

# 임계값 학습 모듈을 적용한 준지도 SAR 이미지 분류

## Semi-supervised SAR Image Classification with Threshold Learning Module

도재준<sup>1</sup> · 김선옥<sup>2\*</sup>

한국항공대학교 인공지능학과<sup>1</sup>, 한국항공대학교 소프트웨어학과<sup>2</sup>

### 요약

준지도 학습(Semi-supervised learning)은 소량의 라벨이 있는 데이터와 다량의 라벨이 없는 데이터를 사용하여 모델을 훈련하는 효과적인 방법이다. 그러나 많은 논문에서 준지도 학습시 하나의 고정된 임계값을 사용하여 각 클래스별 서로 다른 이미지들의 특징별 차이를 고려하지 않고 임의 라벨을 만든다. 본 논문에서는 합성개구 레이더(SAR) 이미지 분류 준지도 학습시 모든 클래스가 하나의 고정된 임계값을 사용하는 대신 각 클래스에 대해 서로 다른 임계값을 적용한다. 모델에 임계값 학습 모듈을 추가하여 임계값을 학습하여 클래스별로 학습되는 차이를 고려하여 클래스별로 서로 다른 임계값을 얻는다. 서로 다른 임계값을 사용한 준지도 학습 기반의 SAR 이미지 분류 방법을 적용유무를 비교하여 클래스별 임계값을 사용하는 이점에 대해 고찰하였다.

■ 중심어 : 준지도 학습, 의사 라벨링, SAR 이미지 분류 학습, 임계값 학습 모듈

### Abstract

Semi-supervised learning (SSL) is an effective approach to training models using a small amount of labeled data and a larger amount of unlabeled data. However, many papers in the field use a fixed threshold when applying pseudo-labels without considering the feature-wise differences among images of different classes. In this paper, we propose a SSL method for synthetic aperture radar (SAR) image classification that applies different thresholds for each class instead of using a single fixed threshold for all classes. We propose a threshold learning module into the model, considering the differences in feature distributions among classes, to dynamically learn thresholds for each class. We compare the application of a SSL SAR image classification method using different thresholds and examined the advantages of employing class-specific thresholds.

■ Keyword : Semi-supervised Learning, Pseudo Labeling, SAR Classification, Threshold Train Module

2023년 11월 24일 접수; 2023년 12월 06일 수정본 접수; 2023년 12월 11일 게재 확정.

\* 이 출판물은 2021년도 한국항공대학교 교비지원 연구비에 의하여 지원된 연구의 결과임.

† 교신저자 (sunok.kim@kau.kr)

## I. 서론

합성 개구 레이더인 SAR 이미지는 지상이 아닌 상공에서 찍은 이미지로 구름이 잔뜩 끼거나 밤에는 촬영이 어려운 일반적인 RGB 영상과는 달리 날씨와 시간에 관계없이 촬영이 가능하다는 장점이 있다. 이런 SAR 이미지의 장점은 밤낮과 구름 낀 날씨 등, RGB 이미지로는 지상의 상황을 파악하기 힘든 상황에서 활용도가 매우 높아 군사적 목적의 감시나 재난 상황 파악, 지하 광물 파악, 선박 수 파악 등 다양한 응용 분야에서 사용 가능한 데이터셋이다.

최근 딥러닝과 인공지능 모델을 이용한 RGB 이미지의 분류 모델[1-3]들의 성능이 크게 향상되었다. SAR 이미지 분류 분야도 딥러닝과 인공지능 모델을 적용한 논문[4-6]들이 있다. 이 논문들은 딥러닝의 방식 중 하나인 지도학습 방식을 이용한다. 이 논문들은 공통적으로 SAR 이미지에서 유용한 특징 정보들을 모델에 학습시켜 과거의 수식적인 분류 방식보다 큰 성능적 향상을 확인할 수 있다. 하지만 지도학습의 특성상 많은 수의 품질이 좋은 라벨링이 된 데이터셋이 필요하다는 단점이 있다. 특히 촬영에 많은 시간과 비용이 드는 SAR 이미지는 사람이 알아보기도 힘들어 라벨링이 된 고품질의 대규모 데이터셋을 구하기가 힘들다는 문제점이 있다.

이런 데이터의 양과 질에 대한 문제점들을 해결하기 위해 최근에는 라벨이 있는 소수의 데이터와 라벨이 없는 다수의 데이터를 섞어서 사용하는 준지도 학습 방법론에 대한 많은 연구가 이뤄지고 있다. 대표적인 준지도 학습 방법으로는 noisy student[7], pseudo label[8], MixMatch[9], ReMixMatch[10], FixMatch[11], SimCLR[12] 등이 있다.

이런 준지도 학습 방법의 가장 큰 장점은 데이터셋을 구성하는데 많은 시간이 소요되지 않는다는 것이다. 지도 학습은 수많은 이미지들에 일

일이 라벨을 붙여 데이터셋을 구성하기까지 많은 시간이 소요되지만 준지도 학습은 일부분만 라벨을 붙여도 충분하기에 데이터셋을 구성하는 것에 있어 큰 장점이 된다. 특히 본 연구에서 사용하고자 하는 SAR 데이터셋은 획득이 힘들고 라벨링을 하는데 많은 시간과 인력이 소모되는 이미지 데이터셋이다.

본 논문은 SAR 이미지를 데이터셋으로 사용하기에 기존에 제안된 일반적인 RGB 이미지 데이터셋들을 기반으로 하는 데이터 증강 방법들을 사용하지 않는다. 대신 SAR 이미지 데이터 분류 모델의 성능을 향상시키기 위해 고안된 데이터 증강 기술인 speckle noise를 사용한다. 그리고 대부분의 준지도 학습법은 라벨이 없는 데이터를 학습할 때, 클래스 별로 서로 다른 특징으로 인한 학습 난이도를 고려하지 않은 고정된 임계값을 사용하여 학습을 진행한다. 그 때문에 유의미한 정보를 담은 많은 정보들이 미세한 차이로 임계값을 넘지 못했다는 이유만으로 버려진다. 클래스 별로 서로 다른 학습 난이도를 고려하여 클래스 별로 임계값을 학습하는 threshold learning module을 추가하여 클래스 간의 미세한 차이를 고려한 pseudo labeling을 하여 학습을 진행한다.

## II. 관련 연구

### 2.1 준지도 학습법

준지도 학습은 양질의 많은 라벨이 있는 데이터셋이 필요한 지도 학습의 단점을 보완하고자 나온 딥러닝 모델의 학습 방법으로 등장했다. 준지도 학습은 라벨이 지정된 데이터와 지정되지 않은 데이터의 정보 모두를 활용하여 모델 성능과 일반화 성능을 개선하는 것이 목표이다. 준지도 학습 방법론은 지도 학습과는 달리 라벨이 지정되지 않은 데이터도 학습에 활용하여 데이터셋의 제약에서 벗어나 다양하고 포괄적인

데이터셋을 사용할 수 있게 한다.

준지도 학습의 이점은 다양한 분야에서 활용되며, 자연어 처리[13,14], 컴퓨터 비전[11,12] 및 음성 인식[15,16] 등 다양한 분야에서 효과적이고 쓸모 있는 딥러닝 모델 학습 방법으로 사용된다. 컴퓨터 비전의 이미지 분류 분야에서 사용되는 대표적인 방법으로 Consistency Regularization[17]과 pseudo labeling[8], Mixup 정규화 방식을 사용한 MixMatch[9]와 이를 한층 발전시킨 ReMixMatch[10], 신뢰도 임계값을 사용한 Fix-Match[11] 등이 있다.

Consistency regularization은 현재 존재하는 대부분의 준지도 학습 방법들의 근간이 되는 개념이다. Sajjadi의 논문[17]에서 제안된 개념으로 딥러닝 모델의 일반화와 입력 데이터의 노이즈에 대한 견고성을 개선하는 데 효과적인 성능을 보여 널리 사용되는 기술이다. 데이터를 변형하고 왜곡시켜 증강한 데이터를 입력으로 주었을 때 학습이 완료된 모델이 일관된 예측을 내놓도록 손실 함수에 규제항을 추가하는 아이디어에 기반한다.

Pseudo labeling[8]은 최근 몇 년간 큰 관심을 받은 인기 있는 준지도 학습 기술 중 하나이다. 이 방법은 먼저 라벨이 지정되지 않은 데이터에 대한 모델의 예측을 사용하여 pseudo label을 생성한다. 이렇게 만들어진 pseudo label을 이용하여 라벨이 없는 데이터들로 모델을 학습하는 방법을 의미한다. 이러한 방법은 라벨이 지정된 데이터가 적거나 구하기 어려운 경우에 효과적으로 사용된다.

MixMatch[9]는 라벨이 지정되지 않은 데이터를  $k$ 번 증강시켜 얻은 예측 값들의 평균을 계산하여 temperature scaling 기법으로 sharpening 시킨 값을 soft label로 활용하는 방법으로 데이터들의 entropy를 최소화한다. 이렇게 증강된 라벨이 있는 데이터들과 없는 데이터들을 합치고 섞어 하나의 새로운 데이터셋으로 만든다. 새로운

데이터셋으로 MixUp 계산을 하여 라벨이 지정된 데이터와 지정되지 않은 데이터를 전부 고려하여 학습을 진행하는 학습방법이다.

ReMixMatch[10]는 MixMatch에서 2가지 추가적인 방법으로 학습 성능을 높인 논문이다. 첫 번째 방법은 모델의 예측 평균을 클래스 별로 나눠서 클래스가 적은 데이터셋에 대해서는 정규화된 예측값이 큰 contribution을 주는 distribution alignment이다. 두 번째 방법은 입력 이미지 데이터에 weak과 strong augmentation을 적용한다. weak augmentation을 이용해 증강한 이미지 데이터에 대해 모델이 예측한 pseudo label을 사용하여 strong augmentation을 이용해 나온 이미지 데이터와 entropy를 최소화하여 학습하는 augmentation anchoring을 제시하였다.

하지만 준지도 학습에서는 라벨이 지정된 데이터와 지정되지 않은 데이터를 같이 사용하기 위한 적절한 분배 알고리즘과 클래스 불균형 및 노이즈 데이터 처리 등의 도전 과제가 여전히 존재한다.

본 논문에서는 SAR 이미지 데이터셋의 분류에 대한 준지도 학습의 새로운 방법과 개요를 제안하며, 그에 따른 이점을 소개한다. 더불어, 준지도 학습에서 직면한 도전 과제 중 일부인 클래스 불균형 문제를 해결하기 위한 새로운 방법인 임계값 학습 모듈을 소개하고, 실제 SAR 이미지 데이터셋인 MSTAR 데이터셋 분류 실험을 통해 그 효과를 증명한다.

## 2.2 데이터 증강

준지도 학습은 라벨이 지정된 데이터와 지정되지 않은 데이터 정보를 모두 활용하여 딥러닝 모델 성능을 개선하고 학습에 필요한 라벨이 지정된 데이터 수를 줄일 수 있어 최근에 활발한 연구가 이뤄지고 있다. 준지도 학습을 이용한 딥러닝 모델의 학습 성능을 더욱 향상시키기 위한 방법 중 한 가지는 데이터 증강 기술을 활용하는

것이다. 데이터 증강은 기존의 라벨이 지정된 데이터와 라벨이 지정되지 않은 데이터에 뒤집기나 색 변형, 가리기 같은 변형을 적용하여 새로운 학습 데이터를 생성하는 과정이다.

본 논문에서의 라벨이 없는 데이터에 사용되는 데이터 증강 기술은 Wang의 논문에서 제시된 논문[18]이나 Fixmatch[11]에서 사용되는 weak과 strong augmentation 기법이다. weak augmentation은 데이터를 뒤집는 등의 변형 정도가 약한 간단한 증강 방법으로 원본 이미지와의 차이가 크지 않다. 반대로 strong augmentation에서는 일반적으로 RandAugment[19]나 AutoAugment[20], CTAugment[10] 등의 이미지의 변형 정도가 큰 증강 방법을 사용한다.

다만, 광학 센서로 가시광선을 촬영하는 RGB 이미지 데이터와 달리 SAR 이미지 데이터는 가시광선보다 파장이 긴 마이크로파로 레이더를 이용해 능동적으로 촬영한다. 때문에 SAR 이미지 데이터는 촬영 방법과 전자기파의 특성에 의해 촬영 대상 표면의 거칠기 정도에 따라 전자기파가 산란되어 촬영이 안 되거나 도플러 현상으로 왜곡이 일어난다는 특징이 있다. 이러한 특징 때문에 SAR 이미지 데이터들은 일반적인 strong augmentation 방식이 아닌 다른 방식들을 사용한다. 대표적으로 Lei의 논문[21]에서 나오는 speckle noise나 spatial correlation, background reflectivity 등의 기법을 사용하여 데이터 증강을 하거나 최근 제시되는 딥러닝 모델[22,23]을 이용하여 데이터를 증강하는 방법들이 있다. 본 논문에서는 물체의 경계면에서 전자기파가 산란과 간섭을 일으켜서 생기는 현상인 speckle noise를 사용한다.

### 2.3 동적 임계값

준지도 학습시, pseudo labeling 단계에서 고정된 임계값을 사용하는 논문은 pseudo label[8],

MixMatch[9], ReMixMatch[10], FixMatch[11], SimCLR[12] 등으로 모델이 예측한 정확도가 설정한 임계값을 넘긴 것들만 pseudo labeling을 하여 학습에 반영이 되는 구조를 취하고 있다. 임계값을 너무 높게 설정하면 대부분의 예측 결과들이 임계값을 넘기지 못해 학습되지 못하고 버려져 낭비되는 데이터들이 많아진다. 하지만 임계값을 너무 낮게 설정하면 틀릴 가능성이 높은 낮은 정확도의 예측값도 학습을 하여 모델 학습에 방해가 될 가능성도 높다.

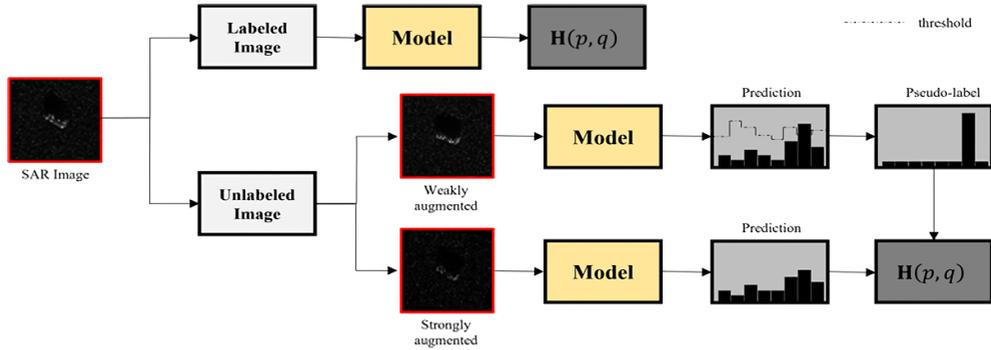
이 문제를 해결하기 위해 학습 도중에 임계값을 변경하여 낭비되는 데이터를 줄이고 학습 능력을 높이기 위한 Dash[24], ADT-SSL[25], FlexMatch[26] 등의 논문들이 있다.

Dash[24]는 loss를 계산하는 과정에서 loss 값을 이용해 정해진 수식대로 계산을 하여 임계값을 조절하여 학습을 진행한다.

ADT-SSL[25]은 학습과정에서 예측과 라벨은 일치하지만 정확도가 임계값 보다 낮아 버려지는 예측 결과들을 사용하여 임계값을 조절하며 학습하는 알고리즘이다.

Dash[24]와 ADT-SSL[25]는 학습 과정에서 임계값이 변동하지만 모든 클래스가 동일한 임계값을 사용한다는 한계가 있다. 그래서 학습 과정에서 값이 변하는 임계값을 사용하면서도 클래스 별로 서로 다른 값을 가지는 FlexMatch[26]가 있다. FlexMatch[26]는 예측 결과와 라벨이 일치하는 개수를 고려하여 클래스 별로 학습 효율을 고려하여 각 클래스마다 서로 다른 임계값을 부여한다.

하지만 이 논문들은 전부 수식을 이용해 수동적으로 계산되는 임계값들이다. 본 논문에서는 임계값 학습 모듈을 학습 모델에 추가하여 각 학습 단계에서 클래스 별로 적절한 임계값을 얻을 수 있다.



〈그림 1〉 네트워크 프레임워크

### III. 제안 방법

#### 3.1 FixMatch

FixMatch[11]는 준지도 학습 방법론들 중에서 최근 발표된 학습 방법으로 RGB 이미지 데이터 분류 분야에서 좋은 성능을 보여준 준지도 학습 방법이다. FixMatch를 기반으로 하는 SAR 이미지 데이터 분류 문제를 해결하기 위해 우리는 먼저 SAR 이미지 데이터셋을 레이블이 지정된 SAR 이미지 데이터셋을  $X = \{(x_b, p_b) : b \in (1, \dots, B)\}$ 와 레이블이 지정되지 않은 SAR 이미지 데이터셋을  $U = \{u_b : b \in (1, \dots, \mu B)\}$ 로 나눈다. Weak augmentation과 strong augmentation은 각각  $\alpha(\cdot)$ 과  $A(\cdot)$ 로 둔다. 먼저 레이블이 지정된 데이터  $X$ 는 지도 학습 손실함수  $l_s$ 를 계산하여 지도 학습을 한다.

$$l_s = \frac{1}{B} \sum_{b=1}^B H(p_b, p_m(y|\alpha(x_b))) \quad (1)$$

다음으로는 준지도 학습 과정을 완료하기 위해 레이블이 지정되지 않은 데이터를 학습하기 위해  $\alpha(u_b)$ 와  $A(u_b)$ 를 하여 데이터를 증강해준다. 본 논문에서 사용하는  $\alpha(\cdot)$ 는 SAR 이미지 데이터를 resize하고 좌우로 뒤집기만 한다.  $A(\cdot)$ 는 기존의 FixMatch에서 RandAugment를 사용하였지만 RGB 이미지 데이터가 아닌 색이 없는 흑백의

SAR 이미지 데이터의 특성에 맞게 RandAugment 대신 speckle noise를 사용한다.

레이블이 지정되지 않은 데이터를 학습하기 위해 레이블이 지정된 데이터셋으로 지도 학습을 완료한 모델로  $\alpha(u_b)$ 의 예측 결과를 계산하여 임계값  $\gamma$ 를 초과한 값들만 pseudo label  $q_b = p_m(y|\alpha(u_b))$ 로 만든다. 그 후, 동일한 모델에  $A(u_b)$ 의 예측 결과와 pseudo label  $q_b$ 를 이용하여 비지도 학습 손실함수  $l_u$ 를 계산해 준지도 학습을 진행한다.

$$l_u = \frac{1}{B} \sum_{b=1}^B (\max(q_b > \gamma)) H(\hat{q}_b, p_m(y|A(u_b))) \quad (2)$$

최종적으로 지도 학습 손실함수  $l_s$ 와 비지도 학습 손실함수  $l_u$ 를 더하여 최종적인 손실함수  $L = l_s + \lambda_u l_u$ 를 최소화시키는 준지도 학습을 한다.  $\lambda_u$ 는 레이블이 지정되지 않은 데이터셋이 준지도 학습의 전체 과정에서 얼마나 영향을 끼칠지 정하는 hyperparameter이다.

그림 1은 전체적인 본 논문에서 제시하는 준지도 학습 방법의 전체적인 흐름을 보여주는 그림이다. 레이블이 존재하는 데이터들은 보통의 지도학습 형태로 학습을 하고, 학습된 모델에 레이블이 지정되지 않은 데이터들을 weak과 strong으로 증강하여 추가 학습을 진행한다. weak augmentation된 데이터들은 모델을 통과하여 예측한

결과값들을 클래스 별로 서로 다른 임계값을 적용하여 pseudo label을 만들고 strong augmentation된 데이터들을 모델에 통과시켜 나온 결과값들과 서로 crossentropy 연산을 하여 학습을 한다.

### 3.2 Data augmentation with speckle noise

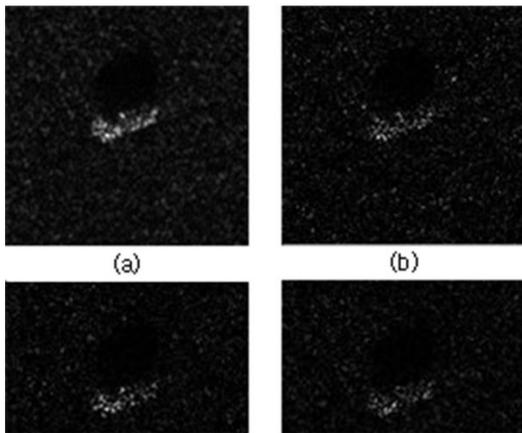
Speckle noise는 SAR 이미지 데이터셋에 적합한 데이터 증강 방법으로 gamma distribution을 이용하여 speckle noise의 강도를 조절한다.

$$N = \frac{\Lambda^\alpha}{\Gamma(\alpha)} n^{\alpha-1} e^{-\Lambda n} \quad (3)$$

$\Gamma$ 는 gamma function으로 파라미터  $\alpha$ 와  $\lambda$ 를 이용하여 speckle noise의 모양과 크기를 조절한다. SAR 이미지 데이터에 speckle noise를 입히는 건 아래의 식과 같다.

$$I_N = I \cdot N \quad (4)$$

(3)과 (4)의 수식을 이용해 우리는 다양한 강도의 speckle noise를 생성하여 이미지에 씌울 수



<그림 2> MSTAR 데이터셋[27]의 “BTR60” 클래스의 예시이다.

(a) 원본, (b) noise level 2, (c) noise level 5, (d) noise level 10으로 증강한 이미지

있다. 그림 2에서 본 논문에서 사용한 speckle noise의 예시를 볼 수 있다.

### 3.3 임계값 학습 모듈

기존의 FixMatch는 라벨이 지정되지 않은 데이터를 학습할 때, 모든 클래스에 동일한 임계값  $\tau$ 를 적용하여 학습을 한다. 모든 클래스에 동일한 임계값을 사용하여 두 가지 문제가 발생한다.

먼저, 학습이 제대로 되지 않은 초기 학습 단계에서부터 높은 임계값을 모든 클래스에 일괄 적용하니 답은 맞지만 정확도가 임계값을 넘기지 못해 학습에 이용되지 못하고 버려지는 많은 데이터들이 존재한다. 두 번째로 학습하는 데이터셋에 특정 클래스가 많거나 적다면 특정 클래스에 대한 정보를 과하게 학습하거나 제대로 학습할 수 없어 학습에 지장이 생기는 데이터 불균형 문제가 생긴다.

우리는 학습 단계마다 각 클래스 별로 서로 다른 임계값을 적용하여 학습하기 위해 임계값 학습 모듈을 추가하여 실험을 하였다.

<표 1> 임계값 학습 모듈의 구조

Layer	Input shape	Output shape
Dense 1	128	64
Batchnorm	64	64
ReLu	64	64
Dense 2	64	10
Sigmoid	10	10
min_max_scale	10	10
Dense 3	10	128
Relu	128	128

표 1이 임계값 학습 모듈의 구조이다. 3개의 Dense layer를 사용하고 1개의 batchnorm layer를 사용한다. 활성화 함수로는 Relu와 sigmoid를 사용하며 두 번째 dense layer와 세 번째 dense layer 사이에 min\_max\_scale layer가 있다. 작동 순서는

먼저 본래의 모델 흐름에서 128의 shape를 지닌 feature가 input으로 들어온다. 그 다음 순서대로 layer들을 통과하며 진행을 하고 min\_max\_scale layer를 지난다. min\_max\_scale layer는 활성화 함수인 sigmoid를 지나간 값들을 적절한 임계값의 범위로 정규화시켜주는 역할을 한다. 그렇게 정규화된 값들을 임계값으로 사용하기 위해 따로 복사를 하고 남은 layer를 통과해 본래의 모델 흐름과 합쳐진다.

본 논문에서는 min\_max\_scale layer의 최소범위를 0.8, 최대범위는 1로 두고 실험을 하였다.

#### IV. 실험 및 결과

##### 4.1 실험

본 논문의 실험에서 사용한 데이터셋은 MSTAR 데이터셋을 사용했다. MSTAR 데이터셋은 탱크와 장갑차, 트럭과 불도저 등을 포함한 10개의 클래스(BMP-2, BRDM-2, BTR-60, BTR-70, T-62, T-72, 2S1, ZSU-234, ZIL-131, D7)로 이루어진 데이터셋으로 SAR 이미지 분류에 많이 사용되는 데이터셋이다. 학습에 사용한 데이터는 전체 3671장 중 20%와 50%만 사용하여 학습하였다.

네트워크 모델은 MSTAR 이미지 데이터를 88 x 88의 크기로 resize하고 speckle noise의 강도는 2로 주어 32의 batch로 100회 학습을 하였다.

실험 환경 구성은 Nvidia 3080 GPU에 Cuda 11.4, pytorch 환경에서 이루어졌다.

##### 4.2 backbone 모델 구조

본 논문에서 기반으로 삼은 FixMatch는 원래 wide ResNet-28-2 모델을 backbone 모델로 삼아 학습을 진행한다. 하지만 SAR 이미지 데이터셋인 MSTAR 데이터셋을 본 논문에서 제시하는 방법으로 학습시키기에는 너무 큰 모델이라 Coman

이 제시한 모델[28]을 backbone으로 한다.

모델은 두 개의 convolution 레이어와 두 개의 dense 레이어, 두 개의 dropout 레이어, 하나의 flatten 레이어와 하나의 max pooling 레이어로 구성되어있다. 모델의 데이터 입력 크기는 (n x n x 1)이다. 입력 데이터는 활성화 함수로 ReLu를 사용하는 3 x 3 크기 커널의 convolution 레이어를 두 번 통과한다. Feature map은 2 x 2 크기의 커널을 가진 max pooling 레이어를 통과하고 dropout 레이어도 통과한다. 마지막으로 flatten 레이어를 통과하고 두 개의 dense 레이어들도 통과한 후, softmax를 거쳐 예측 결과를 계산한다. 표 2에서 사용한 backbone의 구조를 볼 수 있다.

기존의 Fixmatch가 사용하였던 wide ResNet-28-2보다 훨씬 가벼운 모델로도 좋은 성능을 보여준다.

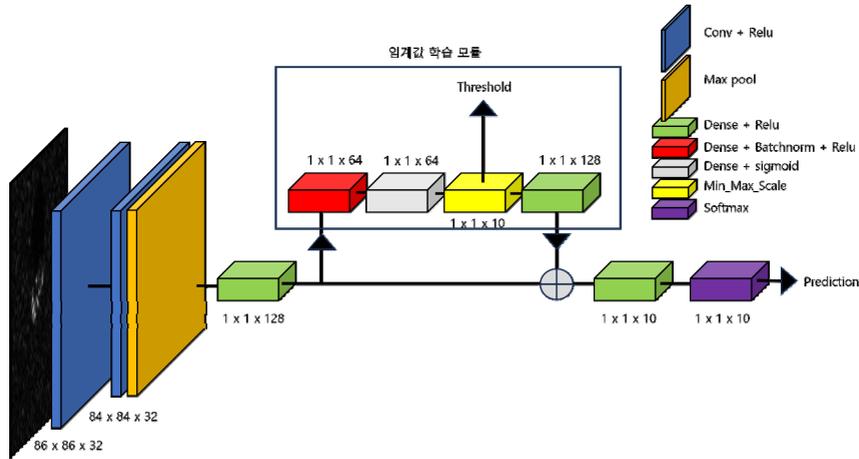
그림 3은 학습에 사용된 backbone 모델에 임계값 학습 모듈이 적용된 모습을 나타낸 그림이다. 첫 번째 dense layer를 통과한 후, 임계값 학습 모듈로 feature를 통과시켜 min\_max\_scale layer를 통과시킨 결과값을 threshold로 사용한다. 그 후, dense layer를 하나 더 통과하여 본래의 학습 흐름에 합쳐준다.

##### 4.3 실험 결과

표 3은 입력 이미지를 88 x 88의 크기로 조정

〈표 2〉 backbone 모델의 구조

Layer	Input shape	Kernel	Output shape
Conv2D	88 x 88 x 1	3 x 3	86 x 86 x 32
Conv2D	86 x 86 x 32	3 x 3	84 x 84 x 32
MaxPool	84 x 84 x 32	2 x 2	42 x 42 x 32
Dropout	42 x 42 x 32	0.25	42 x 42 x 32
Flatten	42 x 42 x 32	1	56448
Dense	56448	1	128
Dropout	128	0.25	128
Dense	128	1	10



〈그림 3〉 backbone과 임계값 학습 모듈 구조도

하고 학습시켰을 때의 결과이다. 전체 데이터 3671장 중 20%의 데이터인 735장을 라벨이 있는 데이터로 사용하고, 나머지 80%를 라벨이 없는 데이터로 사용했다. 50%의 경우는 전체 데이터의 50%인 1836장을 라벨이 있는 데이터로 사용하고, 나머지 50%를 라벨이 없는 데이터로 사용했다. 테스트는 3203장의 test 데이터셋을 사용하여 결과를 얻었다.

88 x 88 크기의 이미지를 아무것도 적용하지 않은 backbone만 사용한 baseline에 20%의 라벨이 있는 데이터를 학습시켰을 때의 결과는 83.51% 이고 50%의 라벨이 있는 데이터를 학습시켰을 때의 결과는 91.99%이다.

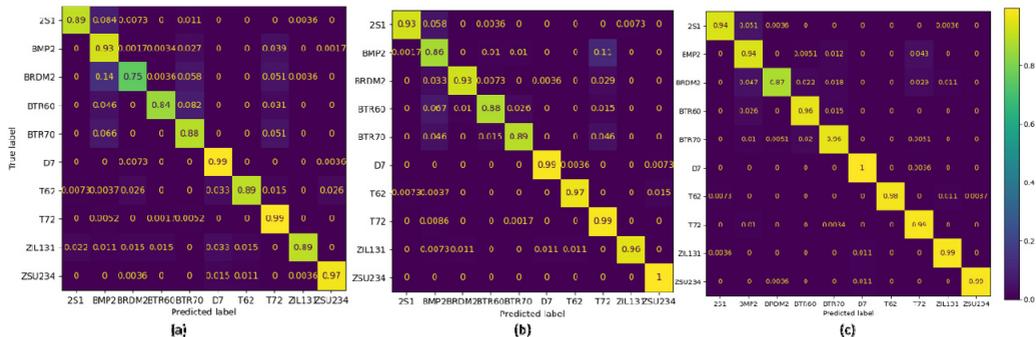
FixMatch만을 적용한 경우는 20%와 50%의 결과가 각각 87.23%와 94.05%이다.

마지막으로 임계값 학습 모듈을 FixMatch에 적용한 경우는 20%와 50%의 결과가 각각 88.14%와 94.86%로 임계값 학습 모듈을 적용하였을 때의 결과가 성능적 우위에 있다.

〈표 3〉 88 x 88 이미지를 input 한 실험 결과

	20%	50%
baseline	83.51%	91.99%
FixMatch	87.23%	94.05%
FixMatch + 임계값 학습 모듈	88.14%	94.86%

그림 4는 88 x 88의 크기로 resize한 MSTAR 데이터를 50%만 사용하였을 때의 결과를 confusion matrix로 나타내었을 때의 모습이다. (a)는



〈그림 4〉 88 x 88 크기의 데이터를 50%만 사용하였을 때의 confusion matrix 결과

baseline의 결과이고 (b)는 FixMatch만을 사용하였을 때의 결과이다. (c)는 FixMatch에 임계값 학습 모듈을 추가하였을 때의 결과이다. (a)부터 ©로 갈수록 오분류의 수가 점차적으로 줄어드는 것을 볼 수 있다.

## V. 결론 및 향후 결과

임계값 학습 모듈을 backbone에 추가하여 학습을 진행한 경우, 학습 과정에 따라 클래스 별로 서로 다른 임계값을 얻어 학습 과정에 좀 더 유연함을 부여해 클래스 간의 데이터 수의 차이 혹은 데이터 자체의 난해함을 해결할 수 있는 것을 볼 수 있다.

기존에도 학습 과정에 서로 다른 임계값을 부여하거나 클래스 별로 서로 다른 임계값을 부여하는 연구들은 존재하였다. 하지만 그 연구들은 전부 수식적으로 임계값들을 구하는 연구들이었고, 이 논문에서 제시한 바와 같이 모델을 학습시키는 과정에서 모델이 자체적으로 임계값을 구하는 연구는 없었기에 임계값도 학습할 수 있다는 사례를 보인 의미 있는 연구 사례라 평가한다.

이번 연구는 SAR 이미지 데이터셋인 MSTAR 데이터셋에 최적화되어있고 임계값 학습 모듈을 하나의 모델에만 시험하였다. 추후 연구는 보다 다양한 모델에 임계값 학습 모듈을 추가하여 여러 데이터셋 별로 최적화된 임계값과 성능 향상을 보일 수 있는지 추가적인 연구가 진행될 수 있을 것이다.

## 참고 문헌

- [1] F, Pierre, et al. "Sharpness-aware minimization for efficiently improving generalization". arXiv preprint arXiv:2010.01412, 2020.
- [2] M, Wortsman, et al. "Model soups: averaging weights of multiple fine-tuned models improves accuracy without increasing inference time." International Conference on Machine Learning. PMLR, 2022.
- [3] H. M. Kabir, "Reduction of Class Activation Uncertainty with Background Information." arXiv preprint arXiv:2305.03238, 2023.
- [4] Z, Yang, et al. "SAR image classification method based on improved capsule network." Journal of Physics: Conference Series. Vol. 1693. No. 1. IOP Publishing, 2020.
- [5] S, Chen, et al. "Target classification using the deep convolutional networks for SAR images." IEEE transactions on geoscience and remote sensing 54.8 4806-4817. 2016
- [6] H, Ren, et al. "Extended convolutional capsule network with application on SAR automatic target recognition." Signal Processing 183 : 108021. 2021
- [7] Q, Xie, et al. "Self-training with noisy student improves imagenet classification." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020.
- [8] E, Arazo, et al. "Pseudo-labeling and confirmation bias in deep semi-supervised learning." 2020 International Joint Conference on Neural Networks (IJCNN). IEEE, 2020.
- [9] D, Berthelot, et al. "Mixmatch: A holistic approach to semi-supervised learning." Advances in neural information processing systems 32, 2019.
- [10] D, Berthelot, et al. "Remixmatch: Semi-supervised learning with distribution alignment and augmentation anchoring." arXiv preprint arXiv:1911.09785, 2019.
- [11] K, Sohn, Kihyuk, et al. "Fixmatch: Simplifying semi-supervised learning with consistency and confidence." Advances in neural information

- processing systems 33 : 596-608. 2020
- [12] T, Chen, et al. "A simple framework for contrastive learning of visual representations." International conference on machine learning. PMLR, 2020.
- [14] A, Anaby-Tavor, et al. "Do not have enough data? Deep learning to the rescue!." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 34. No. 05. 2020.
- [15] B, Zhang, et al. "Censer: Curriculum semi-supervised learning for speech recognition based on self-supervised pre-training." arXiv preprint arXiv:2206.08189, 2022.
- [16] V, Tsouvalas, et al. "Federated self-training for semi-supervised audio recognition." ACM Transactions on Embedded Computing Systems 21.6 : 1-26. 2022
- [17] M, Sajjadi, et al. "Regularization with stochastic transformations and perturbations for deep semi-supervised learning." Advances in neural information processing systems 29, 2016.
- [18] X, Wang, et al. "Contrastive learning with stronger augmentations." IEEE Transactions on Pattern Analysis and Machine Intelligence. 2022.
- [19] E, Cubuk et al. "Randaugment: Practical automated data augmentation with a reduced search space." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops. 2020.
- [20] ED, Cubuk, et al. "Autoaugment: Learning augmentation strategies from data." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019.
- [21] Y, Lei et al. "Synthetic Images Augmentation for Robust SAR Target Recognition." 2021 The 5th International Conference on Video and Image Processing. 2021.
- [22] X, Zhang, et al. "A Novel Data Augmentation Method for SAR Image Target Detection and Recognition." 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS. IEEE, 2021.
- [23] M, Zhang, et al. "Data augmentation method of SAR image dataset." IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium. IEEE, 2018.
- [24] Xu, Yi, et al. "Dash: Semi-supervised learning with dynamic thresholding." International Conference on Machine Learning. PMLR, 2021.
- [25] Liang, Zechen, et al. "ADT-SSL: Adaptive Dual-Threshold for Semi-Supervised Learning." arXiv preprint arXiv:2205.10571 (2022).
- [26] Zhang, Bowen, et al. "Flexmatch: Boosting semi-supervised learning with curriculum pseudo labeling." Advances in Neural Information Processing Systems 34 (2021): 18408-18419.
- [27] ER, Keydel, et al. "MSTAR extended operating conditions: A tutorial." Algorithms for Synthetic Aperture Radar Imagery III 2757, 228-242, 1996.
- [28] C, Coman, "A deep learning SAR target classification experiment on MSTAR dataset." 2018 19th international radar symposium (IRS). IEEE, 2018.

## 저 자 소 개



### 도 재 준(Jae-Jun Do)

- 2022년 2월: 한국항공대학교 소프트웨어학과 (공학사)
- 2022년 3월~현재: 한국항공대학교 인공지능학과 석사과정 <관심분야> 준지도 학습, 컴퓨터 비전, 분류, 인공지능



### 김 선 옥(Sunok Kim)

- 2014년 2월: 연세대학교 전기전자공학과 (공학사)
- 2019년 2월: 연세대학교 전기전자공학과 (공학박사)
- 2021년 3월~현재: 한국항공대학교 인공지능학과 조교수 <관심분야> 준지도 학습, 컴퓨터 비전, 분류, 인공지능