분류를 위한 Vision Transformer 설계

Design of Vision Transformer for Range-Doppler Map-Based Radar Target Classification

최병찬 · 이성규 · 김지현 · 안세환 · 주지한

Byungchan Choi · Seonggyu Lee · Jihyun Kim · Schwan An · Jihan Joo

요 약

거리-도플러 맵은 레이더를 이용하여 안테나가 지향하고 있는 표적에서 산출할 수 있는 대표적인 데이터이다. 거리-도플러 맵을 통해 표적에 대한 상대거리, 상대속도, 반사 신호 세기를 산출할 수 있다. 표적의 형태에 따라 거리-도플러 맵에서의 신호 세기 분포가 달라진다. 이를 응용하여 레이더 시스템은 표적의 종류를 분류할 수 있다. 본 논문은 거리-도플러 맵으로부터 표적을 분류할 수 있는 vision transformer 기반 딥러닝 네트워크 설계법을 제시하고자 한다.

Abstract

Range-Doppler map is the principal information that a radar system can acquire from an antenna that is illuminating the target. Radar system can extract the relative range, relative velocity, and signal strength of the target from a range-Doppler map. The distribution of the signal strength varies according to the target shape. Using these characteristics, the radar system can distinguish between the target types. This paper proposes the design of vision transformer-based deep learning network for target classification from a range-Doppler map.

Key words: Range-Doppler Map, Radar Target Classification, Deep Learning, Vision Transformer, Supervised Learning

I. 서 론

안테나가 지향하는 방향에서 표적에 대해 반사된 신호로부터 상대거리와 상대속도를 산출하고 해당 표적의 종류를 판단하는 것이 레이더 표적 탐지의 주요 목적이다. 이를 위해서 레이더 시스템은 표적에 대한 거리-도플러 맵을 주로 사용한다. 거리-도플러 맵은 거리 축과 도플

러 축으로 이루어져 있다. 거리 축은 레이더와 표적 사이의 상대거리를 나타내며, 도플러 축은 레이더와 표적 사이의 상대속도를 나타낸다. 거리 축과 속도 축으로 이루어진 2차원 좌표계에서 각 좌표는 상대거리와 상대속도에 따른 표적 반사 신호 세기를 기록한다. 표적의 형태에 따라 반사 신호 세기의 분포가 달라진다. 표적의 형태에 따른 거리-도플러 맵에서의 신호 세기 분포는 표적을 분

LIG Nex1 미사일시스템탐색기연구소(RF Seeker R&D, LIG Nex1)

· Manuscript received October 28, 2024 ; Revised November 12, 2024 ; Accepted February 21, 2025. (ID No. 20241028-118)

· Corresponding Author: Byungchan Choi (e-mail: byungchan.choi@lignex1.com)

류하기 위한 feature로서 사용될 수 있다.

자율 주행, 군사용 등 레이더 기반 표적 탐지, 추적을 수행하기 위해서는 소량의 레이더 신호 데이터를 이용하여 단시간 내에 표적 분류를 수행해야 한다. 표적 신호의 SNR(signal-to-noise ratio)를 높이기 위해서는 여러 timestep에 걸쳐서 입력 신호 데이터를 누적해야 한다. 그러나 입력 신호 누적 시간이 길어질수록 표적 신호 수집에 대한 소요 시간이 증가하기에 표적 분류의 실시간성을 확보하기 힘들다. 또한 딥러닝 네트워크는 규모가 커질수록 연산 소요 시간이 증가한다. 실시간성을 강조하는 자율 주행, 군사용 등 레이더 기반 표적 탐지, 추적을 위해서는 규모가 작은 딥러닝 네트워크와 실행 구조를 구현해야 한다.

본 논문은 기존 연구 대비 적은 timestep의 전처리되지 않은 2D 원시 거리-도플러 맵으로부터 표적을 분류할 수 있는 ViT(vision transformer) 기반 딥러닝 네트워크 설계, 학습 전략, 실행 구조를 제안한다. 본 논문은 자동차, 드론, 사람에 대한 거리-도플러 맵으로 구성된 RDRD 데이터셋을 이용하여 표적을 분류할 수 있는 딥러닝 네트워크를 구현했다^{[1][2]}. 본 논문은 거리-도플러 맵의 대부분 영역이 noise 신호로 구성된 특성으로 인해 단일 ViT 네트워크가 과적합되는 경향을 최소화하고 높은 분류 성능을 도달하기 위해 앙상블 기반의 실행 구조를 제안한다. 본 논문이 제안한 ViT 기반 분류 네트워크는 기존 연구 대비 적은 딥러닝 파라미터와 단일 timestep 입력 신호를 사용했음에도 0.97 이상의 F1 score와 최대 오탐율 4.86 %를 도달했다. 기존 연구와 동일한 timestep의 입력 신호 데이터를 사용할 경우 본 논문의 ViT 기반 분류 네트워크는 상대적으로 적은 딥러닝 파라미터를 사용하면서 기존 연구와 유사한 F1 score 0.99 수준의 성능에 도달한다.

II. 관련 연구

2-1 고전적인 레이더 표적 탐지 및 분류 알고리즘

표적 탐지 레이더의 특성은 설계 및 제품마다 다르기에 어느 수준의 신호 세기가 유효한지에 대한 절대적인 지표가 없다. 고전적인 레이더 표적 탐지 알고리즘은 CFAR(constant false alarm rate)와 같은 adaptive filtering 기

법을 거리-도플러 맵에 적용하여 유효 표적 신호 구간을 산출한다. CFAR 알고리즘은 거리-도플러 맵에서 표적이 존재하는지 여부를 판단하는 지점을 CUT(cell under test)라 지칭한다. CFAR 알고리즘은 CUT의 주변 신호 상태에 따라 변경될 수 있는 adaptive threshold를 산출해야 한다. 이를 위해서 CUT를 기준으로 일정 개수의 cell을 이격시킨 후 threshold를 산출한다. CUT를 중심으로 이격시킨 cell을 guard cell이라 지칭한다. CUT에 대한 adaptive threshold를 산출할 때 사용하는 신호 데이터 범위는 guard cell 이후 N개 만큼 설정된 reference cell을 사용한다.

CFAR 알고리즘에서 guard cell이라는 개념을 사용하는 이유는 guard cell 구간에서는 CUT 성분이 강하게 잔류하기 때문이다. CFAR는 CUT가 주변 noise 신호 대비 강하게 나타나는지 여부를 유동적으로 파악하기 위해 유동적인 threshold를 산출해야 한다. CUT 성분이 강하게 잔류하는 guard cell 구간은 adaptive threshold 산출에 사용하지 않는다. CFAR는 CUT 주변 noise 성분이 주로 반영된 reference cell을 사용하여 adaptive threshold를 산출한다. 그림 1은 입력 신호 배열에 대해 CFAR 연산이 레이더 표적 탐지 및 추적 알고리즘에 사용되는 과정을 보여준다.

Threshold 산출 방식에 따라 CFAR 알고리즘 종류가 달라지게 된다. Reference cell 구간의 신호의 평균을 threshold로 사용할 경우 CA-CFAR(cell-average CFAR)라 지칭한다. CA-CFAR의 경우 CUT의 신호 세기가 reference cell 신호

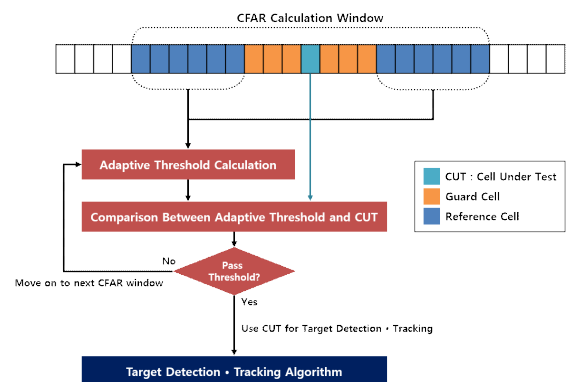


그림 1. 고전적인 CFAR 기반 레이더 표적 탐지 및 추적 알고리즘

Fig. 1. Classical CFAR-based radar target detection and tracking algorithm.

평균보다 높다면 유효 표적 신호로 간주한다. CUT 신호와 reference cell을 신호 세기에 따라 정렬하였을 때 CUT가 일정 수준 이상의 위치에 도달하는지 여부로 신호 유효성을 판단하는 경우 OS-CFAR(order-static CFAR)라 지칭한다. 기존 OS-CFAR의 경우 신호 세기에 따른 정렬을 위한 정렬 알고리즘이 필수로 요구되었다. 표적 신호 데이터 규모가 커질수록 OS-CFAR는 실시간 탐지에 부적합했으나 CUT가 reference cell의 신호 cell들 대비 일정 횟수 이상 큰지 여부를 판단하는 방식으로 정렬 연산을 대체함으로써 실시간 표적 탐지에 사용되고 있다^[3]. 레이더 시스템에서 CFAR는 이와 같은 adaptive threshold 산출하는 방식을 통해 유효 표적 신호 구간을 찾아내기 위한 전처리 과정으로 활발히 사용되어왔다^{[4],[5]}.

2.2 딥러닝 기반 표적 탐지 및 분류 알고리즘

대규모로 수집된 데이터셋을 이용하여 학습하는 딥러닝 기반 탐지 및 분류 네트워크가 고전적인 기법보다 영상처리 분야에서 큰 성과를 거두고 있다. 특히 CNN(convolutional neural network) 기반 네트워크인 AlexNet, ResNet YOLO 등이 물체 인식 및 탐지 분야에서 고전적인 영상처리 알고리즘을 압도하는 성능에 도달했다^{[6]-[8]}. CNN 기반 네트워크의 경우 이미지 전반에 걸쳐서 추출된 local feature를 최종 출력단에서 latent feature vector로 합치면서 이미지 전체에 대한 global feature를 학습하게 된다. 대규모 데이터셋을 이용하여 표적 탐지 및 분류 알고리즘을 딥러닝 네트워크로 만들어냄으로서 기존보다 높은 일반화 특성과 재구성성을 확보할 수 있게 되었다.

참고문헌 [1]은 X-band FMCW(frequency modulated continuous wave) 레이더로 수집한 차량, 드론, 사람에 대한 거리-도플러 맵 데이터셋을 이용하여 CNN 기반 표적 분류 네트워크인 DopplerNet을 제안했다^[1]. 해당 논문 저자들은 FMCW 레이더로 수집된 거리-도플러 맵에서 CFAR 알고리즘을 적용하여 표적 후보군의 위치를 파악했다. 표적 후보군 위치를 중심으로 원시 거리-도플러 맵을 11×61 크기로 잘라내어 데이터셋을 정리했다. 3개 연속적인 timestep에서의 원시 거리-도플러 맵을 이용하여 표적의 종류를 분류하도록 CNN 기반 딥러닝 네트워크를

설계했으며, 이를 평가하기 위해 F1 score를 평가지표로 활용했다. 참고문헌 [1]이 제안한 DopplerNet은 3개의 연속적인 timestep의 거리-도플러 맵을 1개 세트로 누적하여 입력 신호로 사용해야 최대 성능에 도달한다^[1]. 다중 timestep 동안 신호 획득에 따른 지연이 발생하며, 이로 인해 표적 분류 네트워크의 실시간성을 확보하기 힘들다. 본 논문은 참고문헌 [1] 저자들이 Kaggle에 등재한 데이터셋인 RDRD 데이터셋을 ViT 학습에 사용했다^[2].

참고문헌 [9]는 mmWave를 통해 추출된 micro-Doppler map과 ViT를 이용한 레이더 기반 사람 행동 인식 기법을 제안한다^[9]. 해당 논문에서 제안된 vision transformer 기반 딥러닝 네트워크는 원시 거리-도플러 맵이 아닌 2D CFAR를 통해 전처리된 micro-doppler map을 입력으로 사용한다. 2D CFAR를 통해 유효 표적 신호를 강조시킴으로서 ViT와 같은 복잡한 딥러닝 네트워크가 학습에 성공할 수 있도록 유도한다. 하지만 2D CFAR는 입력 신호 규모가 커질수록 연산량이 크게 증가하며, 이로 인해 시스템 전반의 실시간성을 확보하기 어렵다.

참고문헌 [10]은 거리-도플러 맵이 만들어지는 초기 신호처리 과정에서부터 표적 분류를 수행하는 최종단까지 분류 네트워크의 전 과정을 딥러닝으로 구현했다^[10]. 해당 논문 저자들은 레이더를 통해 수집된 원시 표적 신호에 대해 FFT(fast Fourier transform) 대신 딥러닝 기반 신호처리 기법을 적용하여 거리-도플러 맵 변환 시 ViT 학습에 적합한 feature가 반영되게 했다. FFT를 딥러닝으로 대체하면서 참고문헌 [10]이 제안한 딥러닝 네트워크는 증가한 연산량으로 인해 실시간성을 확보하기 힘들다. 또한 FFT와 표적 분류 2가지의 task가 순차적으로 실행되기에 성질이 다른 2가지의 네트워크를 동시에 학습시킬 수 없다. 이로 인해 각각의 네트워크를 학습시키기 위한 연산 자원 소요량과 복잡도가 증가하게 된다.

III. 본 론

3.1 데이터셋 특성 분석 및 운용 방식

본 논문에서 사용하는 RDRD 데이터셋은 X-band FMCW 레이더를 자동차, 드론, 사람에게 조사하여 수집

된 거리-도플러 맵 데이터 등으로 구성되어있다. 표 1은 RDRD 데이터셋에 사용된 FMCW 레이다 성능의 개요이다^[1]. RDRD 데이터셋 설명에 의하면 FMCW 레이다에서 수집된 전체 거리-도플러 맵에서 CFAR를 통해 산출된 유효 표적 신호 부위에 해당되는 원시 신호 데이터를 잘라내어 거리 축 11 cell, 도플러 축 61 cell로 데이터셋을 구성한다^[1]. RDRD 데이터셋은 총 17,405개로 구성되며, 각각 라벨별로 자동차 5,693개, 드론 5,044개, 사람 6,668개 거리-도플러 맵 데이터로 구성되어있다^[2]. RDRD 데이터셋의 거리-도플러 맵은 거리 축 11 cell, 도플러 축 61 cell로 2차원 구조로 구성되어있다. 거리-도플러 맵의 각 지점에서 수집된 반사 신호의 세기는 dB 단위로 표현되어있다.

거리-도플러 맵은 표적 신호에 해당되는 **peak cell** 영역을 제외하고는 **noise**로 구성되어있으며, **peak cell** 영역 면적의 비율은 **noise** 신호가 차지하는 면적 대비 작다. 표적 신호의 SNR이 낮을수록 표적 신호를 **noise**와 구분하기 어려워진다. SNR이 낮은 거리-도플러 맵을 입력받은 딥러닝 네트워크는 표적 **peak cell** 형태를 보고 학습하는 것이 아닌 **noise**를 학습하게 되면서 과적합이 발생하게 된다. 그림 2는 RDRD 데이터셋의 SNR 구간별 데이터 분포를 보여주며, 표 2는 SNR 구간별 데이터셋 샘플을 보여준다. RDRD 데이터셋에서 SNR 25 dB 미만 상황에서는 표적 신호와 신호 수집 환경에서 발생하는 **noise**가 구분되지 않는 것을 볼 수 있다. 딥러닝 네트워크가 표적 신호 대신 **noise**를 학습하는 경향으로 인한 과적합 방지를 위해서 표적 신호와 **noise** 신호가 구분되지 않는 데이터를 학습 및 검증 과정에서 제외해야 한다. 본 논문은 ViT 네트워크가 거리-도플러 맵에서 **peak cell** 대신 배경 **noise**로 표적 분류를 학습하면서 발생하는 과적합을 방지하기 위해 CA-CFAR를 통해 **noise level** 대비 SNR 25 dB 이상인 거리-도플러 맵만 학습 데이터로 선정하여 사용했다. 이러한 전처리 과정을 통해 자동차 4,956개, 드론 2,726개, 사람 6,003개 거리-도플러 맵을 학습 데이터로 선정했다.

레이다에서 수신된 표적 신호를 시간 축 상에서 지정된 개수만큼 신호 데이터를 **sampling**하고, 해당 데이터에 FFT를 적용하여 거리-도플러 맵을 만든다, 거리-도플러 맵은 기본적으로 **grayscale** 이미지로 취급할 수 있다. 그러나 거리 축으로 거리-도플러 맵을 해석할 경우 도플러

표 1. RDRD 데이터셋 FMCW 레이다 성능 개요^[1]

Table 1. Specification summary of FMCW radar used in RDRD dataset^[1].

Radar frequency	X-band
Type	FMCW
Bandwidth	200 MHz
Sample rate	32 MHz
Maximum beat frequency	16 MHz
Range resolution	0.878 m
Doppler resolution	5.58 Hz / 0.34 km/h

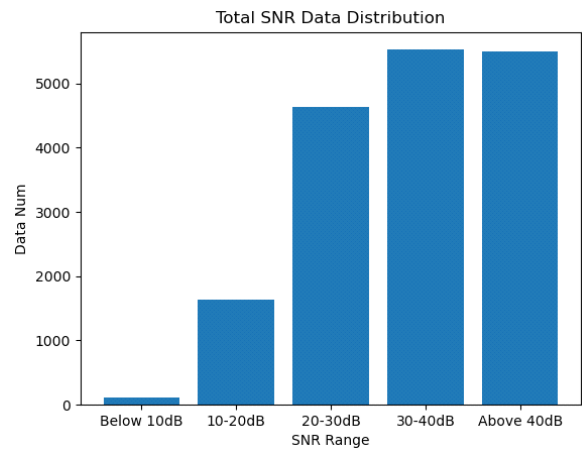


그림 2. SNR 구간별 RDRD 데이터셋 분포

Fig. 2. Data distribution of RDRD dataset by SNR range.

표 2. SNR 구간별 RDRD 데이터셋 샘플 (단위: dB)

Table 2. RDRD dataset sample by SNR range (unit: dB).

SNR Type	Below 10	10~25	25~40	Above 40
Car				
Drone				
People				

스펙트럼이 sampling 시간에 따라 순차적으로 만들어져 하나로 합쳐진 sequential 데이터가 거리-도플러 맵이라 간주할 수 있다. 도플러 축으로 거리-도플러 맵을 해석할 경우 표적 상대 거리에 따른 반사 신호 분포가 도플러 index에 따라 순차적으로 수집되어 하나로 합쳐진 sequential 데이터가 거리-도플러 맵이라 간주할 수 있다. 본 논문은 거리-도플러 맵의 구조를 고려하여 거리-도플러 맵을 이미지 데이터 또는 sequential 데이터라 간주하여 딥러닝 학습을 설계했다.

본 논문은 거리-도플러 맵을 sequential 데이터라 간주하여 학습하는 경우 입력 데이터 전체 sequence에 걸쳐서 전역적으로 학습할 수 있는 transformer를 도입했다^[11]. RNN(recurrent neural network) 계열의 딥러닝 네트워크 또한 sequential 데이터셋 학습에 사용될 수 있다. 하지만 RNN 계열의 딥러닝 네트워크는 sequence가 길어질수록 long-term dependency 현상에 의한 성능 저하가 발생할 수 있다. 특히 프로펠러가 있는 드론의 경우 프로펠러 회전의 상대 속도에 의해서 추가 도플러 스펙트럼이 표적 peak cell 이외의 구간에 발생할 수 있다. 특정 거리 index와 도플러 index에 표적 신호가 모이지 않기에 상대적으로 길어진 sequence를 학습해야 드론 표적을 분류할 수 있을 것이다. 이러한 드론 표적의 신호 특성에 의해 상대적으로 긴 sequence에 걸쳐서 학습하게 되면서 long-term dependency가 발생하여 분류 성능 저하가 발생할 수 있다. 본 논문은 거리-도플러 맵을 sequential 데이터라 간주하여 학습하는 과정에서 long-term dependency에 의한 성능 저하를 방지하고, sequential 데이터인 거리-도플러 맵을 전역적으로 학습할 수 있는 transformer 기반의 ViT 구조를 도입했다.

RDRD 데이터셋의 거리-도플러 맵을 이미지로 취급하여 ViT에 입력하기 위해서는 이미지의 가로, 세로 pixel 개수의 공약수를 기준으로 patch 사이즈를 설정해야한다^[11]. RDRD 데이터셋이 제공하는 거리-도플러 맵의 크기는 11×61로 1 이상의 공약수를 patch 사이즈로 설정할 수 없다. 거리-도플러 맵은 각 축의 index 그 자체에 거리, 속도 정보가 암묵적으로 적용되고, 그에 따라 신호 세기 분포가 설정된다. Resize를 통해 거리-도플러 맵의 크기를 변경할 수 있으나 resize 과정에서 interpolation이 적용되

어 신호 분포 변화가 발생할 수 있다. 본 논문은 거리-도플러 맵의 신호 분포와 거리축, 도플러축 index에 대한 정보를 보존하기 위해 zero padding을 적용하여 12×64 크기의 거리 도플러 맵으로 변환하였다.

분류 네트워크는 입력 데이터에 대해 어떤 label이 제일 유력한지를 판단해서 출력해야한다. 이를 위해 전체 label 개수를 반영하여 1×3 구조의 groundtruth 벡터를 만들며, 각 거리-도플러 맵의 label에 맞춰서 one hot encoding을 적용한다.

3-2 Vision Transformer 학습 설계

본 논문은 총 3가지의 ViT 네트워크를 개별적으로 학습시킨 후 하나의 앙상블로 통합하여 실행 구조를 만든다. Type 1 ViT 네트워크는 거리-도플러 맵을 이미지로 취급하여 표적을 분류하는 네트워크이다. 그림 3은 Type 1 ViT의 실행 프로세스를 보여준다. Type 1 ViT는 입력 이미지를 patch 단위로 분해하여 feature vector로 만든 후 transformer를 통해 모든 patch와 classification vector 사이의 연관성을 학습하여 이미지를 분류하는 네트워크 구조이다^[11]. Patch 단위 embedding vector 생성을 위해 2D CNN 기반의 patch embedding layer를 구현했다. kernel 크

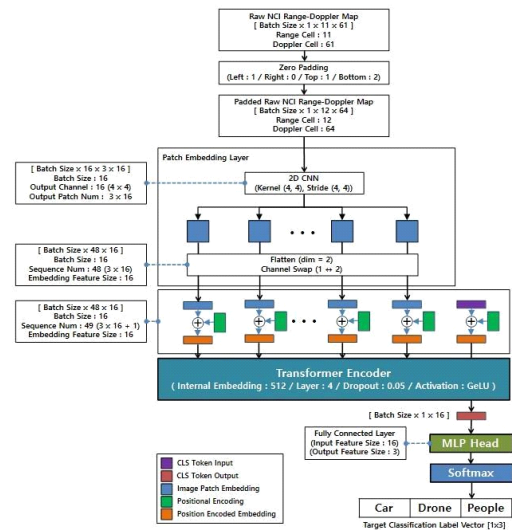


그림 3. Type 1: 이미지 입력 기반 ViT
Fig. 3. Type 1: Image input-based ViT.

기(4×4), Stride 크기(4×4)인 2D CNN layer를 통해 zero padding된 12×64 거리-도플러 맵을 4×4 patch 단위로 분해한다. 각 patch는 1×16의 embedding vector로 변환된다. 이후 flatten 과정과 channel swap 과정을 통해 batch size×sequence×embedding size 순서로 데이터를 정렬한다. patch embedding layer를 통해 거리-도플러 맵은 48개의 sequence로 이루어진 embedding vector로 구성된다. 각 배치 데이터는 48×16 구조의 데이터로 구성된다.

각 이미지 patch가 전체 이미지에서 어느 지점에서 추출된 것인지를 반영하기 위해 positional embedding을 각 이미지 patch의 embedding vector에 더한다. 분류 결과에 상응하는 데이터 출력을 만들기 위해 0으로 구성된 1×16 classification vector를 추가로 연결한다. 각 배치마다 49×16 크기의 데이터가 transformer에 입력된다. Transformer는 attention 메커니즘을 통해 모든 입력 embedding vector 사이의 연관성을 학습하여 49×16 크기의 vector를 출력한다^[12]. Transformer의 출력 중 classification vector 입력에 상응하는 출력 vector를 fully connected layer를 통과 시켜서 groundtruth label과 동일한 크기의 최종 출력 vector를 만들어낸다.

Type 2, Type 3 ViT 네트워크는 거리-도플러 맵을 sequential 데이터로 해석함으로써 표적의 sequence에 따른 도플러 스펙트럼 또는 신호 분포 변화를 인지하여 표적 종류를 분류하도록 딥러닝 네트워크의 내부 실행 과정을 설계했다.

Type 2 ViT 네트워크는 거리-도플러 맵을 거리 축을 sequential axis로 취급하여 sequential 도플러 스펙트럼 데이터에서 표적을 분류하는 네트워크이다. 그림 4는 Type 2 ViT 네트워크의 실행 프로세스를 보여준다. Type 2 ViT 네트워크는 Type 1 ViT 네트워크와 유사하게 12×64 거리-도플러 맵을 입력받아 각 거리 index의 1×64 도플러 스펙트럼 feature vector를 fully connected layer 기반의 embedding layer를 통과시켜 1×64 embedding vector로 변환시켰다. Type 1 ViT와 동일하게 classification vector를 입력 데이터에 추가하여 13×64 구조의 sequential 도플러 스펙트럼 데이터를 구성한다. 각 도플러 스펙트럼과 classification vector가 어느 sequence에서 추출된 것인지를 반영하기 위해 positional embedding을 더한다. Type 2 네트워

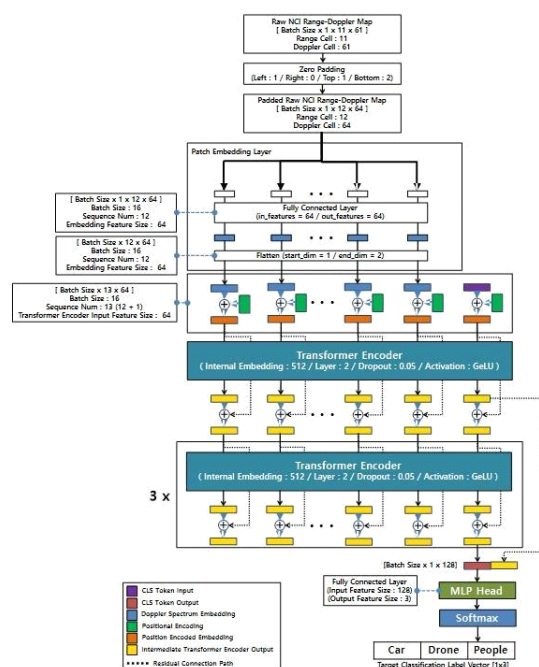


그림 4. Type 2: Sequential 도플러 스펙트럼 입력 기반 ViT
Fig. 4. Type 2: Sequential doppler spectrum input-based ViT.

크는 단계별 transformer 모듈의 입력값이 다음 모듈 입력 값에 통합되어 입력되도록 residual connection을 추가했다. 이를 통해 sequential 도플러 스펙트럼 입력 데이터로 학습하는 ViT 네트워크의 feature reusability를 강화시켜서 과적합 방지 능력을 보강했다. Type 3 ViT 네트워크는 Type 2 ViT 네트워크와 동일하게 fully connected layer 기반 embedding layer와 단계별 transformer 입출력 residual connection을 가진다. 그러나 Type 2 ViT 네트워크와 달리 거리-도플러 맵의 도플러 축을 sequence axis로 취급하여 sequential 신호 분포 데이터를 입력받는다. 그림 5는 Type 3 ViT 네트워크의 실행 프로세스를 보여준다.

본 논문의 ViT의 최종 목표는 표적 label 분류이기에 transformer와 fully connected layer의 최종 출력 vector에서 제일 높은 값에 대한 index가 추론 결과가 된다. 본 논문은 분류 학습을 지도하기 위한 loss function으로 분류 네트워크 학습 시 대표적으로 사용되는 cross entropy loss function을 적용한다. 학습을 지도하는 cross entropy loss function과 상응하기 위해 softmax layer를 최종 출력 fully connected layer에 추가한다. 딥러닝 네트워크 최적화를 위

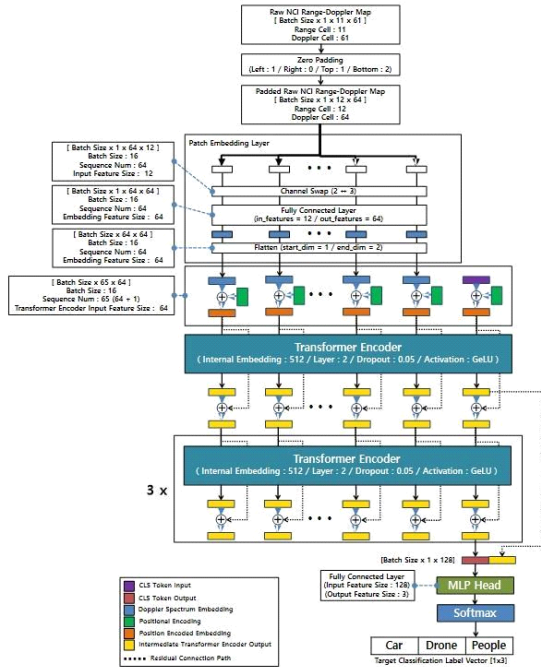


그림 5. Type 3: Sequential 신호 분포 입력 기반 ViT
Fig. 5. Type 3: Sequential signal strength distribution input-based ViT.

한 optimizer로 Adam optimizer를 이용했다^[13]. 본 논문은 학습 과정 중 과적합을 방지하여 학습 안정성을 확보하기 위해 입력 데이터에 대해 min-max normalization을 적용했다.

3.3 앙상블 기반 분류 시스템 실행 구조

거리-도플러 맵은 표적 신호를 제외하고는 대부분이 noise로 구성되어있다. 입력 데이터의 대부분이 noise로 구성된 경우 학습 과정 중 딥러닝 네트워크는 noise로 label 분류를 학습하게 되면서 과적합될 확률이 높다. 이와 같은 학습 조건에서는 단일 네트워크가 F1 score 0.95 이상의 준수한 성능에 도달할 수 있으나 과적합으로 인해 그 이상의 성능을 도달하는 데에 한계가 있다. 본 논문은 거리-도플러 맵을 3가지 다른 관점에서 학습하는 3개 type의 ViT 네트워크를 앙상블 구조로 묶어서 실행하는 과정을 사용했다. 동일한 거리-도플러 맵 입력에 대해 Type 1, 2, 3 ViT 네트워크는 각각의 label 분류를 위한 확

률 분포를 출력한다. 각 네트워크의 제일 높은 확률을 가진 label을 통합한 후 majority voting을 적용하여 최종 분류 결과를 산출한다. 그림 6는 앙상블 기반 분류 시스템 실행 구조를 보여준다. 앙상블 기반 실행 구조는 준수한 성능을 가진 각 개별 네트워크들의 결과를 통합함으로써 과적합이 상대적으로 낮게 나타나도록 추론 성능을 개선할 수 있다.

IV. 실험

4.1 학습 및 실험 설정

본 논문에서는 학습을 위해 전체 데이터셋에서 training set 60 %, validation set 20 %, test set 20 % 비율로 나눠서 운용한다. Training set을 이용하여 딥러닝 네트워크를 학습시킨다. Validation set은 overfitting 없이 학습이 성공적으로 수렴하는지 학습 과정 중 관찰하는 데에 사용한다. Test set은 학습이 완료된 네트워크의 재현성 확인을 위해 사용한다. Pytorch를 이용하여 ViT 기반 표적 분류 네트

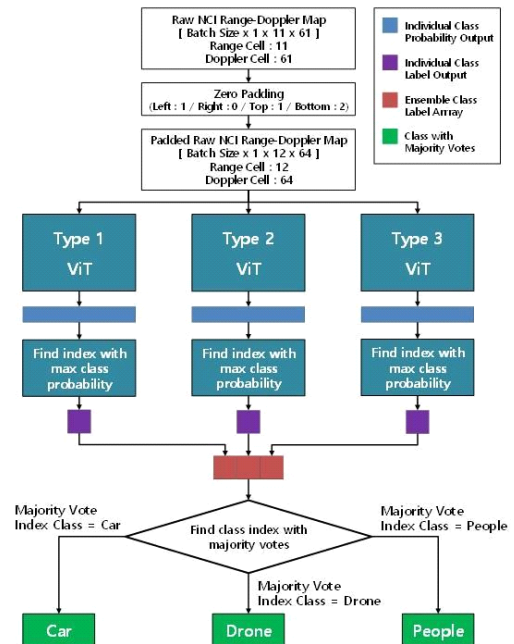


그림 6. 앙상블 기반 분류 시스템 구조 실행 과정
Fig. 6. Processing flow of ensemble-based classification system structure.

워크를 구현했다^[14]. Nvidia GPU를 이용하여 총 100 epoch 학습이 수행되었다. 학습된 딥러닝 네트워크의 성능을 판단하기 위해서 F1 score를 사용했다.

4.2 학습 결과

거리-도플러 맵에 대한 표적 분류 학습 가능성을 확인하기 위해 단일 timestep 거리-도플러 맵으로 본 논문의 ViT 네트워크를 학습시켰다. 매 epoch의 양상블 실행 결과의 loss function 평균과 F1 score를 그림 7 및 그림 8과 같이 기록했다. 그림 7에 제시된 loss function 기록을 확인한 결과 총 100회 학습을 수행하면서 전반적으로 training set에서의 loss function과 validation set에서의 loss function이 일치하여 같이 낮아지는 것을 볼 수 있다. 그림 8을 통해 네트워크의 분류 정확도와 오분류 정도를 모두 반영한 성능 지표인 F1 score 또한 training set과 validation set에서 동일하게 epoch마다 개선되는 것을 볼 수 있다. 학

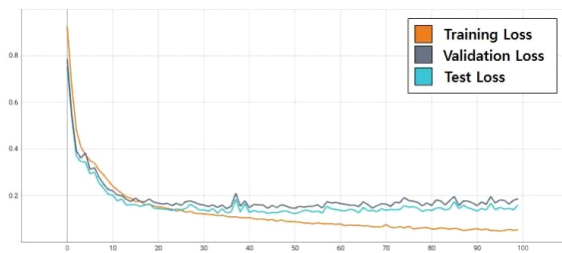


그림 7. 단일 timestep 입력 시 epoch마다 기록된 loss graph
Fig. 7. Loss graph recorded at every epoch with single timestep input.

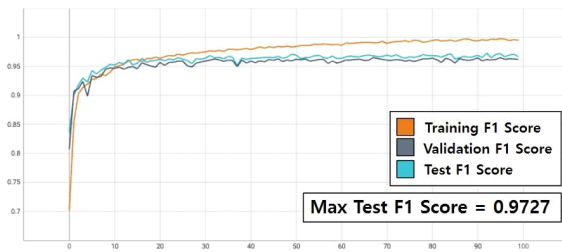


그림 8. 단일 timestep 입력 시 epoch마다 기록된 F1 score
Fig. 8. F1 score recorded at every epoch with single timestep input.

습을 지도하는 loss function과 성능 지표인 F1 score가 training set과 validation set에서 매 epoch마다 개선되고 있기에 본 논문이 제시한 ViT 기반 네트워크가 학습 가능한 것을 판단할 수 있다. test set에 대해 단일 timestep을 사용하는 경우 본 논문이 제시한 네트워크는 최대 F1 score 0.9727에 도달했다. 그림 9는 최대 F1 score에 도달했을 시의 confusion matrix이며, 본 논문이 제시한 네트워크의 최대 오분류 확률은 4.86 %이다.

4.3 비교 실험 결과

레이다 표적 탐지에서 표적 신호에 대한 SNR을 증가시키기 위해 NCI(non-coherent integration)를 사용한다. 본 논문은 NCI가 적용되지 않은 단일 timestep에서의 학습 진행 결과를 통해 본 논문이 제시한 네트워크의 학습 가능성을 확인했다. Timestep을 증가시키면서 NCI를 적용한 상황에서 본 논문의 분류 네트워크 성능을 DopplerNet, ResNet, VGG, LeNet과 비교했다^{[11],[7],[15],[16]}.

표 3는 네트워크 성능 비교 결과를 보여준다. Timestep을 증가시키면서 NCI를 적용한 거리-도플러 맵을 사용할 경우 본 논문의 F1 score가 증가하는 것을 볼 수 있다. 3개

Groundtruth	Cars	95.14%	4.40%	0.46%
	Drones	4.10%	95.66%	0.24%
	People	0.29%	0.00%	99.71%
		Cars	Drones	People
		Prediction		

그림 9. 최대 F1 score에서의 test set confusion matrix
Fig. 9. Test set confusion matrix at max F1 score.

표 3. 제안된 방식, DopplerNet, pretrained 모델의 F1 score 비교

Table 3. F1 score comparison between proposed method, DopplerNet, and pretrained models.

Model type	Input timestep	Integration type	F1 score	Neural network parameter
LeNet ^[15]	1	-	0.9641	5,602,979
	2	NCI	0.9749	5,602,979
	3	NCI	0.9840	5,602,979
ResNet18	1	-	0.9571	11,178,051
	2	NCI	0.9749	11,178,051
	3	NCI	0.9825	11,178,051
ResNet50	1	-	0.9592	23,514,179
	2	NCI	0.9691	23,514,179
	3	NCI	0.9798	23,514,179
VGG19 ^[16]	1	-	0.9559	139,593,539
	2	NCI	0.9746	139,593,539
	3	NCI	0.9813	139,593,539
NasNetMobile ^{[11],[17]}	3	Channel-wise concatenation	0.9769	4,272,887
MobileNetV2 ^{[11],[18]}	3	Channel-wise concatenation	0.9894	3,325,043
DopplerNet ^[1]	3	Channel-wise concatenation	0.9948	3,818,755
Proposed method	1	-	0.9727	1,406,473
	2	NCI	0.9850	1,406,473
	3	NCI	0.9912	1,406,473

연속 timestep의 거리-도플러 맵을 channel-wise concatenation한 DopplerNet 연구 결과와 비교했을 시 본 논문의 네트워크는 DopplerNet 대비 37 %의 모델 파라미터를 사용했음에도 동일하게 F1 score 0.99 수준에 도달한다. 동일 timestep을 사용한 경우 본 논문의 네트워크가 ResNet, VGG, LeNet, NasNetMobile 보다 높은 F1 score에 도달하는 것을 볼 수 있다. 이를 통해 본 논문이 제시하는 ViT 기반 딥러닝 네트워크와 앙상블 기반 분류 시스템 실행 구조가 거리-도플러 맵을 이용한 표적 분류에 적합한 방식이라는 것을 확인할 수 있다.

V. 결 론

본 논문은 전처리되지 않은 원시 거리-도플러 맵을 이용하여 레이더 표적 분류를 수행할 수 있는 ViT 기반 딥러닝 네트워크를 제안했다. 본 논문은 FMCW로 수집된 거리-도플러 맵 데이터셋을 이용하여 학습을 수행했다. 본 논문은 거리-도플러 맵을 sequential 도플러 스펙트럼 또는 sequential 신호 분포 데이터로 운용하는 관점을 제시했다. 본 논문은 기존 연구보다 적은 모델 파라미터로 레이더 표적 분류를 할 수 있는 딥러닝 네트워크 설계 및 학습법을 제안했다. 그리고 거리-도플러 맵의 표적 신호 이외 noise가 강하게 발생하는 것을 고려하여 과적합을 최소화시킬 수 있는 학습 전략과 앙상블 기반 실행 구조를 제안했다. 본 논문이 제안한 네트워크와 실행 시스템 구조는 입력 데이터의 timestep이 증가할수록 성능이 개선되었으며, 대부분의 pretrained 모델보다 좋은 분류 성능에 도달했다.

본 논문은 데이터셋에서 제공되는 단일 표적에 대한 거리-도플러 맵만을 사용하여 학습을 수행했다. 레이더는 자율주행, 로봇 등 여러 분야에서 핵심 인지 센서로 사용되며, 다중 표적 탐지 능력이 요구된다. 본 논문의 연구 결과를 바탕으로 다중 표적 탐지를 위한 딥러닝 네트워크로 발전시킬 수 있다.

References

- [1] I. Roldan, C. R. del-Blanco, Á. D. de Quevedo, F. I. Urzaiz, J. G. Menoyo, and A. A. López, et al., "Dopplernet: A convolutional neural network for recognising targets in real scenarios using a persistent range-Doppler radar," *IET Radar, Sonar & Navigation*, vol. 14, no. 4, pp. 593-600, Apr. 2020.
- [2] I. Roldan, "Real Doppler radar database," 2019. Available: <https://www.kaggle.com/datasets/ioldan/real-doppler-radar-database>
- [3] Y. J. Kong, S. K. Woo, S. Park, S. Y. Shin, Y. H. Jang, and E. Yang, "OS CFAR computation time reduction technique to apply radar system in real

- time," *The Journal of Korean Institute of Electromagnetic Engineering and Science*, vol. 29, no. 10, pp. 791-798, Oct. 2018.
- [4] J. H. Nam, I. S. Koh, S. H. Park, and J. M. Lee, "Target detection and tracking simulation of air defense radar in high-density clutter environments using a real-time simulator," *The Journal of Korean Institute of Electromagnetic Engineering and Science*, vol. 31, no. 8, pp. 738-745, Aug. 2020.
- [5] T. H. Kim, J. H. Shin, S. W. Lee, J. H. Park, S. H. Jang, and S. J. Kim, "Design of waveform and signal processing of target detection for detecting closely spaced airborne targets in airborne radar," *The Journal of Korean Institute of Electromagnetic Engineering and Science*, vol. 31, no. 2, pp. 154-164, Feb. 2020.
- [6] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, vol. 25, 2012.
- [7] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition(CVPR)*, Las Vegas, NV, Jun. 2016, pp. 770-778.
- [8] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *2016 IEEE Conference on Computer Vision and Pattern Recognition(CVPR)*, Las Vegas, NV, Jun. 2016, pp. 779-788.
- [9] S. Huan, Z. Wang, X. Wang, L. Wu, X. Yang, and H. Huang, et al., "A lightweight hybrid vision transformer network for radar-based human activity recognition," *Scientific Reports*, vol. 13, p. 17996, Oct. 2023.
- [10] J. Giroux, M. Bouchard, and R. Laganiere, "T-FF-TRadNet: Object detection with swin vision transformers from raw ADC radar signals," in *2023 IEEE/CVF International Conference on Computer Vision Workshops(ICCVW)*, Paris, Oct. 2023, pp. 4030-4039.
- [11] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, and T. Unterthiner, et al., "An image is worth 16x16 words: Transformers for image recognition at scale," 2020. Available: <https://doi.org/10.48550/arXiv.2010.11929>
- [12] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, and A. N. Gomez, et al., "Attention is all you need," *Advances in Neural Information Processing Systems*, vol. 30, 2017.
- [13] D. P. Kingma, J. Ba, "Adam: A method for stochastic optimization," 2014. Available: <https://doi.org/10.48550/arXiv.1412.6980>
- [14] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, and G. Chanan, et al., "Pytorch: An imperative style, high-performance deep learning library," 2019. Available: <https://doi.org/10.48550/arXiv.1912.01703>
- [15] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, and D. Anguelov, et al., "Going deeper with convolutions," in *2015 IEEE Conference on Computer Vision and Pattern Recognition(CVPR)*, Boston, MA, Jun. 2015, pp. 1-9.
- [16] K. Simonyan, A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014. Available: <https://doi.org/10.48550/arXiv.1409.1556>
- [17] B. Zoph, V. Vasudevan, J. Shlens, and Q. V. Le, "Learning transferable architectures for scalable image recognition," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR)*, Salt Lake City, UT, Jun. 2018, pp. 8697-8710.
- [18] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR)*, Salt Lake City, UT, Jun. 2018, pp. 4510-4520.

최 병 찬 [LIG Nex1 미사일시스템탐색기연구소/선
임연구원]

<https://orcid.org/0000-0001-6902-6370>



2016년 2월: 한양대학교 ERICA 전자통신
공학과 (공학사)

2022년 2월: 한양대학교 전자공학과 (공학
석사)

2022년 2월~현재: LIG Nex1 미사일시스
템 탐색기연구소 선임연구원

[주 관심분야] 딥러닝, RF, 임베디드 SW 등

안 세 환 [LIG Nex1 미사일시스템탐색기연구소/수
석연구원]

<https://orcid.org/0000-0002-8955-9067>



2004년 2월: 숭실대학교 정보통신공학과
(공학사)

2006년 2월: 숭실대학교 정보통신공학과
(공학석사)

2007년 1월~현재: LIG Nex1 미사일시스
템탐색기연구소 수석연구원

[주 관심분야] 초고주파 회로 설계,

마이크로파 송수신기 등

이 성 규 [LIG Nex1 미사일시스템탐색기연구소/연
구원]

<https://orcid.org/0009-0000-1033-7940>



2024년 2월: 광운대학교 전자융합공학과
(공학사)

2024년 2월~현재: LIG Nex1 미사일시스
템탐색기연구소 연구원

[주 관심분야] RF, 임베디드 SW 등

주 지 한 [LIG Nex1 미사일시스템탐색기연구소/팀장]

<https://orcid.org/0000-0003-0643-6246>



2002년 8월: 충북대학교 전파공학과 (공학
사)

2004년 8월: 광운대학교 전파공학과 (공학
석사)

2008년 8월: 광운대학교 전파공학과 (공학
박사)

2008년 7월~현재: LIG Nex1 미사일시스
템탐색기연구소 팀장

[주 관심분야] 마이크로웨이브 탐색기, 초고주파 회로 시스템
등

김 지 현 [LIG Nex1 미사일시스템탐색기연구원 수석
연구원]

<https://orcid.org/0000-0001-5417-4522>



2004년 2월: 부산대학교

전자전기통신공학과 (공학사)

2007년 2월: 부산대학교 전자공학과
(공학석사)

2007년 2월~현재: LIG Nex1 미사일시스
템 수석연구원

LIG Nex1 미사일시스템탐색기연구소수

석연구원

[주 관심분야] RF, 임베디드 SW 등