

인공위성 영상 기반 도로 검출에서 손실 함수 특성 분석

김나경*, 윤영선*†

Characteristics of Loss Functions for Road Detection on Satellite Images

Nagyeong Kim*, Young-Sun Yun*†

요 약

본 연구는 인공위성 영상 기반의 도로 검출에서 손실 함수(Loss Function)와 성능의 변화를 정량적으로 분석하였다. AI Hub에서 제공하는 위성영상 도로 데이터셋을 활용하였으며, 장거리 특징을 모델링하는 NL-LinkNet 모델을 기반으로 실험을 수행하였다. 본 논문에서는 대표적인 화소 기반 손실 함수인 Cross Entropy와 Focal Loss, 영역 기반 손실 함수인 Dice와 Lovasz Loss, 그리고 이들을 결합한 복합 손실 함수를 비교 분석하였다. 실험 결과, Dice 및 Dice 기반 복합 손실 함수가 전반적으로 안정적인 성능과 낮은 표준 편차를 보이며 우수한 결과를 나타냈다. 반면, Lovasz 손실 함수는 IoU를 직접 손실 함수에 반영하고 있어 높은 성능이 예측되었으나, 도로 검출 실험에서는 초기 낮은 학습 안정성과 성능 저하를 보였다. 본 연구로부터 도로와 같이 복잡하고 클래스 불균형이 심한 객체를 검출하는 작업에서는 복합 손실 함수의 이용이 안정적인 성능과 각 클래스의 불균형을 해소할 수 있을 것으로 보이며, 특히 Dice와 Focal 기반의 복합 손실 함수 또는 Lovasz 기반의 복합 손실 함수를 이용하는 것이 균형 잡힌 성능과 초기 학습 안정성에 적합한 것으로 판단된다.

Abstract

This study quantitatively analyzes the relationship between loss functions and performance variations in road detection using satellite imagery. Road datasets from AI Hub were utilized, and experiments were conducted based on the NL-LinkNet model, which models long-range feature dependencies. This paper compares and analyzes representative pixel-wise loss functions, Cross Entropy and Focal Loss; region-based loss functions, Dice and Lovasz Loss; and hybrid loss functions combining these. Results showed Dice and Dice-based hybrid loss functions generally provided stable, high-performing results, while Lovasz Loss, though theoretically effective due to its IoU-based formulation, demonstrates poor initial training stability and lower performance in practice. The results suggest that employing hybrid loss functions can provide stable performance and address class imbalance for complex object detection tasks with significant class imbalance, such as road detection. In particular, Dice and Focal-based hybrid loss functions, or Lovasz-based hybrid loss functions, are deemed suitable for balanced performance and initial training stability.

한글키워드 : 인공위성 영상, 도로 검출, 도로 분할, 손실 함수

keywords : satellite images, road detection, road segmentation, loss functions

* 한남대학교 정보통신공학과

접수일자: 2025.05.18. 심사완료: 2025.05.23.

† 교신저자: 윤영선(email: ysyun@hnu.kr)

게재확정: 2025.06.20.

1. 서론

영상 분할(Image Segmentation)은 컴퓨터 비전 분야의 핵심 기술 중 하나로, 디지털 영상을 의미 있는 영역 또는 객체로 분할한 후, 고유한 레이블을 할당하는 것을 말한다. 이는 영상 내의 각 구성 요소를 정확하게 이해하고 분석할 수 있도록 함으로써, 다양한 응용 분야에서 중요한 역할을 수행한다. 초기 영상 분할 연구는 주로 전통적인 영상 처리 기법에 기반을 두었으나, 최근 심층 학습(Deep Learning)의 발전과 함께 뛰어난 성능 향상을 이루며 그 중요성이 강조되고 있다. 특히, 컨볼루션 신경망(Convolutional Neural Network, CNN) 기반의 다양한 분할 모델들은 복잡한 고해상도의 영상에서도 뛰어난 분할 결과를 보여주며, 의료 영상 분석, 자율 주행, 로봇 인식 등 첨단 기술 발전에 핵심적으로 기여한다.

영상 분할 기술의 활용은 다양한 분야에서 빠르게 증가하고 있으며, 특히 인공위성 영상 분석 분야는 그 잠재력과 중요성이 매우 크다. 고해상도 인공위성 영상은 광범위한 지역에 대한 시각 정보를 효율적으로 제공하며, 도시 계획, 환경 모니터링, 재난 관리, 농업 분석 등 다양한 목적을 위해 활용된다. 이러한 인공위성 영상 분석에 이미지 분할 기술을 적용함으로써, 특정 객체나 지형지물을 자동으로 추출하고 분석할 수 있게 되었다. 예를 들어, 토지 활용 분류, 건물 추출, 산림 변화 감지 등 다양한 분석 작업을 통하여 정책 입안이나 구획 관리 등에서 매우 중요한 역할을 담당한다.

본 연구에서는 인공위성 영상 분석의 중요한 활용 사례 중 하나인 도로 검출(Road Extraction/detection)에 초점을 맞춘다. 정확하고 신뢰성 있는 도로 정보는 도시 계획, 교통 관리, 내비게이션 시스템 구축, 재난 발생 시 긴급 구조 경로 파악 등 사회 기반 시설 관리 및 운영에

필수적인 요소이다. 기존의 수동적인 방법으로는 광범위한 지역의 도로 정보를 추출하는 데 많은 시간과 노력이 소요되며, 주관적인 오류 발생 가능성 또한 존재하였다. 따라서, 인공위성 영상으로부터 도로를 자동으로 검출하는 기술은 효율성과 정확성 측면에서 매우 중요한 의미가 있다.

최근 심층 학습 기반의 영상 분할 모델들은 인공위성 영상에서의 도로 검출 분야에서도 상당한 발전을 이루었으나, 여전히 다양한 환경 변화와 복잡한 배경 요소로 인해 완벽한 성능을 달성하는 데 어려움이 존재한다. 특히, 손실 함수(Loss Function)는 심층 학습 모델의 학습 과정을 유도하고 성능을 결정하는 핵심적인 요소이며, 모델이 예측한 결과와 실제 값 간의 차이를 줄여 최종 시스템의 성능을 결정짓는 중요한 요소이다. 도로 검출과 같은 복잡한 문제에서는 적절한 손실 함수의 선택은 모델의 성능을 크게 좌우하기 때문에, 본 연구에서는 다양한 손실 함수가 도로 검출 정확도에 미치는 영향을 평가하고, 최적의 손실 함수 선택에 도움이 되고자 한다. 적절한 손실 함수의 선택은 AI 모델이 도로 영역의 특징을 효과적으로 학습하고, 다양한 잡음 요소와 방해 요소에도 강인한 예측 결과를 도출하는 데 결정적인 영향을 주기 때문이다.

본 연구는 AI Hub에서 제공하는 다목적 실용 위성 영상 데이터 중 도로 검출용 데이터를 활용 [1]하였으며, 데이터와 같이 배포하는 AI 모델[2]을 수정하여 다양한 손실 함수에 대한 특성을 분석하였다. 객체 탐지 및 영상 분할, 인공위성 영상 기반의 연구는 활발히 진행되고 있으나, 정확한 손실 함수의 특성 분석 없는 단순한 적용은 목적 시스템의 성능에 영향을 줄 수밖에 없다.

본 논문의 구성은 다음과 같다. 2장에서 널리 사용되는 손실 함수와 개략적인 특성을 소개하고, 3장에서는 도로 검출에 사용되는 모델을 살펴본 후, 4장에서 실험 및 결과를 정리하여 분석

하고, 마지막으로 5장에서 요약 및 결론을 맺는다.

2. 손실 함수

이번 장에서는 영상 분할에 널리 사용되는 손실 함수들을 주요 그룹별로 분류하고 각 그룹의 특징을 살펴본다. 최종적으로 본 연구에서 비교 분석할 손실 함수인 Cross Entropy Loss, Focal Loss, Dice Loss, 그리고 Lovasz Loss 함수의 이론적 배경 및 특징을 상세히 기술한다.

2.1 영상 분할 손실 함수 분류

영상 분할에 사용된 손실 함수는 다양한 기준에 따라 분류하고 있으나, 기능적 특징에 따라 크게 다음과 같은 그룹으로 분류한다 [3]

표 1. 영상 분할에 사용되는 손실 함수들
Table 1. Loss functions of image segmentation

Type	Loss functions
Pixel-wise (Distribution)	Cross-Entropy, TopK, Focal, Distance map derived CE, etc.
Region-based	Dice, Log-Cosh Dice, Generalized Wasserstein Dice, IOU (Jaccard), Lovasz-Softmax, Tversky, Focal Tversky, Robust-T, etc.
Boundary-based	Boundary, Hausdorff Distance, Boundary-aware, Active Boundary, InverseForm, etc.
Combination/Hybrid	Dice-Cross Entropy, Focal-Tversky, etc.

- 화소 또는 분포 기반 손실 함수

영상의 각 화소에 대하여 예측값과 실제 레이블 간의 차이를 독립적으로 계산하여 손실을 계산한다. 주로 분류 문제에서 사용되는 손실함수들이 적용된다. 이 방식은 각 화소에 대한 손실이 독립적으로 계산되므로, 클래스 불균형 문제

에 취약할 수 있으며, 배경 클래스의 화소 수가 많은 경우 배경 클래스에 치우쳐서 학습될 가능성이 높다. 대표적으로 적용되는 손실 함수는 Cross Entropy (CE) Loss나 가중치 CE, Focal Loss 등이 있다.

- 영역 기반 손실 함수

예측된 분할 영역과 실제 분할 영역간의 유사성 또는 차이를 측정하여 손실을 계산하는 부류이다. 영상 내의 객체 또는 영역 전체의 속성을 고려하며 화소 단위의 정확성뿐만 아니라 영역의 형태나 크기와 같은 정보도 학습에 반영할 수 있다. 예측 영역과 실제 영역의 겹침 정도를 손실에 반영하기 때문에 클래스 불균형 문제를 화소 기반 손실 함수에 비하여 더 강인한 경향이 있다. 대표적인 손실 함수로는 Dice Loss나 IoU Loss, 그리고 Lovasz Loss 등이 있다.

- 경계 기반 손실 함수

예측된 객체의 경계와 실제 객체의 경계 간의 차이를 측정하여 손실을 계산한다. 객체의 내부 영역보다는 경계에 정확성에 초점을 맞추는 부류이다. 대표적인 손실 함수로는 Boundary Loss 방식이 있다.

- 복합 손실 함수

화소 기반 손실 함수와 영역 기반 손실 함수의 장점을 결합한 방식이다. 복합 손실 함수는 클래스 불균형에 더 효과적으로 대처하고, 분할 결과의 정확성과 영역의 일관성을 동시에 향상하는 것을 목표로 한다. 대표적인 방식으로 CE와 Dice를 결합하는 방식이나 Focal Loss와 영역 기반 손실 함수를 결합하는 방식이 있다.

2.2 손실 함수 선택

본 연구에서는 영상 분할 분야에서 널리 사용

되며, 서로 다른 특징을 갖는 대표적인 손실 함수인 Cross Entropy Loss, Focal Loss, Dice Loss, 그리고 Lovasz Loss를 선택하여 인공위성 영상 기반 도로 검출에서 각 손실 함수의 특성과 성능을 비교 분석하고자 한다. 각 손실 함수의 특성은 다음과 같다.

- Cross Entropy Loss : 픽셀 기반의 기본적인 손실 함수로서, 다른 손실 함수들의 성능을 비교하는 기준(baseline) 역할을 수행할 수 있다.

$$L_{CE} = -\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^C y_{ic} \log(p_{ic}) \quad (1)$$

식(1)의 CE 손실은 모든 화소 N 에 대하여 클래스별로 (총 클래스 수 C) 소프트맥스 활성화가 적용된 예측 확률 p_{ic} 과 실제 레이블 y_{ic} 간의 음의 로그 우도(negative log likelihood)를 최소화한다. CE 손실은 구현이 간단하고 직관적이며, 학습 초기 단계에서 빠른 수렴을 보여 안정적인 학습을 유도하는 경향이 있지만, 화소 간의 공간적 관계를 고려하지 않는다는 특징이 있다.

- Focal Loss: 화소 기반 손실 함수이지만, 클래스 불균형 문제를 완화하고 학습이 어려운 샘플에 집중하도록 설계되어 도로와 배경 클래스 간의 불균형이 심할 수 있는 인공위성 영상 분석에 적절할 것으로 판단하여 고려하였다. Focal 손실 함수는 CE 손실 함수에 감마 ($\gamma \geq 0$)를 적용하여 예측 확률 p_{ic} 에 대한 기여 가중치로 사용한다[4].

$$L_{FL} = -\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^C y_{ic} (1 - p_{ic})^\gamma \log(p_{ic}) \quad (2)$$

식(2)에서 γ 값이 커질수록 쉽게 분류되는 샘플에 대한 손실 감소 효과가 커지므로, 학습이 어려운 적은 양의 클래스에 더 집중할 수 있다.

일반적으로 $\gamma = 2$ 로 기본값을 설정한다.

- Dice Loss: 영역 기반의 손실 함수로서, 예측된 도로 영역과 실제 도로 영역 간의 겹침 정도를 직접적으로 측정하므로, 도로와 같이 길고 얇은 형태의 객체 분할에 효과적일 수 있다. 클래스 불균형 문제에도 비교적 강인한 것으로 알려져 있다. IoU(Intersection over Union)가 $A \cap B / A \cup B$ 로 표현되므로, Dice 손실은 식(3)과 같이 직접 Dice 계수를 계산하여 손실 값을 결정한다[5].

$$\begin{aligned} L_{Dice} &= 1 - \frac{2 * |A \cap B|}{|A| + |B|} \\ &= 1 - \frac{2 \sum_i^N p_i y_i}{\sum_i^N p_i^2 + \sum_i^N y_i^2} \end{aligned} \quad (3)$$

- Lovasz Loss: IoU 기반의 영역 손실 함수이며, 미분할 수 있는 형태로 최적화가 가능하도록 설계되었다. 특히 다중 클래스 분할 문제에서 좋은 성능을 보여, 도로 검출에서 Dice Loss와 비교하여 특성을 파악하고자 한다[6]. Lovasz 손실은 각 화소에 대하여 예측 확률과 실제 레이블 간의 에러를 정의하고, 잘못 분류된 픽셀의 개수를 의미하는 Lovasz 확장 $\Delta J(m)$ 을 이용하여 클래스별 손실 값 $L_c(p, y) = \tilde{\Delta} J(m(p))$ 을 계산한 후 식(4)와 같이 합하여 최종적으로 계산한다.

$$L_{Lovasz} = \sum_{c=1}^C L_c(p_c, y_c) \quad (4)$$

3. 도로 검출 모델

AI Hub 사이트¹⁾에서 제공하는 AI 모델은 NL-LinkNet을 구현한 모델로서 다음의 특성이

- 1) 위성영상 객체 판독,

<https://aihub.or.kr/aihubdata/data/view.do?dataSetSn=73>

있다.

NL-LinkNet[7]은 비지역적 블록 (non-local blocks)을 통하여 이미지 내의 전역 특징 간의 장거리 의존성을 모델링하고 있다. 이 특징은 도로가 위성 영상에서 그림자나 나무, 건물 등에 의하여 부분적으로 가려질 수 있는 경우를 고려하였다. NL-LinkNet의 비지역적 연산은 모든 위치에서 특징의 가중합으로 특징 맵 값을 계산하여 모델이 장거리 의존성과 원거리 정보를 효율적으로 모델링하도록 하였다. 구현 모델은 ResNet34를 기본 모델로 하고, 비지역 블록을 인코더의 중후반부에 배치하여, 다음의 모델 구조를 가진다.

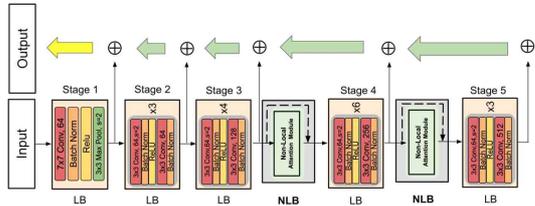


그림 1. NL-LinkNet 구조 [7]
Fig. 1. An architecture of NL-LinkNet [7]

기본 NL-LinkNet모델은 ResNet34를 기본으로 Adam Optimizer에 Binary Cross Entropy와 Dice 손실 함수를 사용하고 있으며[8], 배포된 AI 모델은 Lovasz 손실 함수를 사용하였다[9]. 또한 모델 추론 결과의 정확도를 높이기 위하여 Test Time Augmentation을 통하여 하나의 이미지를 다양하게 변형하여 일반화 성능을 높이는 방식을 적용하였다.

4. 실험 및 결과

4.1 도로 데이터셋

실험에 사용된 데이터셋은 AI Hub 사이트에

표 2. 버전별 도로 종류 비교

Table 2. Comparisons of road types by version

Labels v1.0	Labels v1.5
Motorway	Motorway/Truck
Primary	Primary
Secondary	Secondary
Tertiary	Tertiary
Residential	Path
Unclassified	Railway/Runway /Construction
Background	Background

서 제공되는 것으로 해상도가 각각 70cm, 55cm인 3호와 3A호 광학 영상을 원천 데이터로 사용하고 필터링과 라벨링 작업을 한 것이다[1].

원시데이터 영상의 크기는 약 $24000 \times 24000 px^2$ 이며, $1024 \times 1024 px^2$ 패치 형태로 제작되었고, 촬영 지역은 총 4개 지역으로 미국의 로스앤젤레스, 독일 볼프스부르크, 중국 상하이, 이집트 뉴카이로이다. 최초의 레이블링은 7가지 종류로 고속도로, 제1도로, 제2도로, 제3도로 (기타 포장도로), 주거 도로, 미분류 도로로 배포되었으나, 개정 작업을 거쳐 9가지 종류로 변경되었다. 기존 연구와의 유사성을 높이기 위하여 표 2와 같이 분리형 고속도로와 고속도로를 통합하고, 철도 등을 미분류 도로로 대응시켰다.

표 3. 학습/검증 데이터의 도로 종류 비율

Table 3. Road type percentage of training/validation dataset

	Training	Validation
Motorway	0.90	0.78
Primary	1.02	0.67
Secondary	0.93	0.74
Tertiary	7.11	4.60
Residential	0.73	0.55
Unclassified	0.61	0.87
Background	88.69	91.78

클래스 분류 기준 v1.5를 기준으로 학습과 검증 데이터셋에 대하여 각 도로 종류 비율²⁾을 계산하면 표 3과 같다. 학습데이터나 검증데이터에서 배경을 제외하고 가장 많은 부분을 차지하는 클래스는 기타 도로 (Tertiary)이며, 검증 데이터셋의 도로 비율이 학습 데이터셋에 비하여 작은 분포로 관찰되었다.

표 4. 시스템 및 실험 환경

Table 4. System information and experimental environments

Type	Description
OS	Ubuntu 24.04.2
Main Memory	128G
Graphic Card	NVIDIA GeForce 1080 Ti x 2
python	3.10.17
torch	1.13.1
torchvision	0.14.1
batchsize_per_gpu	1, 4, 8, 16, 20, 24
patch size	512 x 512
Learning rate	0.0008
Epoch	1,000

4.2 실험 환경 및 조건

인공위성 영상 기반 도로 검출 연구에서 손실 함수의 특성을 파악하기 위한 환경은 표 4와 같이 구축하였다. 배포 모델에서는 GPU수 = 4, batch_size_per_gpu = 20, 패치 이미지 768x768 크기로 설정되었으나, 본 연구에서는 GPU 수 = 2, batch_size_per_gpu = 1, 4, 8, 16, 20, 24, 패치 이미지 크기 512x512로 설정하였다. GPU의 한계로 인하여 batch_size_per_gpu >= 20 인 경우에는 기울기 누적 (gradient accumulation) 방식을 적용하여 큰 배치 크기를 모사하였다.

본 연구의 목적이 도로 검출 성능 향상이 아니기 때문에 다른 연구들에서 제시한 SOTA

2) 도로별 레이블이 차지하는 화소 수 비율로 계산

(State of the Art) 알고리즘을 적용하지 않고, 데이터와 같이 배포되는 기본 모델을 이용하여 손실 함수에 따른 성능 변화를 비교 분석하였다. 배포 모델과 비교한 SegFormer를 이용한 연구 결과[10]가 있으나 클래스 분류 기준과 실험 환경에 따른 패치 크기가 달라 직접적인 비교가 불가능하였다. 참고로 데이터셋과 같이 배포된 학습모델을 이용하여 검증 데이터셋을 평가한 결과 28.6%의 mIoU를 보여, v1.0 클래스 분류 기준 38.5%에 현저히 못 미치는 것을 확인했다. 오류의 대부분이 주거 도로와 미분류 도로에서 발생한 것으로 보아 v1.0과 v1.5 클래스 분류 정보가 주거 도로와 미분류 도로, 그리고 기타 도로에서 변경되었을 것으로 추측한다.

4.3 성능 비교 및 손실 함수 분석

4.2절의 환경에서 학습률은 고정하고 배치 크기에 따른 손실 함수별 성능을 관찰하였다. 배치 크기가 1인 경우에는 학습 경향을 파악하기 힘들었기 때문에 전체 성능 분석은 참고 값으로만 제시하였으며, 전반적으로 배치 크기가 증가함에 따라 안정적인 학습이 관찰되었다 (그림 2 참조). 특히 focal 손실 함수의 경우 배치 크기가 작은 경우에 성능의 변화가 크게 관측되었으나, 배치 크기가 증가함에 따라 CE 손실에 비하여 안정적인 성능 변화를 확인할 수 있었다. Dice 손실의 경우 배치 크기에 상관없이 안정적인 변화와 높은 성능을 보였다.

예상과 다르게 Lovasz 손실 함수는 저조한 성능을 보이고 있으며, 이는 손실 함수의 특성상 다른 손실 함수들에 비하여 학습 진행 속도가 더디고, 안정하지 않기 때문으로 추정된다. Lovasz 손실을 제안한 연구에서처럼 반복 횟수에 따른 학습률의 지속적인 변화와 epoch를 증가시킨다면 다른 손실 함수와 비교할 수 있는 성능을 보일 수 있을 것으로 예측한다.

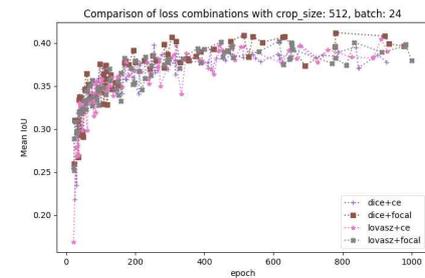
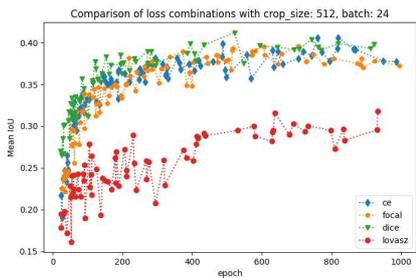
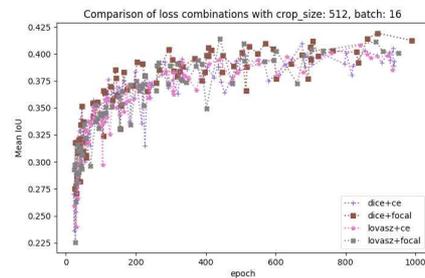
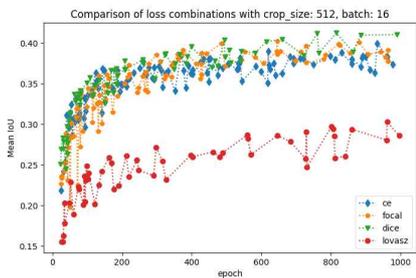
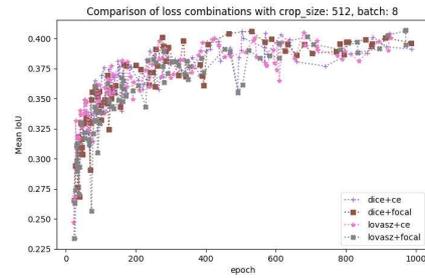
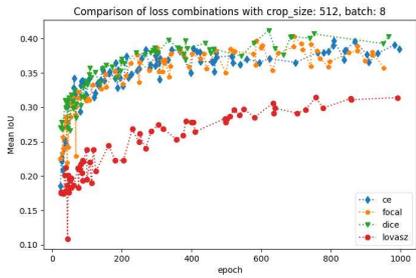
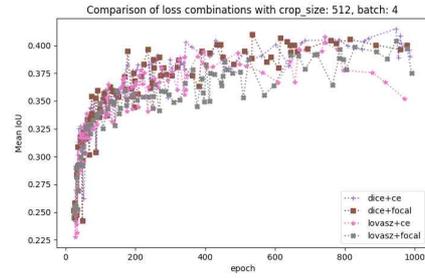
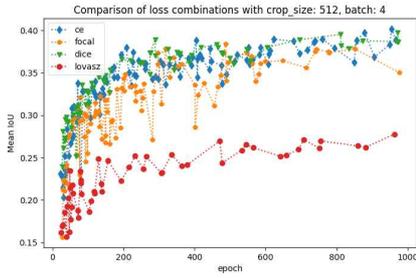


그림 2. 손실 함수의 배치 크기 (위로부터 4, 8, 16, 24)에 따른 성능 비교 (CE: 파랑, Focal: 오렌지, Dice: 초록, Lovasz: 빨강)
 Fig. 2. Performance comparison by batch size of loss function (from top: 4, 8, 16, 24) (CE: blue, Focal: orange, Dice: green, Lovasz: red)

그림 3. 복합 손실함수의 배치 크기 (위로부터 4, 8, 16, 24)에 따른 성능 비교 (Dice+ce: 보라, Dice+focal: 브라운, Lovasz+ce: 핑크, Lovasz+focal: 회색)
 Fig. 3. Performance comparison by batch size of combined loss function (from top: 4, 8, 16, 24) (Dice+ce: purple, Dice+focal: brown, Lovasz+ce: pink, Lovasz+focal: gray)

표 5. 배치 크기에 따른 손실 함수 별 성능 (mIoU) 비교
Table 5. Performance (mIoU) comparison of loss functions by batchsize per gpus

batch size per gpu \ Losses	Single Loss functions				Combined Loss functions			
	CE	Focal	Dice	Lovasz	Dice + CE	Dice + Focal	Lovasz + CE	Lovasz + Focal
1	17.43	18.51	24.27	19.63	22.03	20.32	21.02	21.73
4	40.18	38.43	39.71	27.75	41.51	41.01	40.80	40.39
8	39.69	40.29	41.21	31.45	40.60	40.60	40.50	40.68
16	39.97	40.22	41.24	30.30	40.95	41.92	40.81	41.41
20	39.84	39.12	41.05	33.50	40.17	41.18	41.66	40.12
24	40.61	39.59	41.20	31.80	39.76	41.19	40.43	40.19

표 6. 손실 함수별 최고 성능의 도로 종류별 mIoU, 정확도, 표준 편차 비교
Table 6. Comparison of mIoU, accuracy, and standard deviation for best performing road types by loss functions

Types \ Losses		Single Loss functions				Combined Loss functions			
		CE	Focal	Dice	Lovasz	Dice + CE	Dice + Focal	Lovasz + CE	Lovasz + Focal
batchsize per gpu		24	8	16	20	4	16	20	16
Road Types	Background	92.11	92.15	91.30	88.60	91.86	91.93	92.38	91.44
	Motorway	41.68	40.92	44.62	39.29	43.33	47.15	40.56	45.02
	Primary	30.74	32.45	29.48	19.06	28.55	31.88	30.72	31.28
	Secondary	26.63	28.75	28.41	9.92	25.26	29.22	31.22	25.74
	Tertiary	31.89	31.45	31.27	31.86	32.14	30.69	30.01	31.29
	Residential	15.73	7.75	17.78	11.19	21.10	16.77	18.89	17.31
	Unclassified	45.51	48.56	45.81	34.63	48.34	45.83	47.83	47.83
Mean IoU		40.61	40.29	41.24	33.50	41.51	41.92	41.66	41.41
Accuracy		91.74	91.85	91.00	87.79	91.41	91.60	91.98	91.08
Std. Dev.		24.72	26.11	24.12	26.88	24.23	24.37	24.13	24.46

영역 손실 함수와 분포/화소별 손실 함수를 결합한 복합 손실 함수의 성능 변화를 그림 3에서 살펴보면, Lovasz의 복합 손실 함수는 Dice의 복합 손실 함수에 비하여 성능 변화가 크고, 평균 성능의 저하를 보이고 있다.

Dice 복합 손실 함수는 CE 손실 함수와의 결합보다 Focal 손실 함수와 결합하였을 때 좀 더 안정적인 성능을 보이고 있으나, 표 5에서처럼 최고 성능에서는 큰 차이를 보이지 않았다. 배치 크기에 따른 경향을 살펴보면 배치 크기는 약 16

이나 20이 적절해 보인다. 성능 변화 그래프인 그림 2, 3 그리고 표 5로부터 배치 크기가 작으면 성능 변화의 폭이 크며, 배치 크기가 큰 경우에는 수렴 속도가 늦어 더 많은 학습이 필요할 것으로 보인다.

각 손실 함수별 판별한 도로의 종류와 mIoU 경향을 표 6에 정리하였다. Dice + Focal 손실 함수의 조합은 배치 크기에 상관없이 고른 성능을 보였으나, 다른 조합의 경우에는 큰 성능의 편차를 보인다. 최고 성능만을 고려하였을 때,

Lovasz 손실 함수를 제외하고는 mIoU나 정확률은 거의 유사하였으나, 각 mIoU의 표준 편차를 살펴보면 Dice와 복합 손실 함수의 표준 편차가 단독 손실 함수보다 작아 Dice 또는 복합 손실 함수를 이용하는 것이 바람직해 보인다. 화소 또는 분포 손실 함수인 CE와 Focal 함수를 적용하였을 때 많은 비율을 차지하는 도로 종류에서는 CE, 적은 비율을 차지하는 도로 종류에서는 Focal 손실 함수를 적용하였을 때 약간의 성능 차이를 확인할 수 있었으나 학습 반복 수와 검출 데이터의 분포에 따라 분석 결과도 달라질 수 있을 것으로 보인다.

5. 요약 및 결론

본 연구에서는 인공위성 영상에서 도로를 검출하는 분할 모델의 성능에 손실 함수가 미치는 영향을 분석하였다. 실험에 사용된 데이터셋은 AI Hub에서 제공한 고해상도 KOMPSAT-3 기반 도로 검출 영상이며, NL-LinkNet 모델을 기반으로 다양한 손실 함수에 대해 실험을 수행하였다.

손실 함수에 따른 실험 결과를 분석하기 위하여 화소 기반과 영역 기반의 단일 손실 함수와 서로 다른 손실 함수를 결합한 복합 손실 함수에 따른 성능 변화를 관찰하였다.

단독 손실 함수의 성능 비교에서는 Dice Loss가 가장 안정적인 학습 경향과 높은 성능(mIoU)을 보였으며, Focal Loss는 클래스 불균형 완화 측면에서 유리하였다. 반면, Lovasz Loss는 이론적인 강점에도 불구하고 낮은 학습 안정성과 성능 저하를 보였다.

두 종류의 손실 함수를 결합한 복합 손실 함수의 성능 변화에서는 Dice + Focal 조합이 가장 안정적이고 높은 평균 성능을 기록하였으며,

Dice + CE 조합은 배치 크기가 작은 조건에서 상대적으로 나은 결과를 보였다. Lovasz 기반 복합 손실 함수는 Lovasz의 성능 변동 폭이 크고 평균 성능이 낮았지만, 화소 기반의 손실 함수와 같이 사용되었을 때 Dice 기반의 복합 손실 함수와 비교할 정도로 유사한 성능을 보였다.

배치 크기가 클수록 학습이 안정화되는 경향이 있었으며, 적절한 배치 크기(16~20)에서 성능이 극대화됨을 확인하였다. 단, 너무 큰 배치 크기에서는 학습 수렴 속도가 느려 추가적인 epoch 확보가 필요하다는 것을 추정할 수 있었다.

전체적인 실험 결과로부터 Dice 및 Dice 기반 복합 손실 함수는 높은 정확도와 낮은 mIoU 표준 편차를 보여 전체 도로 종류에 대해 안정적인 성능을 보였으며, Lovasz 손실 함수는 초기 안정적인 학습 수렴 능력과 성능을 위해서 단독으로 사용되기보다는 복합적으로 사용되는 것이 바람직해 보인다.

결론적으로, 도로 검출과 같은 세밀하고 복잡한 구조의 객체 인식 문제에서는 손실 함수의 선택이 모델의 성능을 좌우하는 중요한 요인임을 확인하였다. 특히, Dice 기반 손실 함수 및 Focal과의 조합이 클래스 불균형을 완화하고 도로 검출 성능을 안정적으로 향상하는 데 효과적이었다. 향후 연구에서는 손실 함수 최적화와 더불어 앙상블 기법, 후처리(post-processing) 기법 등과의 통합적 접근이 요구된다.

이 논문은 2024학년도 한남대학교
학술연구비 지원에 의하여 연구되었음

참고 문헌

[1] H. Lee, H. Oh, "AI Dataset for Road

- Detection using KOMPSAT Images”, Data of Geology, Ecology, Oceanography, Space Science and Polar Science 4(1), pp.43-48, 2022, DOI: <https://dx.doi.org/10.22761/DJ2022.4.1.005>
- [2] AI Hub, Satellite Image Object Detection, 2023, <https://aihub.or.kr/aihubdata/data/view.do?dataSetSn=73>
- [3] R. Azad, et al, “Loss Function in the era of semantic segmentation: A survey and outlook:”, arXiv:2312.05391, 2023, DOI: <https://doi.org/10.48550/arXiv.2312.05391>
- [4] T.-Y. Lin, P. Goyal, R. Girshick, K. He, P. Dollar, “Focal Loss for Dense Object Detection”, arXiv:1708.02002, 2017, DOI: <https://doi.org/10.48550/arXiv.1708.02002>
- [5] F. Milletari, N. Navab, S.-A. Ahmadi, “V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation”, arXiv: 1606.04797, 2016, DOI: <https://doi.org/10.48550/arXiv.1606.04797>
- [6] M. Berman, A. R. Triki and M. B. Blaschko, “The Lovasz-Softmax Loss: A Tractable Surrogate for the Optimization of the Intersection-Over-Union Measure in Neural Networks”, 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, pp.4413-4421, 2018, DOI: <https://doi.org/10.1109/CVPR.2018.00464>.
- [7] Y. Wang, J. Seo and T. Jeon, “NL-LinkNet: Toward Lighter But More Accurate Road Extraction With Nonlocal Operations”, in IEEE Geoscience and Remote Sensing Letters, vol. 19, pp.1-5, 2022, Art no. 3000105, DOI: <https://doi.org/10.1109/LGRS.2021.3050477>
- [8] Y. Wang, “NL-LinkNet for Road Extraction”, 2019, <https://github.com/yswang1717/NLLinkNet>
- [9] SI Analytics, “NL-LinkNet for Multi-class Road Extraction”, 2020, <https://github.com/SIAnalytics/nia-road-baseline>
- [10] H. Lee, D. Lee, H. Oh, “Study on Transformer-based Road Segmentation using KOMPSAT Images”, KSAS 2022 Spring Conference, pp.374-376, 2022, <https://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE11076852>

저자 소개



김나경 (Nagyeong Kim)

2019.3-현재 : 한남대학교 정보통신공학과
학·석사 통합과정
<주관심분야> 인공지능, 컴퓨터비전, 위성
영상 분석, 이미지 분할, 객체 탐지



윤영선 (Young-Sun Yun)

2001.2 KAIST 전산학과 박사
2006.4-2007.2 한국전자통신연구원
초빙 연구원
2012.8-2013.7 University of Washington
방문학자
2001.3-현재 한남대학교 교수
<주관심분야> 음성인식, 음성변환, 화자인식,
인공지능, 저작권침해, 유사도, 완성도
감정, 오픈소스 등