

# F-guessed 방식과 준지도 학습을 이용한 화재 데이터 생성 및 인식률 개선 연구

김종식\*, 강대성\*\*

## A Study on Fire Data Generation and Recognition Rate Improvement using F-guessed and Semi-supervised Learning

Jong-Sik Kim\*, Dae-Seong Kang\*\*

---

본 연구는 2022년도 한국통신사업자연합회 재원으로 IT 협동연구센터 지원받아 수행한 연구과제입니다(No. 2022-0069)

---

### 요 약

현재의 준지도 학습 방식은 적은 양의 정답라벨(Labeled data)을 이용하여 정답이 없는 라벨(Unlabeled data)을 증강한 후 예측된 라벨에 대해서 Mixup 등의 방식을 활용하여 정답이 없는 라벨의 손실을 최소화하는 방식으로 인식률을 개선해왔다. 본 논문에서는 화재와 같이 정답라벨 수집이 제한된 경우에 정답라벨의 분포가 전체를 커버하지 못할 수도 있기 때문에 Fine-tuning 방식을 이용하여 화재 자동 데이터 생성 및 인식률 개선을 위한 방법을 제안한다. 인식률 개선을 위해서 F-guessed 방식과 중복적 라벨링 방식을 채용하여 학습 데이터 결과와 예측 결과값이 유사하도록 데이터를 처리하였다. 그 결과 초기 대비 자동 데이터 생성은 5,565 개에서 35,322개로 약 6.3배 정도 증가하였으며, mAP@0.5는 65.9%에서 82.5%로 약 16.6% 정도의 개선된 결과를 얻었다.

### Abstract

The current semi-supervised learning method improved the recognition rate by increasing the Unlabeled data using a small amount of Labeled data and minimizing the loss of Unlabeled data by using a method such as Mixup for the predicted label. This paper proposes a new method for automatic fire data generation and improvement of recognition rate using the Fine-tuning method because the distribution of labeled data may not cover the whole when the collection of labeled data is limited, such as fire. In order to improve the recognition rate, the F-guessed method and the redundant labeling method were adopted to process the data so that the learning data results and the predicted results were similar. As a result, automatic data generation increased by about 6.3 times from 5,565 to 35,322, and mAP@0.5 improved by about 16.6% from 65.9% to 82.5%.

### Keywords

semi-supervised learning, deep learning, fine-tuning, pseudo labeling

---

\* 동아대학교 전자공학과 박사과정  
- ORCID: <https://orcid.org/0000-0002-1459-1943>  
\*\* 동아대학교 전자공학과 교수(교신저자)  
- ORCID: <https://orcid.org/0000-0003-0186-2430>

• Received: Oct. 18, 2022, Revised: Nov. 25, 2022, Accepted: Nov. 28, 2022  
• Corresponding Author: Dae-Seong Kang  
Dept. of Dong-A University, 37 NaKdong-Daero 550, Beon-gil Saha-gu,  
Busan, Korea  
Tel.: +82-51-200-7710, Email: [dskang@dau.ac.kr](mailto:dskang@dau.ac.kr)

## 1. 서론

컴퓨터비전을 위하여 준지도 학습(Semi-supervised learning) 방법은 지난 몇 년 동안 빠르게 발전했다. 현재의 최첨단 방법은 아키텍처 및 손실 함수 측면에서 이전 작업을 단순화하거나 다른 공식을 혼합하여 하이브리드 방법을 도입하고 있다[1]. 그렇지만 아직 현재 딥러닝의 가장 대표적인 방법론은 지도학습(Supervised learning)이다. 하지만 지도학습은 어쩌면 학습 데이터의 패턴을 외우는 학습 방법이다. 그러므로 한 번도 학습하지 않은 데이터에 대해서는 인식하기는 쉽지 않다. 일반화가 잘되기 위해서는 반드시 더 많은 정답라벨(Labeled data)이 요구된다. 성공적으로 딥러닝을 도입한 이미지 분야의 경우도 대용량 정답라벨 이미지를 확보했기 때문에 현재 좋은 성능을 발휘했다고 볼 수 있다. 하지만 정답라벨을 확보하기 어려운 분야들도 상당히 많이 존재하는 게 현실이다.

그래서 많은 연구자가 생각해낸 방법이 적은 정답라벨 데이터만으로도 라벨 예측이 가능한 준지도 학습 방법을 연구하였다[2]. 그중 Pseudo-label 방식은 소량의 정답라벨에서 학습된 모델의 예측값을 기반으로 정답이 없는 라벨에 예측 라벨을 붙이는 방식이며[2], 최근 구글의 Berthelot 등 연구자들은 MixMatch 논문을 통하여 정답 이미지와 정답이 없는 이미지의 Mixup과 평균 이후 Sharpen을 통하여 좋은 결과를 도출했다[2][8]. 그리고 FixMatch 논문

에서는 한 이미지를 약하게 또는 강하게 변형하여 인식을 차이를 줄이는 방식으로 Pseudo-label 방식을 적용하였다[2][9].

하지만 화재와 같이 정답라벨 데이터가 제한된 경우에는 기존 방식을 적용하기에는 어려움이 있었다. 즉 수집된 화재 정답라벨 데이터의 분포가 전체를 커버하지 못할 수도 있기 때문에 학습 데이터에는 없는 새로운 테스트 데이터가 들어왔을 경우 오인식을 하는 경우가 많았다. 그래서 본 논문에서는 기존 Pseudo-label 방식에 정답이 없는 라벨 데이터를 Fine-tuning 학습을 통하여 세밀히 분류하고, 반복적 예측을 통하여 예측한 확률 분포를 줄이는 방식으로 화재 인식을 개선하기 위한 방식을 제안한다 [3][4]. 그림 1은 Fine-tuning과 준지도 학습을 이용한 화재 데이터 생성 개념도이다.

## II. 관련 이론

### 2.1 준지도 학습의 이해

적은 정답라벨 데이터가 있으면서 추가로 활용할 수 있는 많은 정답이 없는 라벨 데이터가 있다면, 준지도 학습을 고려할 수 있다. 준지도 학습은 소량의 정답라벨에는 지도학습을 적용하고, 많은 양의 정답이 없는 라벨에는 비지도 학습(Unsupervised learning)을 적용해 추가적인 성능향상을 목표로 하는 학습 방법이다.

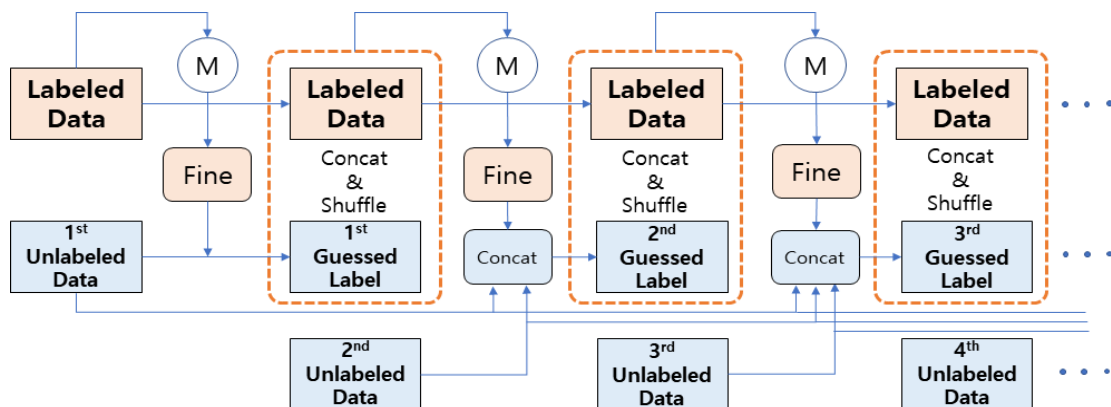


그림 1. Fine-tuning과 준지도 학습을 이용한 화재 데이터 생성 개념도

Fig. 1. Conceptual diagram of fire data generation using fine-tuning and semi-supervised learning

준지도 학습은 바로 정답 데이터를 수집하고, 라벨링' 작업에 드는 많은 자원과 비용을 줄이기 위해 등장하게 되었다. 준지도 학습은 라벨의 정답을 맞히는 모델에서 벗어나 데이터 자체의 본질적인 특성이 모델링 된다면 소량의 정답라벨 데이터를 통해 적은 양의 학습으로도 일반화 성능을 끌어올릴 수 있다는 것이다.

준지도 학습의 목적함수는 지도학습 손실  $\mathcal{L}_s$ 와 비 지도학습 손실  $\mathcal{L}_u$ 의 합을 식 (1)과 같이 최소화하는 것으로 표현할 수 있다. 즉, 지도학습과 비 지도학습을 1-stage로 한 번에 학습한다. 이것이 2-stage로 이루어지는 자기 지도학습(Self-supervised learning)과 전이학습(Transfer learning) 등과의 차이점이다[5].

$$\mathcal{L}_{oss} = \mathcal{L}_S + \mathcal{L}_U \quad (1)$$

정답이 없는 라벨을 어떻게 학습에 사용할 것이냐의 관점에서 다양한 준지도 학습의 방법론들이 등장하였다. 그중 현재 제안한 기술과 유사한 Pseudo labeling 방식과 MixMatch 그리고 FixMatch에 대해 먼저 알아보자.

① Pseudo labeling은 지도학습을 통해 일차적으로 학습된 모델을 이용하여, 라벨링 되지 않은 데이터에 대해 예측을 수행한다.

수행된 예측 결과를 이용해 가짜(Pseudo)로 라벨링 한다고 하여 Pseudo labeling이라고 한다. 따라서 Pseudo labeling을 수행하기 위해서는 학습된 모델이 있어야 하고, 정답이 없는 라벨 데이터가 있어야 한다. 정답이 없는 라벨에 대해서 Pseudo labeling을 한 후에 이 확장된 데이터 세트를 이용하여 2차 학습을 수행하는 방식이다. 그림 2에 Pseudo labeling 수행방식을 그림으로 표시하였다[1][6].

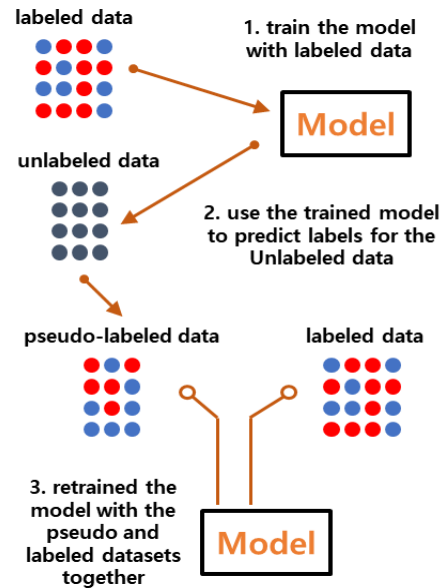


그림 2. Pseudo labeling 방식  
Fig. 2. Pseudo labeling operation

② MixMatch는 정답라벨 데이터(X)와 정답이 없는 라벨 데이터(U)를 주면, 처리된 정답라벨 샘플 ( $X'$ )와 예측 정답라벨(Guessed labeled)( $U'$ )을 생성한다. 그리고 공식적으로 준지도 학습에 대한 결합 손실  $\mathcal{L}$ 은 다음 식 (2)와 같이 정의된다[7][8].

$$\begin{aligned} \chi', u' &= MixMatch(\chi, u, T, K, \alpha) \quad (2) \\ \mathcal{L}_\chi &= \frac{1}{|\chi'|} \sum_{x, p \in \chi'} H(p, P_{model}(y_{vertx}; \theta)) \\ \mathcal{L}_u &= \frac{1}{|Uvertu'|} \sum_{u, q \in u'} \|q - P_{model}(y_{vertu}; \theta)\|_2^2 \\ \mathcal{L} &= \mathcal{L}_\chi + \lambda_u \mathcal{L}_u \end{aligned}$$

$H(p, q)$ 는 분포  $p$ 와  $q$  사이의 교차 엔트로피이고,  $T, K, \alpha, \lambda_u$ 는 하이퍼 파라미터이다. MixMatch는 Entropy minimization, Label consistency regularization 및 Mixup을 모두 적용한 방법이다. 그림 3에 MixMatch 구현 방식을 도식화하였다.

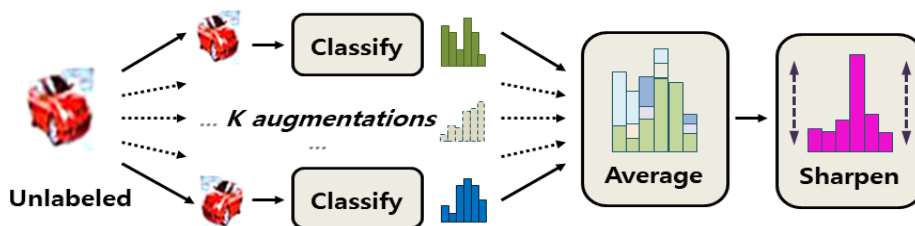


그림 3. MixMatch 방식  
Fig. 3. MixMatch operation

- Entropy Minimization : 분류기가 정답이 없는 라벨 데이터의 예측 엔트로피를 최소화하는 것으로, Entropy minimization의 방법 중에는 Pseudo-labeling 이 있다.

- Mixup : 증강이 완료된 정답라벨과 정답이 없는 라벨을 섞고, 그 데이터에 대하여 정답과 정답이 없는 라벨 데이터 이미지를 서로 겹치는 방식이다.

- Consistency regularization : 정답라벨과 정답이 없는 라벨 데이터를 사용하여 데이터를 학습하는데, 비슷한 데이터 혹은 약간 변형한 데이터를 가지고 학습시켰을 때 예측 결과가 비슷하게 나와야 한다는 의미이다.

③ FixMatch : 크로스 엔트로피 손실을 사용하여 정답라벨 이미지에서 지도학습 모델을 훈련하는 방식이다. 정답이 없는 라벨의 각 이미지에 대해 약한 증강과 강한 증강 방식을 적용하여 두 개의 이미지를 얻는다. 약하게 증강된 이미지가 모델에 전달되고 클래스에 대한 예측을 얻어서 가장 자신 있는 클래스의 확률은 임계값과 비교한다. 임계값 보다 높으면 해당 클래스를 기본 레이블 즉 Pseudo-label 로 사용하는 방식이다. 그런 다음 강하게 증강된 이미지가 모델을 통해 전달되어 클래스에 대한 예측을 진행한다. 이 예측은 정답 Pseudo-label과 비교하는 크로스 엔트로피 손실로 사용한다. 이때 두 손실이 결합하고 모델이 최적화된다. 그림 4에 FixMatch 구현 방식을 도식화하였다[9].

## 2.2 Fine-tuning

Fine-tuning이란 기존에 학습된 모델을 기반으로 아키텍처를 새로운 목적으로 이미지 데이터에 맞게

변형하고 이미 학습된 모델 가중치(weight)로부터 학습을 업데이트하는 방법을 말한다. 즉 딥러닝에서 이미 존재하는 모델에 추가 데이터를 투입하여 파라미터를 업데이트하는 것을 말한다. 좀 더 상세하게 설명하면, Fine-tuning은 정교한 파라미터 튜닝이라고 생각하면 되는데, 정교한 기술과 파라미터가 핵심이다. Fine-tuning을 완료하기 위해서는 기존에 학습된 레이어 내 데이터를 추가로 학습시켜 파라미터를 업데이트해야 한다. Fine-tuning을 할 때 주의할 점은 정교해야 한다.

완전히 랜덤한 초기 파라미터를 쓴다거나, 가장 아래쪽의 레이어 즉, 일반적인 특징을 학습한 덜 추상화된 레이어의 파라미터를 학습해버리면 오버피팅이 일어나 전체 파라미터가 붕괴하는 문제가 생기기 때문이다. Fine-tuning은 전이 학습의 한 분류이다. 사전 훈련된 모델을 자신의 필요에 맞게 용도 변경하려면 그림 5와 같이 네 가지 전략 중 한 가지 모델로 미세 조정해야 한다[10].

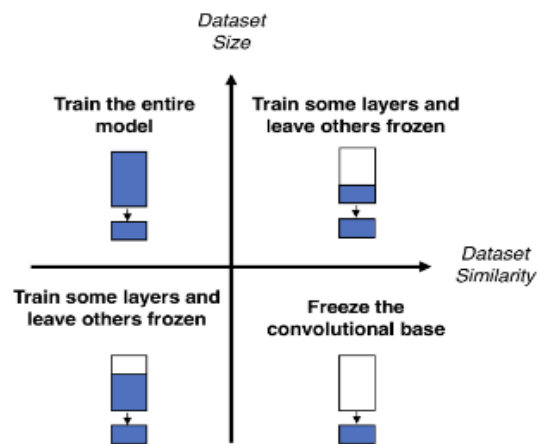


그림 5. Fine-tuning의 종류  
Fig. 5 Types of fine-tuning

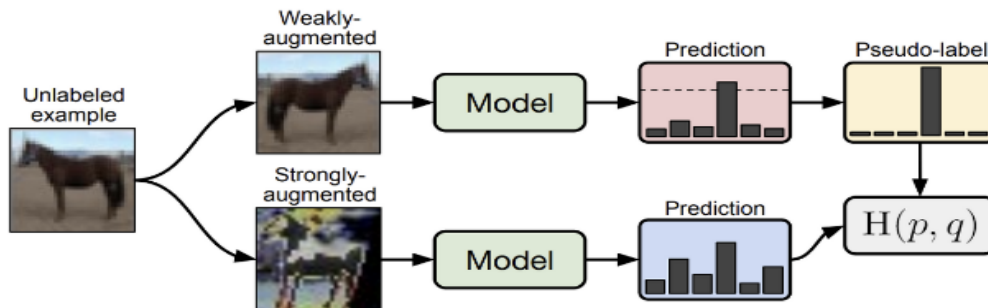


그림 4. FixMatch 방식  
Fig. 4. FixMatch operation

1사분면은 큰 데이터 세트이지만 사전 훈련된 모델의 데이터 세트와 다르다. 데이터 세트가 크기 때문에 처음부터 모델을 훈련하고 원하는 모든 작업을 수행할 수 있다. 2사분면은 사전 훈련된 모델의 데이터 세트와 유사한 대규모 데이터 세트를 사용한다. 데이터 세트가 크기 때문에 과적합이 문제가 되지 않아 원하는 만큼 학습할 수 있다. 3사분면은 사전 훈련된 모델의 데이터 세트와 달리 작은 데이터 세트를 활용한다. 훈련할 레이어 수와 동일한 레이어 수 사이의 균형을 찾는 것이 어렵고, 더 깊이 들어가면 모델이 과적합 될 수 있다. 4사분면은 작은 데이터세트이지만 사전 훈련된 모델의 데이터 세트를 모두 사용한다. 마지막 FC(출력 계층)만 변경하고 새 분류기를 훈련하면 되는 방식이다[11].

### III. 제안하는 방법

#### 3.1 F-guessed 방식

기존 준지도 학습 방식은 단순히 사전 학습된 모델을 이용하여 주어진 정답이 없는 라벨을 예측하는 방식이거나, 데이터 증강을 통하여 정답이 없는 라벨 데이터를 충분히 확보한 후 학습 모델을 통하

여 예측하는 방식으로 크게 2가지로 구분된다. 즉 기존 연구들은 데이터 확보 후 기존 학습 모델을 이용하여 예측(Guessed)에 초점을 맞추었다. 본 논문에서는 기존과 같이 정답라벨 데이터를 이용하여 사전 학습 모델을 생성하는 방식은 동일하나, 사전 학습 모델들의 인식률을 증가시키기 위하여 Fine-tuning을 통하여 파라미터를 업데이트하는 부분에서 차이가 있다.

즉 기존에 학습된 레이어를 추가로 학습시켜 파라미터를 업데이트하여 최적화 과정이 추가된 것을 말한다. Fine-tuning은 기존 신경망을 활용한 재학습 및 최적화 과정을 의미한다. 기존 학습모델의 정교한 파라미터 튜닝을 한 후 정답이 없는 라벨 데이터를 예측(Guessed)하면 더욱 정답라벨과 유사한 라벨을 예측할 수 있기 때문이다. 이렇게 하면 MixMatch, FixMatch 등과 같이 별도의 이미지 처리 없이 정답이 없는 라벨 데이터에 대한 라벨링의 정확성이 향상되며, n차 예측(F-guessed) 데이터와 정답라벨을 이용하여 학습 모델의 정규화가 가능해진다. 그림 6에는 Fine-tuning과 준지도 학습을 이용한 화재 데이터 생성 개념도를 표현하였으며, 그림 7에는 F-guessed 최적화 방식에 대한 개념을 그림으로 표현했다[10][12].

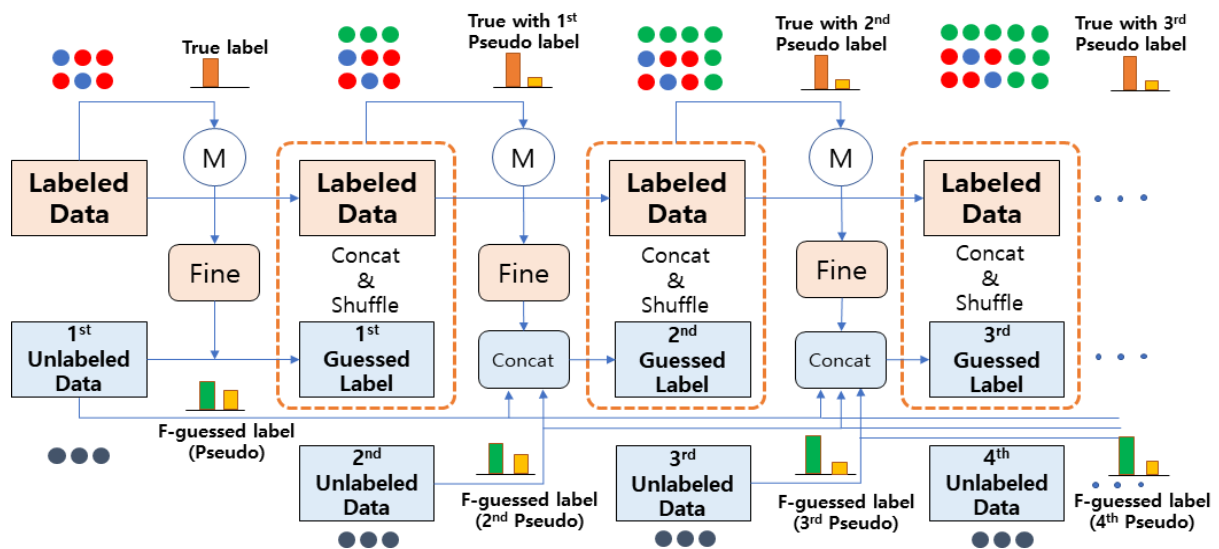


그림 6. Fine-tuning과 준지도 학습을 이용한 화재 데이터 생성 개념도

Fig. 6. Conceptual diagram of fire data generation using fine-tuning and semi-supervised learning

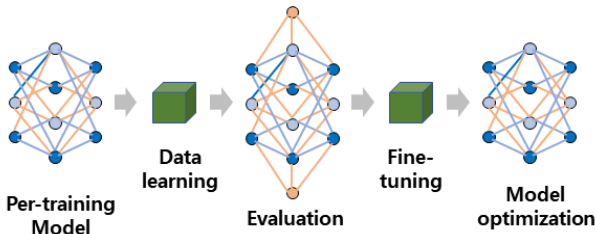


그림 7. F-guessed 최적화 방식  
Fig. 7. F-guessed optimization method

그리고 F-guessed 방식은 새로운 데이터가 추가될 때마다 추가 Fine-tuning을 매번 진행하도록 설계되어 있으며, 추가 Fine-tuning 한 학습 결과를 이용하여 기존 정답이 없는 라벨과 신규 정답이 없는 라벨 데이터를 재혼합하여 새로운 예측 모델을 만들게 하였다. 이렇게 진행하면 1회 학습 데이터의 신경망을 활용한 재학습 과정보다도 더 의미 있고, 정교한 파라미터 튜닝이 전체 정답이 없는 라벨 데이터 이미지(기존 Unlabeled data + 신규 Unlabeled data)를 재예측하여 더욱 정답라벨과 유사한 라벨을 예측이 가능하기 때문이다. 그리고 이렇게 중복적인 F-guessed 방식은 손실이 더 이상 감소하지 않는 범위까지 계속 진행하도록 설계하였다.

### 3.2 중복적 라벨링

중복적 라벨링 방식을 적용하는 이유는 화재 인식을 개선이 목적이다. 기존의 준지도 학습 방식은 정답이 없는 라벨과 정답라벨을 사용하여 이미지 Mixup 방식 등을 활용하여 정답라벨과 예측 결과가 비슷하게 출력되는 방식으로 인식을 개선하였다. 그리고 정답이 없는 라벨과 정답라벨 데이터 모두에 대해서 라벨링이 변경되는 방식이었다. 하지만 본 연구에서는 정답라벨은 정답이 없는 라벨과 라벨링 혼합 및 학습을 위해 데이터 공유는 하지만 정답라벨의 라벨링은 변형을 주지 않고 계속 유지하게 하였다. 하지만 정답이 없는 라벨에 대해서는 차수가 진행될수록 라벨링이 변형이 계속 진행이 되는 방식으로, 1차 학습 라벨링은 n차 학습 라벨링을 진행할수록 더욱 정답라벨과 유사한 라벨을 예측이 가능하도록 설계하였다. 즉 중복적 라벨링 방식은 다음과 같이 3가지 특징으로 간단히 설명될 수 있다[3].

① 정답라벨(Labeled data)은 준지도 학습을 위해 학습 차수와 관계없이 학습데이터로만 활용하고, 데이터나 라벨링 변형이 이루어지지 않는다. 즉 GT(Ground Truth) 및 Teacher로 활용한다.

② 정답이 없는 라벨(Unlabeled data)은 준지도 학습을 위해 학습 차수에 따라 항상 최종 예측 라벨링이 적용되도록 설계하였다. 즉 항상 Student로 활용하였다.

③ 정답라벨 + 예측(F-guessed) 데이터는 준지도 학습용 새로운 모델을 생성하기 위하여 정답라벨과 정답이 없는 라벨을 모두를 합치고 섞는 작업을 중복적으로 진행하며, 이 학습 가중치 값을 활용하여 새로운 Fine-tuning을 진행한다[13][14].

그림 8에는 중복적 라벨링에 대한 개념도이다.

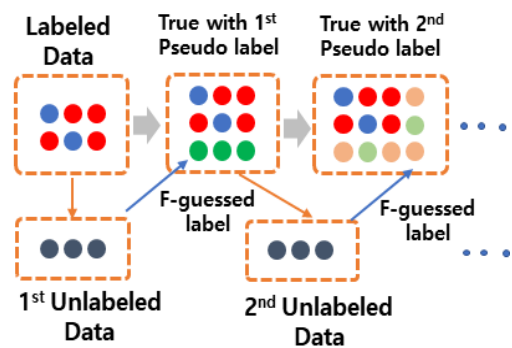


그림 8. 중복적 라벨링을 적용한 학습 방법  
Fig. 8. Learning method with overlapping labeling

### 3.3 데이터 증강 및 인식을 개선

준지도 학습 방식을 통한 인식을 개선하기 위해 앞서 제안한 F-guessed 방식과 중복적 라벨링 방식 외에 화재 정답라벨을 확보하기 위해 CycleGAN 방식과 알파 블렌딩(Alpha blending) 방식을 활용하였다. 현재 인터넷의 화재 데이터는 모두 대형 화재 위주로 화재가 발생하는 초기 데이터 확보가 어려워 CycleGAN 방식을 이용하여 특정 상황에 맞는 화재 데이터를 생성하고(예를 들면 사무실 내부 화재 등), 알파 블렌딩 방식을 통해서 특정 위치의 초기 화재 발화 이미지 데이터를 생성하였다[4].

이렇게 증강된 데이터와 인터넷 데이터를 병합하여 초기 화재 정답라벨 데이터를 확보하였다. 그리고 인식을 개선하기 위해 한 번에 많은 데이터를 확

보하여 학습시키는 것 보다는, 초기 데이터보다 반 정도 적은 또는 적은 수량의 정답이 없는 라벨 데이터를 추가하고 학습을 반복하는 게 인식률 개선에 도움이 된다는 것은 실험을 통하여 알게 되었다. 그리고 정답이 없는 라벨의 추가 횟수는 무한정 반복보다는 손실의 변화량이 적거나 미미할 때 멈추게 하였다. 모델 분류 정확도(Classification accuracy)가 높다고 객체 탐지 정확도(Detector accuracy)가 높은 것이 아니므로 화재 객체 탐지 실수가 최소화되도록 일부 추가 데이터에 Mixup 방식과 Mosaic 방식을 적용하였다.

Mixup 방식은 데이터 증강 기법의 하나로 두 이미지를 일정 비율로 섞어 새로운 데이터를 생성하는 방법이다. Mixup을 이용하면 클래스 중간에 위치하는 데이터에 대해 과도하게 확신하지 않기 때문에 성능이 향상될 수 있다. 또한 Mixup을 활용하면 데이터의 라벨이 잘못된 경우에도 기존 학습 방법보다 쉽게 대처가 가능하다. 그리고 Mosaic 증강 방식은 4장의 이미지를 합치지만 무작위로 잘라 붙이기보다는 크기만 조정하여 이미지가 버려지는 영역 없이 다 활용하였다. 그림 9는 이렇게 확보된 화재 데이터를 활용하여 정답라벨 데이터를 구축하고 활용하는 방식을 그림으로 정리하였다.

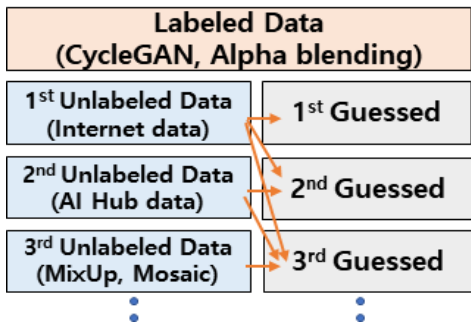


그림 9. 데이터 증강 및 활용 방법  
Fig. 9. How to augment and use data

#### IV. 실험 방법 및 결과

Fine-tuning과 준지도 학습을 이용한 화재 데이터 생성 및 인식률 개선 연구 실험은 CPU: AMD Ryzen 7 3700X 8-Core Processor 3.6 GHz, GPU: NVIDIA GeForce RTX 8000, RAM 32GB 컴퓨터 환경에서 실험하였다[3].

초기 데이터 세트는 표 1에 표시된 것과 같이 인터넷 및 자체 DA-FSL[4] 증강 방식을 이용하여 화재, 연기, 불꽃 이미지 데이터 세트를 확보하였다.

표 1. 초기 정답라벨 데이터 세트 정보  
Table 1. Basic labeled data set information

Data	Internet	Data augmentation	Total
Q'ty	3420	2145	5565

#### 4.1 실험 결과

초기 정답라벨 데이터는 준지도 학습을 위하여 초기 학습이 진행된 결과를 이용하여 표 2에 주어진 정답이 없는 라벨 데이터를 재예측하여 정답라벨과 유사한 예측이 가능하도록 5차에 걸쳐 진행하였다. 이렇게 F-guessed를 매번 진행하면 n차 학습 데이터의 신경망을 활용한 재학습 과정보다도 더 의미 있고, 정교한 파라미터 튜닝이 전체 정답이 없는 라벨 데이터 이미지에 가능하기 때문이다. 4~5차 데이터는 AI Hub의 화재 data를 활용하였다[15]. 그리고 학습은 Darknet 53을 이용하였으며, 화재 객체 검출은 Yolov4[16]를 사용하였다. 그림 10에는 학습을 위한 네트워크 구성 및 학습 순서를 그림으로 정리하였다. 그리고 학습을 위해 Max\_batches = 80,000으로 설정하여 실험을 진행하였다.

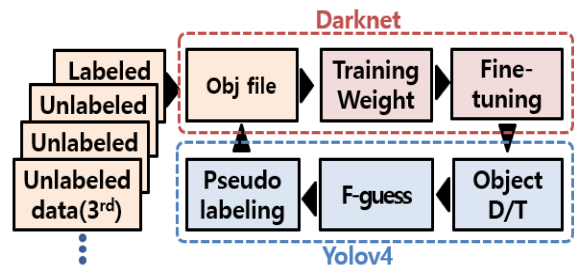


그림 10. 실험 네트워크 구성  
Fig. 10. Experimental network configuration

그 결과 표 2에는 F-guessed와 준지도 학습을 이용한 화재 데이터 생성 전체 이미지 수량을 표시하였다. 실험 결과 이미지는 초기 정답라벨 데이터 5,565개를 기준으로 6.3배 증가한 35,322개의 Pseudo 정답라벨 데이터를 확보한 결과를 얻었다.

표 2. Pseudo 정답라벨 데이터 세트 생성 정보  
Table 2. Pseudo labeled data set augmentation information

Data	Labeled	F-guessed (Pseudo labeled)	Total
0	5,565	0	5,565
1st	5,565	2,187	7,752
2nd	5,565	4,965	10,530
3rd	5,565	11,978	17,543
4th	5,565	21,601	27,166
5th	5,565	29,757	35,322

표 3에는 초기 정답라벨 데이터 학습 모델로부터 Pseudo 정답라벨 데이터가 완료된 화재 데이터를 기준으로 5차에 걸쳐 F-guessed와 준지도 학습을 이용한 화재 인식을 변화 결과를 Loss, mIOU, mAP 순으로 표시하였다. 테스트 결과 초기 정답라벨 데이터 대비 Loss는 최대 1.94% 감소하였으며, mIOU는 26% 증가하였고, mAP@0.5는 16.56%가 개선된 결과를 얻었다. 그리고 5차 이후 추가 학습을 진행하지 않은 이유는 프로그램 종료를 위한 기준을 손실의 변화가 적은 것을 기준으로 설정하였다. 손실이 적다는 것은 라벨링 데이터의 일관성이 확보되었다고 판단했다. 그리고 화재 데이터가 추가로 확보가 되면 추가 실험도 계속 진행하여 보겠다.

표 3. Max\_batches = 8,000 기준 객체 인식을 test 결과  
Table 3. Object precision rate test results based on max\_batch = 8,000

Method	Unit(%)			
	Loss	mIOU	mAP	
0	Train	<b>3.347</b>	<b>52.23</b>	<b>65.93</b>
	Fine-tuning	3.002	56.12	72.00
1st	Train	2.783	56.28	66.18
	Fine-tuning	2.902	59.64	75.42
2nd	Train	2.488	62.95	73.17
	Fine-tuning	2.232	62.83	76.45
3rd	Train	2.038	69.97	79.28
	Fine-tuning	<b>1.95</b>	<b>72.53</b>	<b>84.61</b>
4th	Train	1.77	73.35	74.80
	Fine-tuning	1.47	74.58	77.20
5th	Train	1.47	75.62	80.98
	Fine-tuning	<b>1.41</b>	<b>78.22</b>	<b>82.49</b>

mAP: mean average precision(mAP@50)

그리고 1차~5차에 걸친 이미지별 B/B(Bounding Box)와 클래스 성능 테스트 결과는 그림 11에 표시하였다. 차수별로 테스트에 활용된 이미지 데이터들은 학습에 사용하지 않은 일반 이미지를 인터넷에서 무작위로 선정하여 테스트한 결과이다. 실험은 이 중 7가지 이미지만을 활용하여 실험한 결과를 표시하였다[3]. 그 결과 차수가 반복될수록 초기 학습 데이터 세트 대비 클래스는 큰 변화가 없으나 B/B의 모양과 위치는 차수가 증가할수록 많은 차이를 보인다.

그림에는 1차~5차 B/B와 클래스 성능을 테스트한 전체 결과를 그림으로 표시하였다. 기존 정답라벨 대비 테스트한 1, 4, 6번 이미지에 대해서는 차수와 관계없이 비슷한 B/B와 클래스 결과를 보였으나, 그림 12와 같이 테스트 이미지 2, 5, 7번 대해서는 B/B와 클래스 변화가 테스트 차수에 따라 변화가 많았다. 테스트 이미지 2번은 정답라벨을 학습한 결과에서는 연기를 인식하지 못하였으나 2차 학습 이후부터는 정상 검출이 되었다. 이미지 5번은 정답라벨을 학습한 결과에서는 연기를 인식하지 못하였으며 2차까지 학습에서만 정상적으로 연기를 인식하였고, 5차에서는 여전히 연기를 인식하지 못하고 있다. 그리고 테스트 이미지 7번은 초기에 불을 연기로 잘못 인식하는 등 오인식이 심각했으나 5차 학습 이후 오인식이 완전히 사라진 결과를 보인다.

표 4에는 사람이 직접 수동으로 라벨링 한 36,749개의 정답라벨과 초기 정답 라벨링 5,565개를 기준으로 F-guessed가 완료된 35,322개의 데이터를 비교한 실험이다. 실험 결과 사람이 수동으로 라벨링한 것과 비교하여 Loss는 0.97, mIOU는 8.8%, mAP는 4.15% 개선된 결과를 얻었다.

표 4. 정답라벨 데이터와 F-guessed 데이터 비교 실험  
Table 4. Labeled and F-guessed data comparison experiment

Data	Q'ty	Loss	mIOU	mAP
Labeled	36,749	2.38	69.42	78.34
Pseudo labeled	35,322	1.41	78.22	82.49



Test	Labeled	1 <sup>st</sup> Guessed	2 <sup>nd</sup> Guessed	3 <sup>rd</sup> Guessed	4 <sup>th</sup> Guessed	5 <sup>th</sup> Guessed
1						
2						
3						
4						
5						
6						
7						

그림 11. 1차~5차 B/B(Bounding box)와 클래스 성능 테스트 결과  
 Fig. 11. 1st~5th B/B(Bounding box) and classification performance test results

하지만 인식률이 높다고 현장에서 바로 활용할 수 있는 것은 아니다. 현장의 다양한 환경과 조건을 만족하기 위해서는 더 많은 데이터가 필요로 하고 더 많은 기술적 아이디어가 추가되어야 할 것이다.

다만 이 실험은 사람이 수동으로 라벨링 하는 수고를 조금이나 마 줄이고 준지도 학습을 통하여 사람의 개입을 최소화하면서 재난 현장에 활용 가능성을 시험하였다.



그림 12. 주요 라벨링 성능 비교  
 Fig. 12. Comparison of key labeling performance

## V. 결 론

본 논문의 연구는 화재 및 재난 분야와 같이 데이터 수집이 제한된 특수한 경우 데이터 수집 및 라벨링 작업을 쉽게 하기 위하여 기존의 준지도학습인 Pseudo labeling 방식을 보완한 F-guessed 방식과 중복적 라벨링 기법을 제안하였다. 그 결과 초기 정답라벨 데이터 대비 Loss는 최대 1.94% 감소하였으며, mIOU는 26% 증가하였고, mAP@0.5는 16.56%가 개선된 결과를 얻었다. 그리고 정답이 확보된 데이터 숫자도 초기 정답라벨 데이터 5,565개를 기준으로 6.3배 증가한 35,322개의 F-guessed 정답 데이터가 확보되었다. 그래도 여전히 부족한 화재 데이터 수집을 위해 지속적 추가 실험이 요구되며, 향후 화재 데이터가 아닌 공개된 일반 데이터를 활용한 추가 성능 검증도 병행하여 연구를 진행할 예정이다.

## References

- [1] Amit Chaudhary, "Semi-Supervised Learning in Computer Vision", <https://amitnss.com/2020/07/semi-supervised-learning> [accessed: Sep. 10, 2022]
- [2] Yassine Ouali, Céline Hudelot, and Myriam Tami, "An Overview of Deep Semi-Supervised Learning", Machine Learning(cs.LG), arXiv:2006.05278, Jul. 2020. <https://doi.org/10.48550/arXiv.2006.05278>.
- [3] Jong-Sik Kim and Dae-Seong Kang, "A Study on Automatic Data Generation and Object Recognition Rate Improvement using Heterogeneous Object Detection and Duplicate Labeling Method", The Journal of Korean Institute of Information Technology, Vol. 20, pp. 21-29, May 2022. <https://doi.org/10.14801/jkiit.2022.20.5.21>.

- [4] Hye-Youn Lim, Jun-Mock Lee, and Dae-Seong Kang, "A Method for Improving Learning Convergence Curve and Learning Time of DA-FSL Model using Knowledge Distillation", *The Journal of Korean Institute of Information Technology*, Vol. 18, pp. 25-32, Oct. 2020. <https://doi.org/10.14801/jkiit.2020.18.10.25>.
- [5] Xiaokang Chen, Yuhui Yuan, Gang Zeng, and Jingdong Wang, "Semi-Supervised Semantic Segmentation with Cross Pseudo Supervision", *Computer Vision and Pattern Recognition(CVPR)*, pp. 2613-2622, Jun. 2021.
- [6] Vinko Kodžoman, "Pseudo-labeling a simple semi-supervised learning method", <https://datawhathow.com/pseudo-labeling-semi-supervised-learning> [accessed: Sep. 10, 2022]
- [7] David Berthelot, Nicholas Carlini, Ian Goodfellow, Nicolas Papernot, Avital Oliver, and Colin A Raffel, "MixMatch: A Holistic Approach to Semi-Supervised Learning", *Neural Information Processing Systems 32*, 2019.
- [8] David Berthelot, Nicholas Carlini, Ekin D. Cubuk, Alex Kurakin, Kihyuk Sohn, Han Zhang and Colin Raffel, "ReMixMatch: Semi-Supervised Learning with Distribution Alignment and Augmentation Anchoring", *Machine Learning (stat.ML)*, arXiv:1911.09785, Feb. 2020. <https://doi.org/10.48550/arXiv.1911.09785>.
- [9] Kihyuk Sohn, David Berthelot, Nicholas Carlini, Zizhao Zhang, Han Zhang, Colin A. Raffel, Ekin Dogus Cubuk, Alexey Kurakin, and Chun-Liang Li, "FixMatch: Simplifying Semi-Supervised Learning with Consistency and Confidence", *Neural Information Processing Systems 33*, 2020.
- [10] Ananya Kumar, Aditi Raghunathan, Robbie Jones, Tengyu Ma, and Percy Liang, "Fine-Tuning can Distort Pretrained Features and Underperform Out-of-Distribution", *Computer Vision and Pattern Recognition(cs.CV)*, arXiv:2202.10054, Feb. 2022. <https://doi.org/10.48550/arXiv.2202.10054>.
- [11] Pedro Marcelino, "Transfer learning from pre-trained models. In *Towards Data Science (2018)*", <https://towardsdatascience.com/transfer-learning-from-pre-trained-models-f2393f124751> [accessed: Sep. 10, 2022]
- [12] Marius Mosbach, Maksym Andriushchenko, and Dietrich Klakow, "On the Stability of Fine-tuning BERT: Misconceptions, Explanations, and Strong Baselines", *Machine Learning(stat.ML)*, arXiv:2006.04884, Mar. 2021. <https://doi.org/10.48550/arXiv.2006.04884>.
- [13] Barret Zoph, Golnaz Ghiasi, Tsung-Yi Lin, Yin Cui, Hanxiao Liu, Ekin Dogus Cubuk, and Quoc Le, "Rethinking Pre-training and Self-training", *Neural Information Processing Systems 33*, 2020.
- [14] Mengde Xu, Zheng Zhang, Han Hu, Jianfeng Wang, Lijuan Wang, Fangyun Wei, Xiang Bai, and Zicheng Liu, "End-to-End Semi-Supervised Object Detection with Soft Teach", *IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 3060-3069, 2021. <https://doi.org/10.48550/arXiv.2106.09018>.
- [15] AI Hub data, <https://aihub.or.kr> [accessed: Apr. 10, 2022]
- [16] Alexey Bochkovskiy, Chien-Yao Wang, Hong-Yuan, and Mark Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection", *Computer Vision and Pattern Recognition(cs.CV)*, arXiv:2004.10934, Apr. 2020. <https://doi.org/10.48550/arXiv.2004.10934>.

저자소개

김 종 식 (Jong-Sik Kim)



1991년 2월 : 부경대학교  
전자공학과(공학사)  
2020년 8월 : 동아대학교  
전자공학과(공학석사)  
2020년 9월 ~ 현재 : 동아대학교  
전자공학과 박사과정  
관심분야 : 영상처리, AI

강 대 성 (Dae-Seong Kang)



1994년 5월 : Texas A&M 대학교  
전자공학과(공학박사)  
1995년 ~ 현재 : 동아대학교  
전자공학과 교수  
관심분야 : 영상처리, AI, 패턴인식