

선별적인 임계값 선택을 이용한 준지도 학습의 SAR 분류 기술

도재준¹⁾ · 유민정¹⁾ · 이재석²⁾ · 문효이²⁾ · 김선옥^{*,1)}¹⁾ 한국항공대학교 인공지능학과²⁾ 한화시스템(주) 레이더연구소Semi-Supervised SAR Image Classification
via Adaptive Threshold SelectionJaejun Do¹⁾ · Minjung Yoo¹⁾ · Jaeseok Lee²⁾ · Hyo Moon²⁾ · Sunok Kim^{*,1)}¹⁾ Department of Artificial Intelligence, Korea Aerospace University, Korea²⁾ Radar Research and Development Center, Hanwha Systems Co., Ltd., Korea

(Received 9 August 2023 / Revised 14 March 2024 / Accepted 15 March 2024)

Abstract

Semi-supervised learning is a good way to train a classification model using a small number of labeled and large number of unlabeled data. We applied semi-supervised learning to a synthetic aperture radar(SAR) image classification model with a limited number of datasets that are difficult to create. To address the previous difficulties, semi-supervised learning uses a model trained with a small amount of labeled data to generate and learn pseudo labels. Besides, a lot of number of papers use a single fixed threshold to create pseudo labels. In this paper, we present a semi-supervised synthetic aperture radar(SAR) image classification method that applies different thresholds for each class instead of all classes sharing a fixed threshold to improve SAR classification performance with a small number of labeled datasets.

Key Words : Semi-Supervised Learning(준지도학습), Synthetic Aperture Radar(합성개구레이더), Classification(분류), Data Augmentation(데이터 증강)

1. 서론

합성 개구 레이더인 SAR 이미지는 지상이 아닌 상공에서 찍은 이미지로 구름이 잔뜩 끼거나 밤에는 촬영이 어려운 일반적인 컬러(RGB) 영상과는 달리 날씨

와 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

* Corresponding author, E-mail: sunok.kim@kau.ac.kr

Copyright © The Korea Institute of Military Science and Technology

이미지 분류 분야도 딥러닝과 인공지능 모델을 적용한 논문^[4]들이 있다. 이 논문들은 딥러닝의 방식 중 하나인 지도학습 방식을 이용한다. 위 방법들은 공통적으로 SAR 이미지에서 유용한 특징 정보들을 모델에 학습시켜 과거의 수식적인 분류 방식보다 큰 성능적 향상을 확인 할 수 있었다. 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 라벨이 없는 데이터를 활용하여 두 모델이 서로를 학습시키는 teacher-student 모델인 noisy student^[7], 라벨이 없는 데이터들의 예측값 중 가장 높은 값을 가상의 라벨로 부여하는 pseudo label^[8], 라벨이 없는 데이터셋에 다양한 학습 기법들을 적용하여 학습을 하는 MixMatch^[9], ReMixMatch^[10], FixMatch^[11], SimCLR^[12] 등이 있다.

본 논문은 SAR 이미지를 데이터셋으로 사용하기에 기존에 제안된 일반적인 RGB 이미지 데이터셋들을 기반으로 하는 데이터 증강 방법들과 달리 SAR 이미지 데이터 분류 모델의 성능을 향상시키기 위해 고안된 데이터 증강 기술인 간섭 잡음(speckle noise)을 사용한다. 그리고 대부분의 준지도 학습법은 라벨이 없는 데이터를 학습 할 때, 고정된 임계값을 사용하여 학습을 진행한다. 그렇기 때문에 초기 학습 단계의 pseudo labeling^[8] 단계에서 잘못 만들어진 pseudo label로 인해 학습 정확도 측면에서 많은 손해를 보며 학습을 진행한다. 본 연구에서는 학습 초기의 unlabeled 데이터의 잘못된 pseudo label의 수를 줄여 학습의 안정성을 올리고자 임계값을 클래스별로 상이하게 선정하여 학습을 진행한다.

2. 관련 연구

2.1 준지도 학습법

준지도 학습은 딥러닝 모델 학습 분야에서 지도 학습을 대신할 딥러닝 모델 학습 방법으로 등장하여, 레

이블이 지정된 데이터와 지정되지 않은 데이터의 정보 모두를 활용하여 모델 성능과 일반화 성능을 개선하는 것을 목표로 한다. 지도 학습과는 달리 준지도 학습 방법론은 레이블이 지정되지 않은 데이터도 학습에 활용하여 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 효과적인 딥러닝 모델 학습 방법으로 사용된다. 그 중, 컴퓨터 비전 분야의 이미지 분류 영역에서 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 FixMatch^[11] 등이 있다.

Consistency regularization은 딥러닝 모델의 일반화와 입력 데이터의 노이즈에 대한 견고성을 개선하는 데 효과적인 성능을 보여 널리 사용되는 기술이다. 데이터를 왜곡시켜 증강한 데이터를 입력으로 주었을 때 학습이 완료된 모델이 일관된 예측을 내놓도록 손실 함수에 규제항을 추가함으로써 딥러닝 모델의 일반화와 정확도를 보장한다.

Pseudo labeling^[8]은 레이블이 지정되지 않은 데이터에 대한 모델의 예측을 사용하여 pseudo label을 생성하고, 이를 이용하여 지도 학습 방식으로 모델을 학습하는 방법을 의미한다. 이러한 방법은 레이블이 지정된 데이터가 적거나 구하기 어려운 경우에 효과적으로 사용된다.

MixMatch^[9]는 레이블이 지정되지 않은 데이터를 k번 증강시켜 얻은 예측 값들의 평균을 계산하여 temperature scaling 기법으로 sharpening 시킨 값을 soft label로 활용하는 방법으로 데이터들의 엔트로피(entropy)를 최소화시킨다. 이렇게 증강된 레이블이 있는 데이터들과 없는 데이터들을 합치고 섞어 하나의 새로운 데이터셋으로 만들어준다. 새로운 데이터셋으로 MixUp 계산을 하여 레이블이 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 약한 데이터 증강과 강한 데이터 증강을 적용한다. 약한 데이터 증강을 이용해 증강한 이미지 데이터에 대해 모델이 예측한 pseudo label을 사용하여 강한 데이터 증강을 이용해 나온 이미지 데이터와 entropy를 최소화 하여 학습하는 augmentation anchoring을 제시하였다.

하지만 준지도 학습에서는 레이블이 지정된 데이터와 지정되지 않은 데이터를 같이 사용하기 위한 적절한 알고리즘과 클래스 불균형 및 노이즈 데이터 처리, 대규모 데이터셋에 적용 가능한 알고리즘 설계 레이블이 지정되지 않은 데이터 학습의 신뢰도 및 안정성 등의 도전 과제가 여전히 존재한다.

본 논문에서는 준지도 학습에서 직면한 도전 과제 중 일부인 레이블이 지정되지 않은 데이터에 대한 pseudo labeling 과정에서 클래스별로 동일한 임계값을 사용하였을 때 야기되는 문제를 해결하고자 한다. pseudo label 학습의 신뢰도 및 안정성 문제를 해결하기 위한 새로운 기술인 클래스 별 서로 다른 임계값 사용을 제안하고, 실제 SAR 이미지 데이터셋인 MSTAR 데이터셋 실험을 통해 그 효과를 증명한다.

2.2 데이터 증강 방법

준지도 학습은 레이블이 지정된 데이터와 지정되지 않은 데이터 정보를 모두 활용하여 딥러닝 모델 성능을 개선하고 학습에 필요한 레이블이 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 기술의 활용이다. 데이터 증강은 기존의 레이블이 지정된 데이터와 레이블이 지정되지 않은 데이터에 뒤집기나 가리기 같은 변형을 적용하여 새로운 학습 데이터를 생성하는 과정이다.

본 논문에서의 레이블이 없는 데이터에 사용되는 데이터 증강 기술은 X. Wang의 논문에서 제시된 방법^[8]이나 FixMatch^[11]에서 사용되는 약한 데이터 증강과 강한 데이터 증강 기법이다. 약한 데이터 증강은 데이터를 뒤집는 등의 변형 정도가 약한 간단한 증강 방법으로 원본 이미지와의 차이가 크지 않다. 반대로 강한 데이터 증강에서는 일반적으로 RandAugment^[19]나 AutoAugment^[20], CTAugment^[10] 등의 이미지의 변형 정도가 큰 증강 방법을 사용한다.

다만, 광학 센서로 가시광선을 촬영하는 RGB 이미

지 데이터와 달리 SAR 이미지 데이터는 가시광선 보다 파장이 긴 마이크로파로 레이더를 이용해 능동적으로 촬영하기 때문에 SAR 이미지 데이터는 촬영 방법과 전자기파의 특성에 의해 촬영 대상의 표면 거칠기 때문에 전자기파가 산란되어 촬영이 안 되거나 도플러 현상으로 왜곡이 일어난다는 특징이 있다. 이러한 특징 때문에 SAR 이미지 데이터들은 일반적인 강한 데이터 증강 방식이 아닌 다른 방식들을 사용한다. 대표적으로 Y. Lei의 논문^[21]에서 나오는 spatial correlation, background reflectivity 등의 기법을 사용하여 데이터 증강을 하거나 noise의 일종인 speckle noise를 제거하여 SAR 이미지 데이터의 잡음을 제거하거나^[22], 딥러닝 모델^[23,24]을 이용하여 데이터를 생성하여 증강하는 방법들이 있다. 최근에는 speckle noise를 추가하는 학습을 시켜 speckle noise를 제거하는 복합적인 방법^[25]처럼 다양한 방법들이 제시되고 있다. 본 논문에서는 일반적으로 생성되는 SAR 영상의 경우 speckle noise가 존재한다고 가정하고 speckle noise를 추가하여 증강하는 방법을 사용한다.

3. FixMatch기반 클래스 별 임계값 설정 학습 모델

3.1 FixMatch 프레임워크

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. 레이블이 있는 데이터셋은 X 는 B 개의 데이터 x_b 와 라벨 p_b 로 구성되어있다. 레이블이 지정되지 않은 데이터셋 U 는 μB 개의 데이터 u_b 로 구성되어있다. 여기서 μ 는 X 와 U 의 수를 결정하는 hyperparameter이다. 약한 데이터 증강과 강한 데이터 증강은 각각 $\alpha(\cdot)$ 과 $A(\cdot)$ 로 둔다. 먼저 레이블이 지정된 데이터 X 는 지도 학습 손실함수 l_s 를 계산하여 지도 학습을 한다. 손실함수 l_s 에 사용하는 H는 cross entropy 공식을 따른다. p_m 은 학습 모델의 데이터 x_b 에 대한 예측결과를 의미한다.

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

다음으로는 준지도 학습 과정을 완료하기 위해 레이블이 지정되지 않은 데이터를 학습하기 위해 $\alpha(u_b)$ 와 $A(u_b)$ 를 하여 데이터를 증강 해준다. 본 논문에서 사용하는 $\alpha(\cdot)$ 에서 데이터를 증강하는 방식은 기존의 FixMatch가 RGB 데이터셋에서 사용하였던 여러 증강 방법들 중 일부를 선정하여 사용하였다. 다양한 실험을 통해 가장 좋은 성능을 보인 resize와 random horizontal flip을 사용하였다. 여기서의 resize는 선형보

간법을 이용하여 이미지의 크기를 조절하는 방식으로 서로 다른 크기의 SAR 이미지들을 동일한 크기로 통일 시켜 학습이 가능하게 해준다. $A(\cdot)$ 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 τ 를 초과한 값들만 pseudo label $q_b = p_m(y|\alpha(u_b))$ 로 만든다.

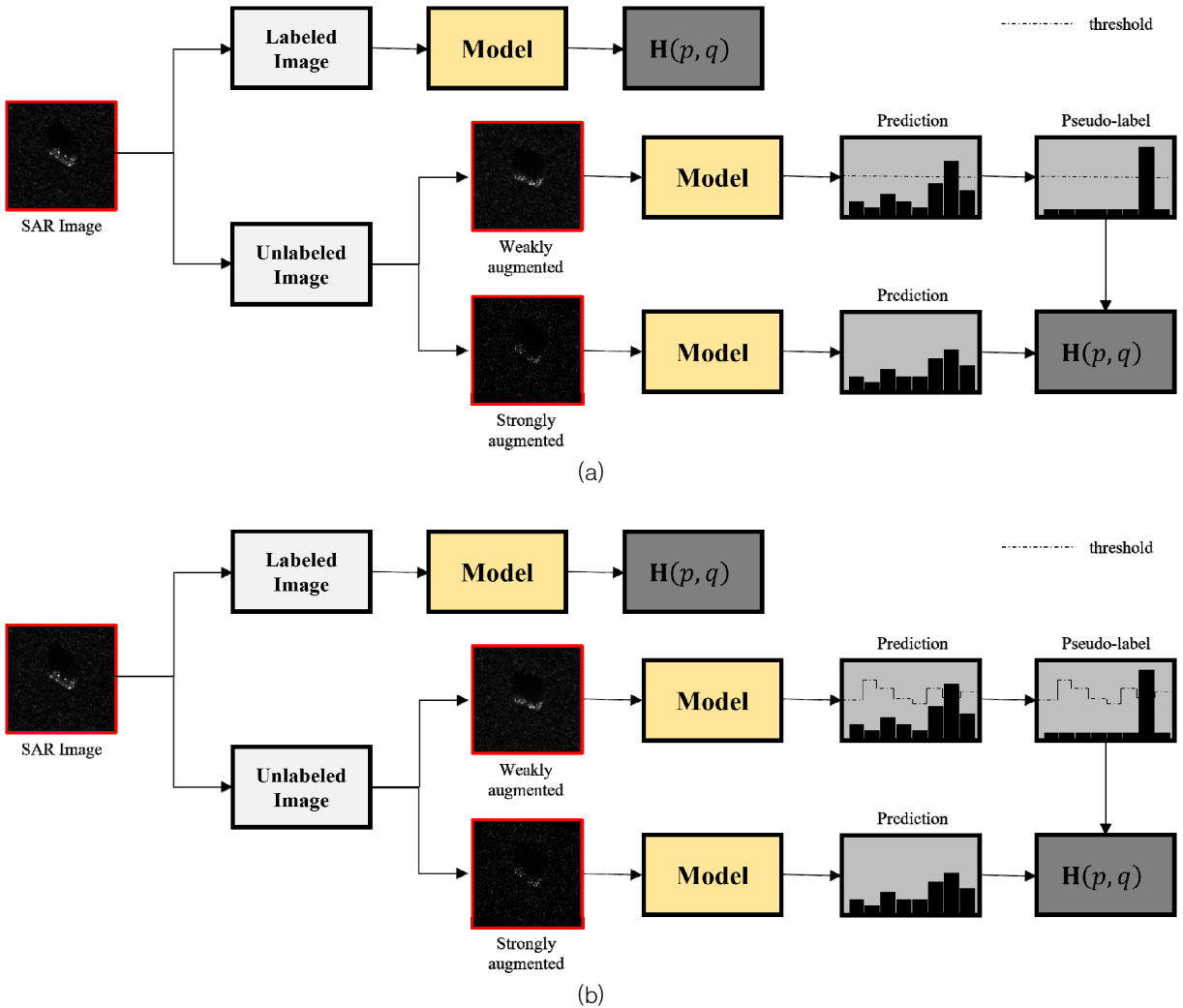


Fig. 1. (a) Overall framework of FixMatch networks, (b) Overall framework of proposed networks

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.

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

최종적인 손실함수는 지도 학습 손실함수 l_s 와 비지도 학습 손실함수 l_u 를 더하여 $L = l_s + \lambda_u l_u$ 가 되어 최소화 시키는 준지도 학습을 한다. λ_u 는 레이블이 지정되지 않은 데이터셋이 준지도 학습의 전체 과정에서 얼마나 영향을 끼칠지 결정하는 hyperparameter로 0 이상의 양수값을 가진다. λ_u 값에 따라 비지도 학습 손실함수 l_u 이 학습에서 가지는 가중치가 달라지며, 본 논문에서는 레이블이 지정되지 않은 데이터셋의 학습을 온전하게 반영하기 위해 λ_u 의 값을 1로 설정하였다.

Fig. 1. (a)는 FixMatch의 전체적인 흐름을 나타낸 그림이고 Fig. 1. (b)는 본 논문에서 제시하는 FixMatch 기반의 준지도 학습 방법의 전체적인 흐름을 보여주는 그림이다. 레이블이 존재하는 데이터들은 보통의 지도학습 형태로 학습을 하고, 학습된 모델에 레이블이 지정되지 않은 데이터들을 약한 데이터 증강과 강한 데이터 증강 방법으로 데이터들을 증강하여 추가 학습을 진행한다. 약한 데이터 증강이 된 데이터들은 모델을 통과하여 예측한 결과값들로 pseudo label을 만드는데 Fig. 1. (b)에서 클래스 별로 서로 다른 임계값을 적용하여 pseudo label을 만들고 강한 데이터 증강이 된 데이터들을 모델에 통과시켜 나온 결과값들과 서로 crossentropy 연산을 하여 학습을 한다.

3.2 Data augmentation with speckle noise

SAR 이미지 데이터에 추가되는 Speckle noise는 SAR 이미지 데이터셋에 적합한 데이터 증강 방법이다. 본 논문에서 제시하는 speckle noise를 추가하는 방법은 K. Choi^[25]에서 제시된 개념으로 speckle noise를 SAR 이미지에 추가할 수 있는 간단한 방법이다. speckle noise는 데이터가 증강되는 distribution을 넓혀 학습 모델을 일반화 시킬 수 있다. Speckle noise는 gamma distribution을 이용하여 speckle noise N 의 강도를 조절한다.

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

Γ 는 gamma function으로 speckle noise의 모양과 크기를 조절하는 파라미터인 α 와 λ 를 이용하여 speckle noise의 모양과 크기를 조절한다. n 은 양수의 값을 가지는 상수이다. SAR 이미지 데이터에 speckle noise를 추가하는 방법은 아래의 식과 같다.

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

(3)과 (4)의 수식을 이용해 다양한 강도의 speckle noise를 가지는 이미지를 생성할 수 있다. 식 (4)에서 I 는 SAR 이미지로 0~1 사이의 값을 가지는 행렬이다. I 에 gamma distribution 기반의 speckle noise N 을 곱하여 최종적인 이미지 I_N 을 얻을 수 있다. Fig. 2에서는 다양한 강도의 speckle noise를 추가하여 생성한 예시 이미지를 볼 수 있다. Speckle noise의 강도가 강해지면 강해질수록 모델이 데이터를 학습하는데 방해가 되어 성능이 오히려 떨어지는 것을 확인하였다. 본 논문에서는 α 와 λ 를 가장 적절한 강도인 2와 0.5로 사용하였다.

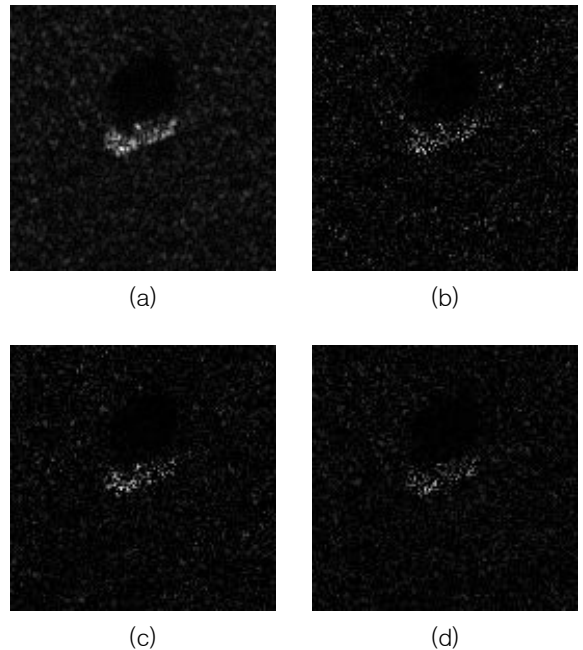


Fig. 2. Example of class “BTR60” in MSTAR dataset^[26]. (a) Original image and augmented images with (b) noise level 2, (c) noise level 5, and (d) noise level 10

3.3 클래스 별 임계값

기존의 FixMatch는 레이블이 지정되지 않은 데이터를 학습 할 때, 모든 클래스에 동일한 임계값 τ 를 적용하여 학습을 한다. 모든 클래스에 동일한 임계값을 사용하여 두 가지 문제가 발생한다.

먼저, 학습이 제대로 되지 않은 초기 학습 단계에서부터 높은 수준의 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 학습 할 수 없어 학습에 지장이 생기는 데이터 불균형 문제가 생긴다. 이 연구에서는 클래스 별 동일한 임계값을 사용함으로써 인해 pseudo label에 불균형이 생겨 레이블이 없는 데이터를 학습 할 때, 특정 클래스로 학습이 편중되는 것을 의미한다.

우리는 Table 1처럼 각 클래스 별로 서로 다른 임계값을 적용하여 상기한 두 가지 문제를 해결하였다.

Table 1. Adaptive thresholds by class

2S1	BMP 2	BRD M2	BTR 60	BTR 70	D7	T62	T72	ZIL1 31	ZSU 234
0.95	0.99	0.99	0.99	0.99	0.95	0.98	0.98	0.98	0.95

기존의 FixMatch는 RGB 이미지 데이터셋에 사용되는 준지도 학습 방법으로 원래 0.95의 고정된 임계값을 가진다. FixMatch를 SAR 이미지 데이터 분류에 적합하게 FixMatch를 적용한 Fig. 1. (a)의 모델을 만드는 과정에서, 여러 임계값을 실험하며 임계값을 0.98로 설정할 때 가장 좋은 결과를 보이는 것을 확인하였다. Fig. 1. (a)의 모델로 실험을 진행하며 레이블이 없는 데이터를 증강한 $\alpha(u_b)$ 와 $A(u_b)$ 의 예측 결과와 pseudo label을 비교한 추이를 확인하였다. 확인 결과 예측 결과가 동일한 비율이 높은 클래스들은 Fig. 3. (c)의 경향이 나타나는 것을 확인하였다. 반대로, 정답률이 낮은 클래스들은 Fig. 3. (a)의 경향이 나타나는 것을 확인하였다. 이런 이유로 잘못된 pseudo label이 생겨 전체적인 모델 학습에 악영향을 끼친다. 위의 문제를 해결하기 위해 우리는 클래스 별로 서로 다른 임계값들을 적용하였다. 클래스 별로 서로 다른 임계값은 정답률이 낮은 4개의 클래스들과 높은 3개의 클래스를 골라내어 $\alpha(u_b)$ 와 $A(u_b)$ 의 예측 결과가 많이 다른 클래스에는 높은 임계값인 0.99를 적용하여 Fig. 3. (b)처럼 틀린 pseudo label을 줄였다. $\alpha(u_b)$ 와 $A(u_b)$ 의 예측 결과가 많이 같은 클래스에는 낮은 임계값인 0.95를 주어 정확한 pseudo label이 생길도록 유도하였다.

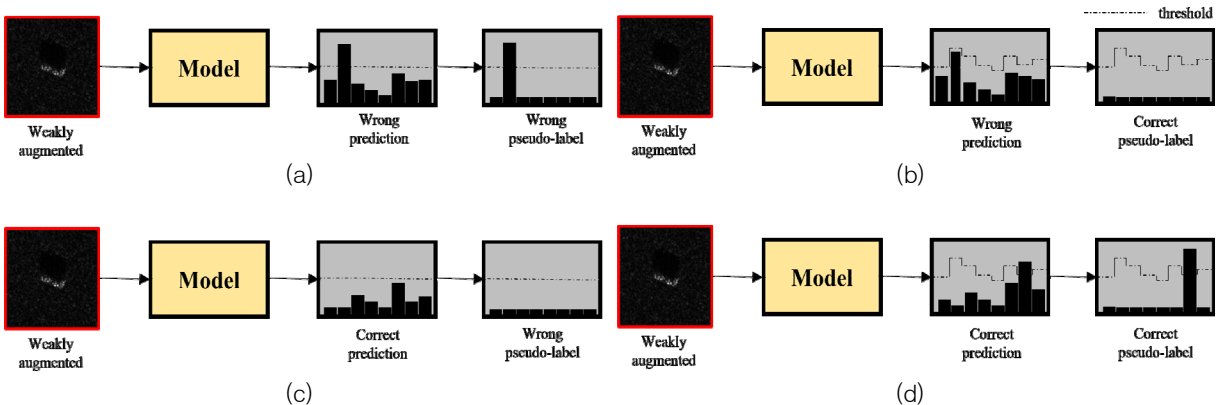


Fig. 3. Effectiveness of the proposed framework. (a) shows wrong prediction with confidence over the fixed threshold in FixMatch frameworks lead to wrong pseudo label, (b) shows wrong prediction with confidence under the threshold in the proposed framework lead to correct pseudo label, (c) shows correct prediction with confidence under the threshold in FixMatch framework lead to wrong pseudo label, and (d) shows correct prediction with confidence over the threshold in the proposed framework lead to correct pseudo label

위와 같이 클래스별 서로 다른 임계값을 통해 더 정확한 학습을 할 수 있도록 하였다.

Fig. 3에서 예시를 볼 수 있다. (a)와 (c)에서 볼 수 있듯이 모든 클래스들이 동일한 임계값을 가진 경우는 학습 모델이 틀린 예측을 했지만 임계값을 넘는 정확도(confidence)나 맞은 예측을 했지만 임계값을 넘지 못하는 confidence로 인해 잘못된 pseudo label들이 생성되어 모델의 학습에 방해가 된다. 하지만 (b)에서 틀린 예측의 confidence가 높아 해당 클래스의 임계값이 높아 pseudo label로 사용되지 않는 것을 볼 수 있다. 또한 (d)에서 낮은 confidence의 옳은 예측은 낮은 임계값 덕분에 pseudo label로서 올바르게 기능하는 것을 확인 할 수 있다.

4. 실험 및 결과

4.1 실험 환경 및 데이터셋

본 논문의 실험에서 사용한 데이터셋은 MSTAR 데이터셋으로 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 SAR 이미지 분류에 많이 사용되는 데이터셋이다. 학습에 사용한 데이터는 전체 3671장 중 데이터 수에 따라 학습 편차가 생기지 않도록 각 클래스 별로 200개씩 골라내어 클래스 간 데이터 수를 맞춰 총 2000장 중 10%, 20%, 50%만 레이블이 있는 데이터로 사용하였고 나머지 90%, 80%, 50%는 레이블이 없는 데이터로 사용하여 학습하였다. 각 클래스 별로 수를 통일시켜 데이터셋 자체의 균형을 보장하여, 레이블이 없는 데이터셋을 임계값을 넘는 pseudo label들만 사용하며 생기는 클래스 불균형의 영향을 확인 할 수 있다.

레이블이 있는 데이터와 없는 데이터는 랜덤하게 추출하였다.

테스트 데이터셋은 클래스당 150장씩 랜덤하게 추출하여 만든 1500장의 고정된 데이터셋으로 테스트하여 실험 결과들을 얻었다.

네트워크 모델은 MSTAR 이미지 데이터를 각각 44×44 , 88×88 의 크기로 resize하고 speckle noise의 강도를 2로 주어 32의 batch로 100회 학습을 한 세트 총 10번 학습하여 Table 3과 4에 평균(표준편차)의 형태로 나타내었다.

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 모델^[27]을 backbone 모델로 삼아 학습을 진행한다. 하지만 SAR 이미지 데이터셋인 MSTAR 데이터셋을 본 논문에서 제시하는 방법으로 학습시키기에 너무 큰 모델이라 C. Coman이 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 구성 되어있다. 모델의 데이터 입력 크기는 $(n \times n \times 1)$ 이다. 입력 데이터는 활성화 함수로 ReLu를 사용하는 3×3 크기 커널의 convolution 레이어를 두 번 통과 한다. Feature map은 2×2 크기의 커널을 가진 max pooling 레이어를 통과하고 dropout 레이어도 통과한다. 마지막으로 flatten 레이어를 통과하고 두 개의 dense 레이어들도 통과 한 후, softmax를 거쳐 예측 결과를 계산한다.

Table 2. Backbone network architecture

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

Table 2에서 사용한 backbone의 구조를 볼 수 있다. 기존의 FixMatch가 사용하였던 wide ResNet-28-2보다 훨씬 가벼운 모델로도 좋은 성능을 보여준다.

4.3 실험 결과

Table 3, 4는 입력 이미지를 각각 44×44 , 88×88 의 크기로 조정하고 학습 시켰을 때의 결과이다.

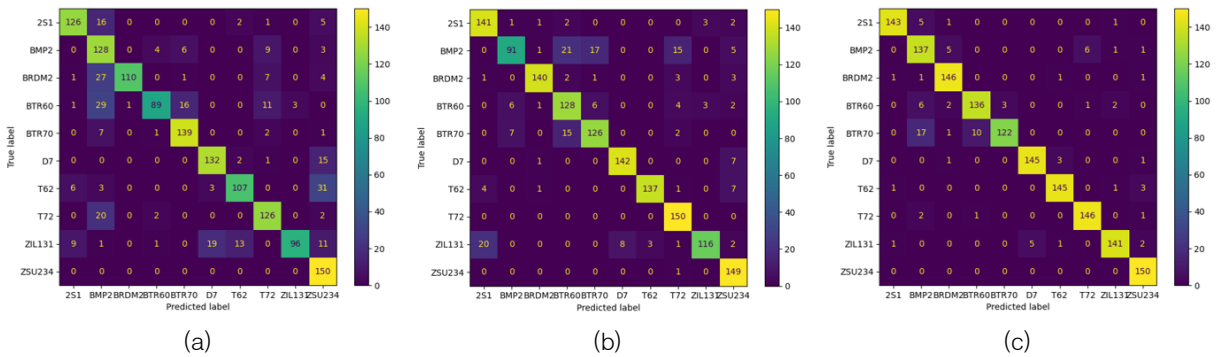


Fig. 4. Result of 44×44 size confusion matrix of adaptive threshold. (a) Result when only 10 % of the data is used, (b) result when only 20 % of the data is used, and (c) result when only 50 % of the data is used

Table 3. Comparison between baseline model, FixMatch model and our proposed model in image size 44×44

	10 %	20 %	50 %
baseline	73.63 % (1.88)	83.94 % (1.07)	91.09 % (0.90)
FixMatch	75.81 % (1.67)	85.95 % (1.67)	93.19 % (0.76)
FixMatch + adaptive ths	76.23 % (2.29)	86.71 % (0.86)	93.25 % (0.62)

Table 4. Comparison between baseline model, FixMatch model and our proposed model in image size 88×88

	10 %	20 %	50 %
baseline	72.16 % (2.55)	83.51 % (1.61)	91.99 % (0.91)
FixMatch	72.57 % (2.28)	85.79 % (1.69)	94.87 % (0.48)
FixMatch + adaptive ths	77.05 % (2.38)	87.99 % (1.39)	95.04 % (1.07)

Baseline은 C. Coman이 제시한 모델^[27]만으로 레이블이 있는 데이터들만을 사용해 학습한 결과이다. FixMatch는 기존의 FixMatch를 SAR 이미지를 분류하기 적합하게 적용한 Fig. 1. (a)의 모델의 결과이다. FixMatch + adaptive ths는 Fig. 1. (b)의 모델로 Table 1의 임계값을 적용한 결과이다.

입력 이미지의 크기, 사용한 데이터의 양과 상관없이 대부분의 모든 경우에서 baseline은 물론 FixMatch와 비교시에도 성능적 우위를 보인다.

Fig. 4는 44×44 크기의 입력 이미지로 클래스 별로 서로 다른 임계값을 주었을 때, 실험 결과의 confusion matrix이다. 사용한 데이터의 크기와 양이 많아지면 많아질수록 정답을 맞추는 정확도가 높아지는 것을 확인 할 수 있다.

5. 결론

본 논문은 SAR 이미지 데이터셋 분류 학습 모델에 준지도 학습 방법론을 적용하여 학습시키는 새로운 방법을 제안하였다. RGB 데이터셋을 사용하여 우수한 성능을 증명한 FixMatch를 SAR 이미지 데이터셋에 적용하여 FixMatch만으로도 SAR 이미지 데이터셋 분류 성능 향상에 도움이 된다는 것을 검증하였다. 이미지 분류 성능을 더욱 높이기 위해 SAR 이미지 데이터셋의 특성에 어울리는 데이터 증강 방식인 speckle noise를 사용하고, 각 클래스 별로 서로 다른 임계값을 주어 기존의 FixMatch보다 훨씬 더 높은 정확도를 보여주었다.

후 기

이 연구는 2022년 정부(방위사업청)의 재원으로 국방 과학연구소의 지원을 받아 수행된 미래도전국방기술 연구개발사업임(No. 915029201).

References

- [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.
- [13] R, Shams, “Semi-supervised classification for natural language processing,” arXiv preprint arXiv:1409.7612, 2014.
- [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] C, Hyunho, and J, Jechang, et al., "Speckle noise reduction technique for SAR images using statistical characteristics of speckle noise and discrete wavelet transform," *Remote Sensing* 11.10 (2019): 1184.
- [23] 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.
- [24] M, Zhang, et al., "Data augmentation method of SAR image dataset," IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium, IEEE, 2018.
- [25] K, Choi, et al., "Deep Cascade Network for Noise-Robust SAR Ship Detection With Label Augmentation," in *IEEE Geoscience and Remote Sensing Letters*, Vol. 19, pp. 1-5, 2022, Art No. 4514005, doi: 10.1109/LGRS.2022.3205715.
- [26] ER, Keydel, et al., "MSTAR extended operating conditions: A tutorial," *Algorithms for Synthetic Aperture Radar Imagery III* 2757, 228-242, 1996.
- [27] S, Zagoruykond Nikos Komodakis, "Wide residual networks," arXiv preprint arXiv:1605.07146, 2016.
- [28] C, Coman, "A deep learning SAR target classification experiment on MSTAR dataset," 2018 19th international radar symposium(IRS), IEEE, 2018.