

## 복소수 ResNet 네트워크 기반의 SAR 영상 물체 인식 알고리즘

황인수<sup>\*,1)</sup><sup>1)</sup> 국방과학연구소 국방첨단기술연구원

## A Complex Valued ResNet Network Based Object Detection Algorithm in SAR Images

Insu Hwang<sup>\*,1)</sup><sup>1)</sup> The Intelligence & Information Technology Directorate, Agency for Defense Development, Korea

(Received 30 March 2021 / Revised 31 May 2021 / Accepted 18 June 2021)

## Abstract

Unlike optical equipment, SAR(Synthetic Aperture Radar) has the advantage of obtaining images in all weather, and object detection in SAR images is an important issue. Generally, deep learning-based object detection was mainly performed in real-valued network using only amplitude of SAR image. Since the SAR image is complex data consist of amplitude and phase data, a complex-valued network is required. In this paper, a complex-valued ResNet network is proposed. SAR image object detection was performed by combining the ROI transformer detector specialized for aerial image detection and the proposed complex-valued ResNet. It was confirmed that higher accuracy was obtained in complex-valued network than in existing real-valued network.

Key Words : 복소수 네트워크(Complex Valued Network), 합성 개구면 레이더(Synthetic Aperture Radar), 물체인식(Object Detection), 복소수 레즈넷(Complex ResNet)

## 1. 서론

SAR은 EO(Electro-Optica) 장비와는 다르게 기상 조건과 상관없이 전천후로 영상을 얻을 수 있는 감시정찰 장비이다. 또한 SAR 영상의 해상도 또한 지속적으로 향상되고 있다<sup>1)</sup>. 이러한 SAR 영상에서의 물체 인식은 민간 혹은 국방 분야에서 많은 필요성을 가진다<sup>2)</sup>.

가령 해상 환경에서 해상 교통 모니터링, 어선 관리, 해양 환경 보호 등을 위하여 SAR 영상에서의 해양 선박 탐지는 중요한 이슈이다<sup>3-5)</sup>. 이때 딥러닝의 자동 특징 추출 특성을 활용하면 사람이 직접 영상을 판독하거나, 특징을 정의할 필요 없이 편리하게 선박, 항공기 등의 물체 인식을 수행할 수 있다<sup>6)</sup>.

SAR 영상은 지면에 반사된 레이더 신호의 한 종류이기 때문에 진폭과 위상을 가지는 복소수 형태이다. SAR 영상에서의 진폭은 지형이나 물체의 레이더 반사도 정보를 가지고, 위상은 레이더와 물체 사이의 거

\* Corresponding author, E-mail: hciinsu@add.re.kr  
Copyright © The Korea Institute of Military Science and Technology

리 정보를 가진다. 기존의 SAR 영상에서의 물체 인식을 수행할 때는 위상 데이터의 처리가 어렵다는 점 때문에 주로 진폭만을 이용하였다는 한계가 있다<sup>[3-9]</sup>.

본 논문에서는 SAR 영상의 진폭과 위상 데이터를 모두 활용하여 물체 인식을 수행하였다. 기존의 실수 기반의 네트워크로는 복소수 연산을 처리 할 수 없기 때문에 복소수 기반(complex-valued) ResNet 네트워크(network)를 제안하였고, 이를 활용하여 물체 인식을 수행하였다. 기존에 물체인식을 위한 복소수 기반 네트워크는 VGGNet 같은 간단한 모델을 수정하였지만<sup>[10]</sup>, ResNet은 이보다 훨씬 깊고 넓기 때문에, 성능이 좋다는 장점을 가진다. 실험에는 독일의 TerraSAR 위성의 0.6 m, 1 m 해상도를 가지는 영상과 이탈리아 CosmoSkymed의 1 m 해상도를 가지는 영상을 데이터셋(dataset)으로 활용하였다. 각 영상에는 선박과 항공기가 레이블링 되어 있다. 데이터셋을 활용하여 선박과 항공기를 대상으로 탐지 성능을 분석하였다. 진폭과 위상 데이터를 함께 이용하게 되면 단순히 진폭만을 활용하였을 경우보다 더 뛰어난 정확도 성능을 보였다.

## 2. 관련 연구

### 2.1 ROI transformer<sup>[11]</sup>

항공 영상에서의 물체 인식은 물체가 회전된 상태로 복잡한 환경과 섞여 있어 이를 잘 탐지하는 것이 중요하다. 기존의 HROI(Horizontal ROI)를 활용할 경우 회전 되어 있는 물체를 잘 찾지 못하는 문제가 발생할 수 있다. ROI transformer를 활용하면 처음에는 HROI와 rotated ground truth 사이의 오프셋(offset)을 학습하여 RROI를 생성한다. 생성된 RROI를 이용하여 다시 피쳐맵(feature map)에서 피쳐 추출(feature extraction)을 수행하는 방식으로 물체 인식을 수행한다.

하지만 논문에서의 ROI transformer는 RGB 영상의 회전된 물체 인식에만 특화되어 있다는 한계가 있다. SAR 영상은 RGB 형태가 아닌 복소수의 실수부(real), 허수부(imaginary)를 가지는 2 타입의 데이터를 가지는 영상으로, SAR 영상에서 탐지를 위한 방안이 필요하다.

### 2.2 Deep complex network<sup>[12]</sup>

복소수 신호에 대한 연구는 신호 및 이미지 처리 분야에서 중요한 이슈이다. 위상 처리가 다루기 어렵

기 때문에 진폭만 이용하여 신호처리를 수행하는 경우가 많았다. 복소수 데이터를 활용하면 기존에 진폭만 사용하여 신호처리를 한 경우보다 좀 더 깊은 특성을 얻을 수 있다. Deep complex network에서는 복소수가  $z = a + ib$ 의 형태인 점을 이용하여 컨볼루션(convolution)과 배치 노멀라이제이션(batch normalization)을 복소수 연산이 가능하도록 구현 하였다. 초기화는 이때, 자비에 초기화(xavier initialization)을 활용하였다. 다만 학습을 수행시에 복소수 데이터를 활용하지 않았고, 실수 데이터로 복소수를 임의로 생성하여 실험을 수행하였다는 한계가 있다.

본 논문에서는 실제 SAR 영상의 복소수 데이터를 모두 활용할 수 있도록 복소수 기반 ResNet 네트워크를 제안하였고, 초기화는 백본(backbone)인 ResNet50에서의 프리트레인(pretrain)된 값을 활용하여 자비에 초기화를 활용한 초기화보다 뛰어난 탐지 성능을 내었다.

## 3. Complex-Valued Network

### 3.1 Complex Building Blocks

복소수 컨볼루션(complex convolution), 복소수 배치 노멀라이제이션(complex batch normalization), 복소수 렐루(complex ReLU) 등은 기 적용된 수식을 활용하였다<sup>[10,12,13]</sup>.

#### 3.1.1 Complex convolution

필터  $W$ 는 복소수 이므로  $W_r + i^* W_i$ 로 나타낼 수 있고, 입력  $X$ 역시 복소수 이므로  $X_r + i^* X_i$ 와 같이 나타낼 수 있다. 따라서 두 컨볼루션 곱은 수식 (1)과 같이 계산이 가능하다.

$$W \circ X = (W_r + i^* W_i) \circ (X_r + i^* X_i) = (W_r \circ X_r - W_i \circ X_i) + i^* (W_r \circ X_i + W_i \circ X_r). \quad (1)$$

#### 3.1.2 Complex batch normalization

복소수 배치 노멀라이제이션에서는 수식 (2)을 통해서 입력  $x$ 의 분포를  $\bar{x}$ 의 분포로 수정한다. 이때  $V$ 는  $2 \times 2$  covariance matrix이다.  $V^{-\frac{1}{2}}$ 는 수식 (3)을 통해서 계산이 가능하다.

$$\bar{x} = V^{-\frac{1}{2}}(x - E[x]), \quad (2)$$

$$\left( E(\text{평균}) : \begin{bmatrix} E_r \\ E_i \end{bmatrix} V(\text{분산}) : \begin{bmatrix} V_{rr} & V_{ri} \\ V_{ir} & V_{ii} \end{bmatrix} \right)$$

$$V^{-\frac{1}{2}} = \frac{1}{ST} \begin{bmatrix} V_{ii} + S & -V_{ri} \\ -V_{ir} & V_{rr} + T \end{bmatrix}, \quad (3)$$

여기서, S와 T는 아래의 수식 (4), (5)와 같이 표현된다.

$$S = (V_{rr} \circ V_{ii} - V_{ri}^2)^{\frac{1}{2}}, \quad (4)$$

$$T = (V_{rr} + V_{ii} + 2S)^{\frac{1}{2}}, \quad (5)$$

최종으로 BN의 결과는 수식 (6)과 같이 정리 할 수 있다.

$$BN(\bar{x}) = \gamma \bar{x} + \beta. \quad (6)$$

( $\beta$ : shift parameter,  $\gamma$ : scaling parameter)

### 3.1.3 Complex ReLU

실수부와 허수부의 연산을 따로 수행하므로 렐루 함수는 수식 (7)과 같이 분리해서 연산을 수행한다.

$$ReLU_{Complex}(Z) = ReLU(Z_r) + i * ReLU(Z_i). \quad (7)$$

## 3.2 Proposed complex initialization method

일반적으로 복소수 알고리즘에서는 프리트레인된 값을 활용하지 않고, 자비에 초기화 등을 사용하였다. [12,14,15] 본 논문에서는 일반적인 초기화 방법이 아닌 백본(ResNet50)에서 전달 받은 프리트레인 된 학습 파라미터들을 활용하였다.

### 3.2.1 Complex convolution initialization

컨볼루션의 경우에 프리트레인 된 첫 번째 conv1의 가중치(weight)는 (64, 3, 7, 7)의 형태로 전달 받는다.

이때 3은 RGB로 입력 받아서 출력 사이즈가 64가 됨을 의미한다. 복소수 연산의 경우에는 RGB 대신 복소수인 실수, 허수 데이터 형태로 2장을 받아서 사이즈가 맞지 않게 된다. 이 때문에 conv1에서는 오직 R 값만 받아서 컨볼루션을 수행할 수 있도록 (64, 1, 7, 7)의 형태로 값을 받아서 I, Q에 각각 동일하게 복사

하여 입력 받도록 하였다. 첫 번째 컨볼루션 이후에 다른 복소수 컨볼루션은 실수 컨볼루션과 커널(kernel) 사이즈가 동일하여 프리트레인에서의 값 그대로 초기화에 사용하였다. 즉, I, Q 값에 pretrained kernel 실수 값을 동일하게 복사하였다.

### 3.2.2 Complex batch normalization initialization

배치 노멀라이제이션의 경우에 바이어스(bias)와 평균은 각각  $\begin{bmatrix} B_r \\ B_i \end{bmatrix}$ ,  $\begin{bmatrix} E_r \\ E_i \end{bmatrix}$ 의 형태로 실수 배치 노멀라이제이션에서 프리트레인된 값을 실수부, 허수부 각각에 동일하게 초기화 하였다. 그리고 가중치는  $\begin{bmatrix} W_{rr} & W_{ri} \\ W_{ir} & W_{ii} \end{bmatrix}$ 의 형태이므로,  $W_{rr}, W_{ii}$ 에는 실수 배치 노멀라이제이션에서 프리트레인된 가중치 값으로 초기화 하였고,  $W_{ri}$ 에는 0으로 초기화 하였다.  $\begin{bmatrix} V_{rr} & V_{ri} \\ V_{ir} & V_{ii} \end{bmatrix}$ 에서도 가중치와 동일한 방식으로 실수 배치 노멀라이제이션의 프리트레인된 가중치 값으로 초기화 하였다.

## 3.3 Proposed complex residual network

물체 인식을 위해서 일반적으로 프리트레인된 백본 모델을 사용한다. 뉴럴 네트워크는 넓고 깊어질수록 성능이 좋아지는 특성을 보이는데, 백본으로 선택한 ResNet은 기존의 VGGNet이나 GoogLeNet 등의 네트워크보다 훨씬 깊고 학습이 용이하다. 기존에는 간단한 VGGNet 등을 복소수 기반 VGGNet으로 바꾸어 물체 인식에 수행하였다.[10] 본 논문에서는 이보다 복잡하지 만 깊고 학습이 용이한 ResNet50을 백본으로 선택하여 복소수 연산이 가능하도록 한 복소수 기반 ResNet을 제안하였다. ResNet50은 Fig. 1과 같이 첫 번째 residual block으로 입력되기 전에 conv → bn → ReLU → maxpool을 거친다. Residual block 내에서는 conv → bn → conv → bn → conv → bn → ReLU 반복하여 거치게 된다. 복소수 알고리즘을 구현하기 위하여 컨볼루션, 배치 노멀라이제이션, 렐루, 맥스 풀링(maxpooling)을 모두 복소수 연산이 가능하도록 복소수 컨볼루션, 복소수 배치 노멀라이제이션, 복소수 렐루, 복소수 맥스 풀링(complex maxpool)으로 수정하였다.

Fig. 2는 제안한 복소수 ResNet과 ROI transformer 기반의 detector를 합쳐서 만든 복소수 기반 ResNet 네트워크이다. I, Q 신호를 2 채널로 입력받아서 복소수 ResNet을 지나서, 실수, 허수 두 데이터를 RMS(root

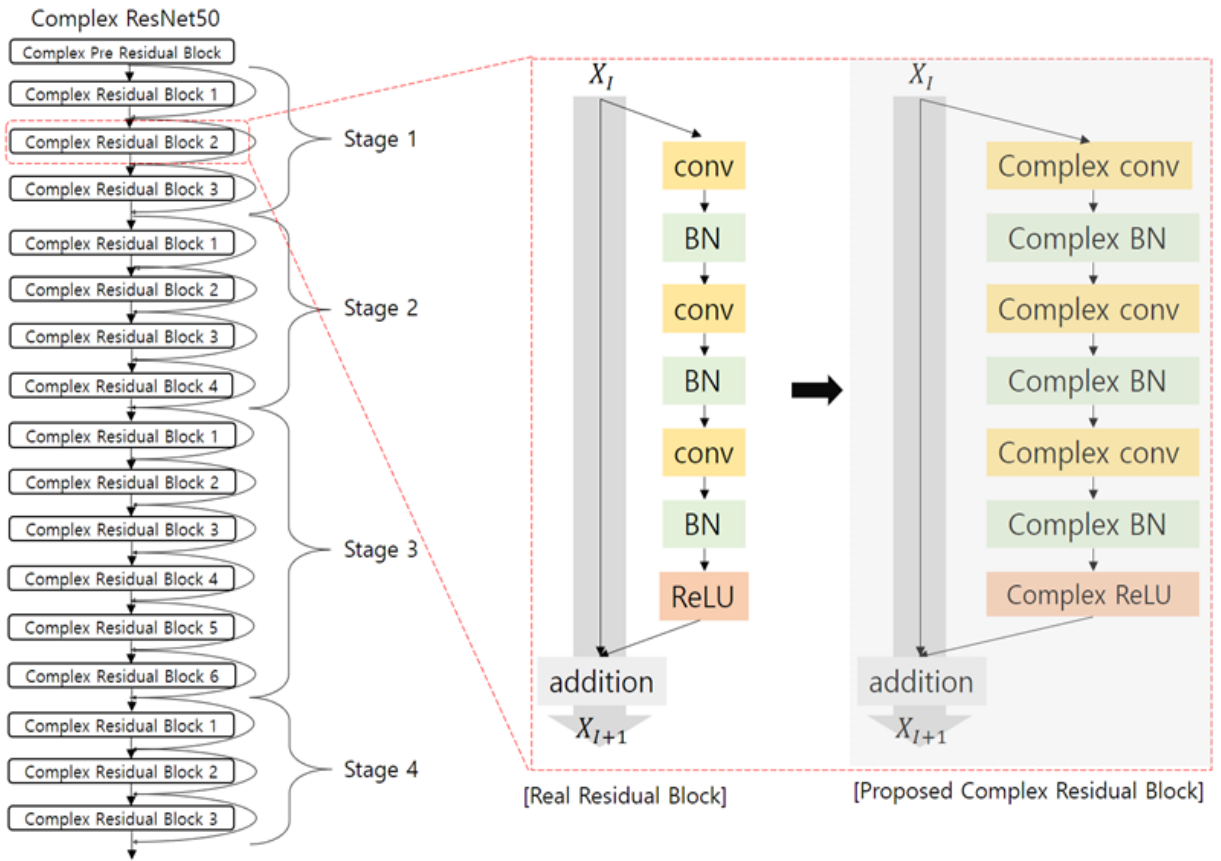


Fig. 1. Complex ResNet50, Comparison between real residual block & complex residual block

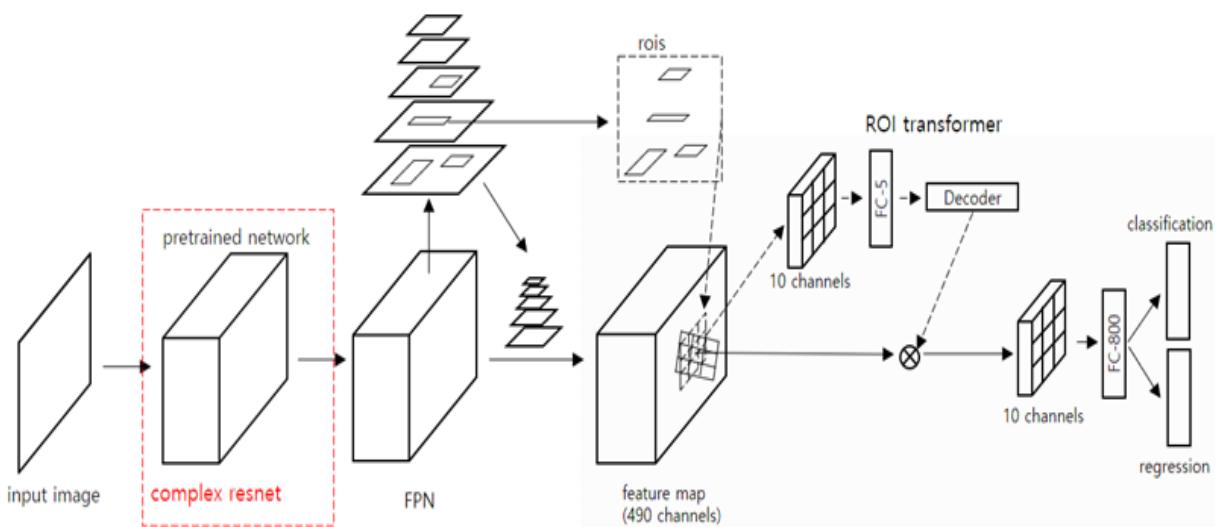


Fig. 2. Complex-valued ResNet network with ROI transformer

mean square) 형태로 합친다. 2 채널 데이터가 복소수 기반 ResNet 네트워크를 지난 후 1채널 실수 데이터로 변환 되고, FPN(feature pyramid network)층에서 5개의 피쳐맵을 뽑아낸다. 그 후 RPN에서 피쳐맵 별로 roi를 선별하고, 선별된 roi들이 ROI transformer에서 HROI로 입력된다. HROI와 rotated ground truth 사이의 오프셋 만큼 학습하여, RROI를 생성하고, 생성된 RROI로 다시 피쳐 추출을 수행한다. 이때 FPN과 ROI transformer는 일반 실수 네트워크와 동일한 구조이다. 따라서 backbone에 대한 부분만 복소수 연산이 가능하도록 구현되었고, neck과 head에 대한 부분은 일반 실수 연산 네트워크와 동일하다.

#### 4. 실험 방법

##### 4.1 Dataset

실험에 사용된 데이터셋은 중대형 위성인 TerraSAR의 0.6 m 해상도 영상 50장과 1 m 해상도 영상 10장, 중대형 위성 CosmoSkymed의 1 m 해상도 영상 55장을 활용하였다. 사용한 SAR 영상은 MGD(Multi-Look Ground Detected) 영상으로, 3 채널의 데이터를 가진다. 첫 번째 채널은 I값이 저장되어 있고, 두 번째 채널은 Q값이 저장되어 있고, 세 번째 채널에는  $\sqrt{I^2+Q^2}$ 이 저장되어 있다. MGD는 SSC(Single-Look Slant Range Complex) 영상에서 노이즈를 제거하고 Multi-looking을 수행하여 이미지의 품질이 향상된 영상이다.

각각의 SAR 영상들에는 선박과 항공기가 레이블링 되어 있다. Table 1에는 해상도 별로 영상 숫자와 레이블링 되어 있는 배와 항공기 객체의 숫자를 정리한 것이다. 총 영상 115장에 배 11811개, 항공기 3730개가 레이블링되어 있다.

Table 1. SAR satellite image labeling status

영상	규격	영상 숫자	배	항공기
Terra-SAR	0.6 m × 0.25 m	50	2906	902
	1 m × 1 m	10	1989	552
Cosmo Skymed	1 m × 1 m	55	6916	2276
합계	-	115	11811	3730

영상 숫자의 한계 때문에 학습 세트(train set)와 검증 세트(validation set)는 같은 영상 세트를 사용하였고, 테스트 세트(test set)는 별도의 영상 세트를 사용하였다. SAR 원본 영상은 한 장의 크기가 20000 × 20000으로 학습시키기에는 크기가 크고, SAR 원본 영상 개수의 한계 때문에 800 × 800 패치로 원본 영상을 잘라서 사용하였다. 학습/검증 : 테스트 영상 패치의 개수 비율은 8:2이며 랜덤으로 분리하였다. 이렇게 하여 학습/검증으로 4661개, 테스트로 1165개의 패치를 사용하였다.

##### 4.2 Implementation details

실험 수행시 비교를 위한 베이스 라인인 실수 기반의 네트워크에서는 SAR의 진폭(intensity) 데이터만을 활용하므로, ResNet50이 RGB 기반의 3영상을 입력 받지 않고, 이중 R 값에만 진폭 데이터를 입력하여 실험을 수행하였다.

복소수 기반 네트워크에서는 복소수 데이터를 활용하여 실험을 위하여, ResNet50의 RGB 3영상을 입력하는 대신 R과 G의 자리에 각각 I, Q 두 데이터를 2채널로 입력 받았다. 즉, 입력 복소수 데이터는 A+B\*i 형태에서 2채널로 각각 실수부(A), 허수부(B)를 입력 받았다. 복소수 데이터는 복소수 ResNet에 입력전에 [-1 1] 사이로 정규화를 먼저 수행을 하였다. 이때 복소수 연산은 실수, 허수 두 데이터를 활용하므로, residual 네트워크 연산이 끝나는 지점에서 2 채널 데이터를 RMS 형태로 합쳐서 1 채널의 실수 데이터를 만들고 난 후에 FPN에서 피쳐맵을 뽑도록 구현 하였다.

#### 5. 실험 결과

##### 5.1 Comparison between real-valued network vs complex-valued network

실험 결과로 기존의 실수 기반의 네트워크와 제안한 복소수 기반 네트워크를 활용하여 SAR 영상에서 물체인식을 수행한 결과를 Table 2에서 비교하였다. 실험에서 학습율은 0.005이고 총 150 에폭(epoch) 실험을 수행하였다. 실험 결과로 복소수 기반 네트워크를 활용한 경우가 mAP 기준으로 49.04 %의 정확도를 보여서 실수 기반의 네트워크에서의 46.56 % 보다 높은 정확도를 보였다. 이때, mAP는 이진 분류기의 성능 지표로 객체 검출 알고리즘의 성능을 평가하는데 주

로 사용된다. 개별적으로 항공기와 선박에 대한 탐지 정확도도 복소수 기반 네트워크에서 51.31 %, 46.76 % 로 실수 기반의 네트워크에서의 46.97 %, 46.15 % 보다 뛰어났다.

Table 2. Comparison of test result between real-valued network & complex-valued network

	항공기(%)	선박(%)	mAP(%)
real-valued network	46.97	46.15	46.56
complex-valued network	51.31	46.76	49.04

5.2 Comparison between convolution & batch normalization parameter train/freeze

복소수 ResNet 내에서의 컨벌루션, 배치 노멀라이제이션의 파라미터들을 학습 혹은 비학습 할 때의 경우를 Table 3과 같이 나누어서 총 100 에폭 만큼 실험을 수행해보았다. 모두 프리트레인된 값을 학습 파라미터 초기값으로 사용하였다. 파라미터로는 컨벌루션의 가중치, 배치 노멀라이제이션의 평균, 분산, 가중치, 바이어스의 학습 유무에 따라 실험을 수행하였다. 실험 결과로 컨벌루션 가중치는 학습을 수행하고, 배치 노멀라이제이션의 평균, 분산은 학습 하지 않고, 배치 노멀라이제이션의 바이어스, 가중치는 학습을 수행한 경우 mAP가 47.36 %로 성능이 가장 높았다.

5.3 Comparison between pretrained initialization model & not pretrained initialization model

Table 4는 Table 3에서 성능이 가장 좋았던 컨벌루션 가중치는 학습을 수행하고, 배치 노멀라이제이션의 평균, 분산은 학습 하지 않고, 배치 노멀라이제이션의 바이어스, 가중치는 학습한 경우에서, ResNet50의 프리트레인 값을 초기값으로 사용 유무에 따른 성능 비교 실험을 총 100 에폭 만큼 수행한 것이다. 표 4의 실험 결과를 보면 초기값으로 프리트레인된 값을 사용하지 않고, 다른 일반적인 복소수 논문과 같이 단순히 자비에 초기화<sup>[12,14,15]</sup>를 사용하였을 경우가 mAP가 40.23 %로 가장 성능이 떨어지는 것을 확인할 수 있었다. 그 다음으로 컨벌루션 가중치만 프리트레인된 값을 사용한 경우가 성능이 좋았고, 컨벌루션과 배치 노멀라이제이션 모두에서 프리트레인된 값을 학습 파

라미터 초기값으로 사용한 경우가 mAP가 47.36 %로 가장 정확도 성능이 좋았다.

Table 3. Comparison of test result with (a) conv weight freeze, bn mean, variance, weight, bias train, (b) conv weight train, bn mean, variance, weight, bias train, (c) conv weight freeze, bn mean, variance freeze & bn weight, bias train, (d) conv weight train, bn mean, variance, weight, bias freeze, (e) conv weight train, bn mean, variance freeze & bn weight, bias train

	conv weight train	bn mean, variance train	bn weight, bias train	항공기 (%)	선박 (%)	mAP (%)
(a)	X (freeze)	O	O	9.09	11.93	10.51
(b)	O	O	O	9.09	11.6	10.35
(c)	X (freeze)	X (freeze)	O	33.35	36.14	34.75
(d)	O	X (freeze)	X (freeze)	42.9	39.68	41.29
(e)	O	X (freeze)	O	<b>49.43</b>	<b>45.29</b>	<b>47.36</b>

Table 4. Comparison of test result with not pretrained, only conv weight pretrained, all parameters (conv, bn) are pretrained

Method	항공기(%)	선박(%)	mAP(%)
not pretrained	40.3	40.15	40.23
only conv weight pretrained	47.71	43.69	45.7
conv, bn parameters pretrained (proposed)	<b>49.43</b>	<b>45.29</b>	<b>47.36</b>

5.4 Comparison between RPN&RCNN nms threshold variation

Table 5는 복소수 기반 네트워크의 RPN에서의 NMS (Non maximum suppression) 임계치(threshold), RCNN detector에서의 NMS 임계치 등의 파라미터를 변화시

켜가면서 총 150 에폭의 실험을 수행한 결과이다. 이때, RPN은 우수한 region proposal을 생성하기 위해 end-to-end로 학습을 하는 모델이고, RCNN은 이미지 분류를 수행하는 CNN과 RPN 알고리즘을 연결한 모델이다. NMS 임계치를 통해서 동일한 물체에 대해서 여러개의 박스가 생성되는 것을 막을 수 있다. 실험 결과로 RPN의 NMS 임계치가 0.7에서 0.8로 커지게 되면 mAP가 48.4 %에서 48.45 %로 높아졌다. 또한 detector의 NMS 임계치가 0.1에서 0.25로 커지게 되면 mAP가 48.45 %에서 49.04 %로 높아졌다. NMS 임계치가 커지게 되면 부정확한 roi 박스는 제거하게 되므로 탐지 정확도가 상승하게 된다.

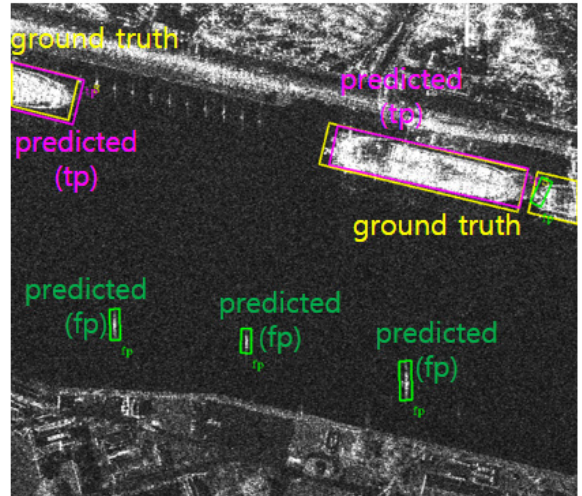
Table 5. Comparison of test result with RPN & RCNN nms threshold change

	Method	항공기 (%)	선박 (%)	mAP (%)
(a)	RPN nms_thr : 0.7 RCNN nms_thr : 0.1	50.56	46.24	48.4
(b)	RPN nms_thr : 0.8 RCNN nms_thr : 0.1	50.61	46.29	48.45
(c)	RPN nms_thr : 0.8 RCNN nms_thr : 0.25	51.31	46.74	49.04

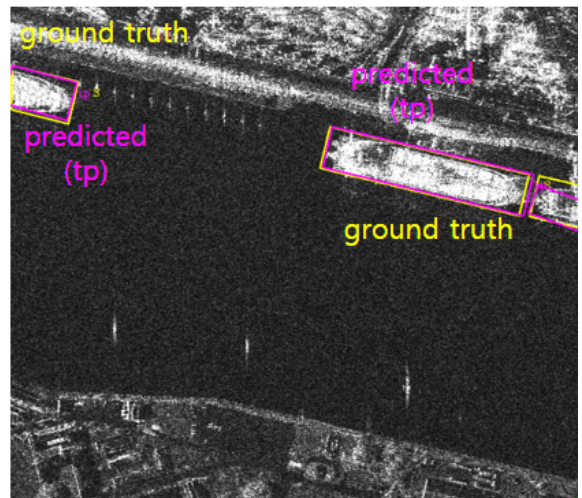
5.5 Qualitative analysis

Fig. 3-5는 SAR 패치 영상으로, 노란색 박스는 ground truth, 분홍색 박스는 true positive, 초록색 박스는 false positive, 주황색 표시는 false negative이다. 테스트를 수행할 때 RROI와 RGT(Rotated Ground truth) 사이의 IoU가 0.5 이상이면 true positive로, 그 이하이면 false positive로 판단한다. 해상 SAR 영상인 Fig. 3을 확인해보면 진폭 정보만을 이용하였을 경우보다 복소수 데이터를 모두 활용하였을 경우에 좀 더 정확한 선박 물체 인식을 수행한 것을 확인할 수 있었다. Fig. 3에서의 mAP는 (a)에서 28.6 %, (b)에서 84.3 %였다. Fig. 4는 공항 SAR 영상으로, Fig. 3과 마찬가지로 복소수 데이터를 모두 활용하였을 경우에 좀 더 정확한 항공기 물체 인식을 수행하였다. Fig. 4에서의 mAP는 (a)에서 50.2 %, (b)에서 99 %였다. Fig. 5는 해상 SAR 영상에서 진폭 정보만을 이용하였을 때 탐

지하지 못한 ground truth인 false negative가 발생하였는데, 복소수 데이터를 활용하여 더 정확한 선박 물체 인식을 수행한 것을 확인하였다. Fig. 5에서의 mAP는 (a)에서 66.7 %, (b)에서 79.6 %였다.



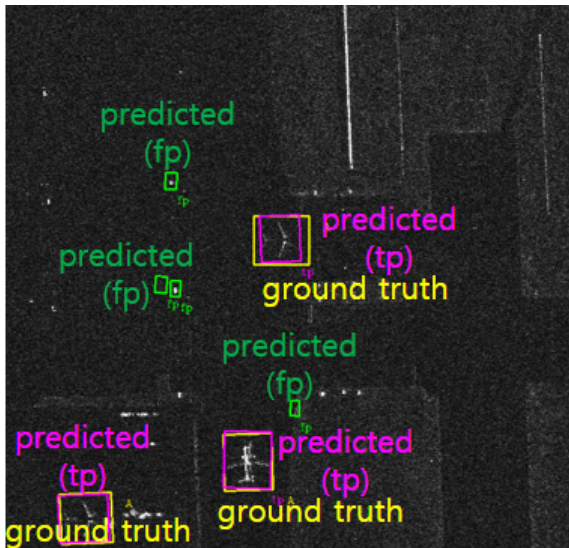
(a) real-valued network



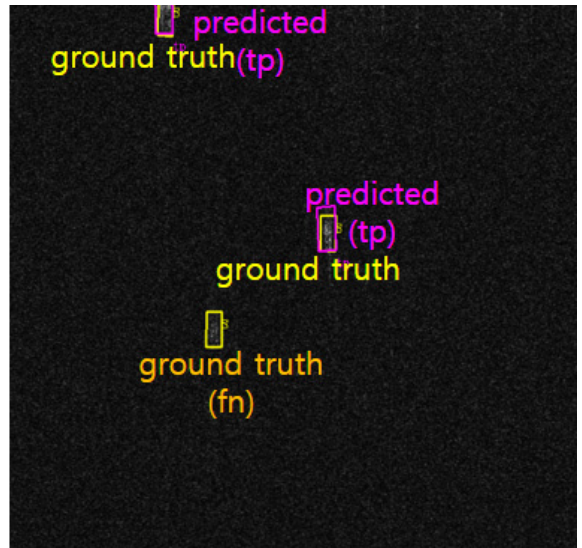
(b) complex-valued network

Fig. 3. Comparison of ship detection between real-valued network detection result(upper) vs complex-valued network detection result (lower); yellow box : ground truth box, pink box : true positive box, green box : false positive box

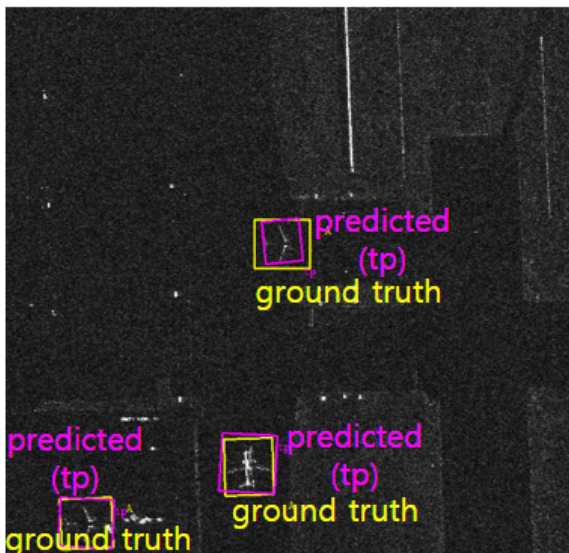




(a) real-valued network

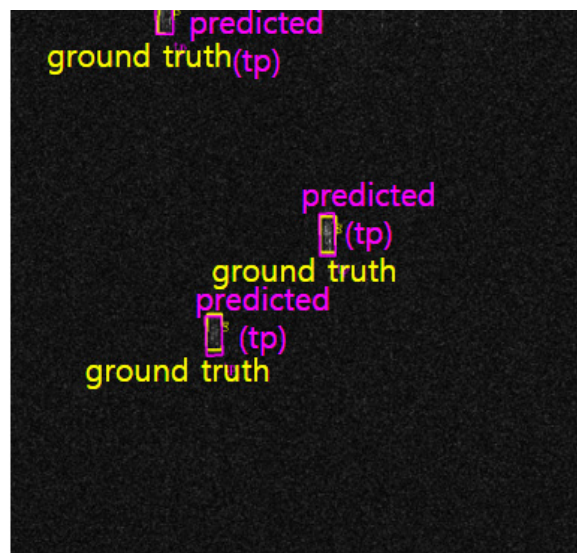


(a) real-valued network



(b) complex-valued network

Fig. 4. Comparison of airplane detection between real-valued network detection result(upper) vs complex-valued network detection result (lower); yellow box : ground truth box, pink box : true positive box, green box : false positive box



(b) complex-valued network

Fig. 5. Comparison of ship detection between real-valued network detection result(upper) vs complex-valued network detection result (lower); yellow box : ground truth box, pink box : true positive box, orange letter : false negative



## 6. 결론

SAR 영상에서의 물체 인식은 해상도의 향상 등으로 점차 활용도가 높아지고 있다. 본 논문에서는 복소수 기반 ResNet 네트워크를 제안하여 SAR 영상에서 물체인식을 수행하였다. 항공 영상은 물체가 회전되어 있는 경우가 많기 때문에 이에 특화된 detector인 ROI transformer 기반에, 제안한 복소수 ResNet을 합쳐서 SAR 영상 물체 인식을 수행하였다. 데이터셋으로는 독일의 TerraSAR 위성의 0.6 m, 1 m SAR 영상과 이탈리아 CosmoSkymed 위성의 1 m SAR 영상을 활용하였다. 선박과 항공기를 대상으로 물체 인식을 수행하였고, 기존의 실수 기반의 네트워크보다 복소수 기반의 네트워크에서 더 높은 정확도가 나오는 것을 확인할 수 있었다. 또한 기존의 복소수 기반의 네트워크들과는 달리 프리트레이닝 값을 초기값으로 활용하여 더 높은 정확도를 얻을 수 있었다.

## References

- [1] G. R. Benitz, "Synthetic Aperture Radar(SAR) Tutorial," MTI Lincoln Laboratory, 2009.
- [2] W. Li, J. Chen, X. Shi, A. Liu, "An Optimization Beam-position Selection Method Based on Genetic Algorithm for Distributed Satellite-borne SAR System," Proceedings of IEEE CIE International Conference on Radar, pp. 895-898, 2011.
- [3] T. Zhang, X. Zhang, J. Shi, S. Wei, J. Wang, J. Li, "Balanced Feature Pyramid Network for Ship Detection in Synthetic Aperture Radar Images," IEEE Radar Conference, pp. 1-5, 2020.
- [4] L. Liu, G. Chen, Z. Pan, B. Lei, Q. An, "Inshore Ship Detection in SAR Images Based on Deep Neural Networks," IEEE Geoscience and Remote Sensing Symposium, pp. 25-28, 2018.
- [5] M. Kang, X. Leng, Z. Lin, K. Li, "A Modified Faster R-CNN Based on CFAR Algorithm for SAR Ship Detection," International Workshop on Remote Sensing with Intelligent Processing, pp. 1-4, 2017.
- [6] Y. Li, Z. Ding, Y. Wang, J. Chen, "SAR Ship Detection Based on ResNet and Transfer Learning," IEEE International Geoscience and Remote Sensing Symposium, pp. 1188-1191, 2019.
- [7] J. Ding, B. Chen, H. Liu, M. Huang, "Convolutional Neural Network with Data Augmentation for SAR Target Recognition," IEEE Geoscience and Remote Sensing Letters, pp. 364-368, 2016.
- [8] T. Yang, J. Zhu, J. Liu, "SAR Image Target Detection and Recognition Based on Deep Network," SAR in Big Data Era, pp.1-4, 2019.
- [9] Y. Wang, Y. Zhang, H. Qu, Q. Tian, "Target Detection and Recognition Based on Convolutional Neural Network for SAR Image," International Congress on Image and Signal Processing, pp. 1-5, 2018.
- [10] S. Gu, L. Ding, "A Complex-valued VGG Network Based Deep Learning Algorithm for Image Recognition," Ninth International Conference of Intelligent Control and Information Processing, pp. 340-343, 2018.
- [11] J. Ding, N. Xue, Y. Long, G.S Xia, Q. Lu, "Learning ROI Transformer for Oriented Object Detection in Aerial Images," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2849-2858, 2019.
- [12] C. Trabelsi, O. Bilaniuk, Y. Zhang, "Deep Complex Networks," International Conference on Learning Representations, 2017.
- [13] Z. Sun, X. Xu, Z. Pan, "SAR ATR Using Complex-valued CNN" Asia-pacific Conference of Image Processing, Electronics and Computers, pp. 125-128, 2020.
- [14] H. Choi, J. Kim, J. Huh, A. Kim, J. Ha, K. Lee, "Phase-aware Speech Enhancement with Deep Complex U-net," International Conference on Learning Representations, 2019.
- [15] X. Glorot, Y. Bengio, "Understanding the Difficulty of Training Deep Feedforward Neural Networks," In Aistats, pp. 249-256, 2019.