



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

2022년 8월

석사학위 논문

Semantic Segmentation

알고리즘을 이용한 드론 영상
기반 도로 추출 및 변화 탐지

조선대학교 대학원

토목공학과

오행열

Semantic Segmentation

알고리즘을 이용한 드론 영상
기반 도로 추출 및 변화 탐지

Road Extraction and Change Detection Based on
Drone Image Using Semantic Segmentation
Algorithm

2022년 8월 26일

조선대학교 대학원

토목공학과

오행열

Semantic Segmentation

알고리즘을 이용한 드론 영상
기반 도로 추출 및 변화 탐지

지도교수 정 명 훈

이 논문을 공학석사학위신청 논문으로 제출함

2022년 4월

조선대학교 대학원

토 목 공 학 과

오 행 열

오행열의 공학석사학위 논문을 인준함

위원장 조선대학교 교수 김성홍 (인)

위원 조선대학교 조교수 이희영 (인)

위원 조선대학교 부교수 정명훈 (인)

2022년 5월

조선대학교 대학원

목 차

ABSTRACT

제 1장. 서 론	1
제1절 연구 배경 및 목적	1
제2절 논문의 구성	3
제 2장. 선행연구	5
제1절 도로 추출	5
1. 머신러닝 방식	5
1. 딥러닝 방식	6
제2절 변화 탐지	9
제 3장. 방법론	11
제1절 Semantic Segmentation	11
1. LinkNet	11
2. D-LinkNet	12
3. NL-LinkNet	15
4. Pretrained Encoder	18

5. IoU(Intersection over Unit)	21
제2절 변화 탐지	21
1. 변화 탐지 프로세스	21
2. Confusion Matrix	22
제 4장. 실험	25
제1절 데이터 처리 및 분석 아키텍처	25
1. AIHub 위성영상 객체 판독 데이터	26
2. 광주광역시 드론 정사영상 데이터	27
가. 드론 정사영상 전처리	28
나. 데이터 라벨링	31
3. 최종 데이터 구성	33
제2절 실험 환경	34
제3절 실험 결과	35
1. 도로 추출 모델 평가 결과	35
2. 도로 추출 모델 하이퍼파라미터 최적화	37
3. 변화 탐지 결과	40
제 5장. 결 론	46

참고문헌 48

표 목 차

표 3.1 ResNet 구조	20
표 4.1 아리랑 위성 정보	27
표 4.2 드론 정사영상 정보	28
표 4.3 실험 환경	34
표 4.4 하이퍼파라미터	35
표 4.5 모델 성능 평가 결과	36
표 4.6 LinkNet-34 하이퍼파라미터	37
표 4.7 하이퍼파라미터 최적화 적용 결과	37
표 4.8 변화 탐지 성능 평가 결과	42

그림 목 차

그림 1.1 우리나라 지도 제작 과정	2
그림 2.1 Patch-level CNN 기반 모델들의 일반적인 아키텍처	7
그림 2.2 FCN 기반 모델들의 일반적인 아키텍처	8
그림 2.3 Deconvolutional network 기반 모델의 일반적인 아키텍처	9
그림 3.1 LinkNet Architecture	12
그림 3.2 D-LinkNet Architecture	13
그림 3.3 Convolution과 Dilated Convolution	14
그림 3.4 Dilated Convolutional Architecture in D-LinkNet	15
그림 3.5 Non-Local Attention Module	17
그림 3.6 NL-LinkNet Architecture	18
그림 3.7 신경망 구조 비교	19
그림 3.8 IoU	21
그림 3.9 변화 탐지 프로세스	22
그림 3.10 Confusion Matrix	23
그림 4.1 시스템 아키텍처	25
그림 4.2 AIHub 데이터	26
그림 4.3 드론 정사영상 전경	28
그림 4.4 그리드 생성 예시	29
그림 4.5 그리드를 활용한 도시철도 공사현장 추출 예시	30
그림 4.6 해상도 비교	31
그림 4.7 Label Studio 라벨링 예시	32

그림 4.8 드론 정사영상 데이터 생성 예시 32
 그림 4.9 드론 정사영상 학습데이터 : 유촌동-백운교차로 구간 33
 그림 4.10 드론 정사영상 평데이터 : 백운교차로-광주역 구간 34
 그림 4.11 각 모델 추론 결과 비교 36
 그림 4.12 추론 결과 비교1 38
 그림 4.13 추론 결과 비교2 39
 그림 4.14 차분 예시 41
 그림 4.15 Threshold 0.25에서의 Window 크기 별 변화 탐지 결과1 44
 그림 4.16 Threshold 0.25에서의 Window 크기 별 변화 탐지 결과2 45

ABSTRACT

Road Extraction and Change Detection Based on Drone Image Using Semantic Segmentation Algorithm

Oh, Haeng Yeol

Advisor : Prof. Jeong, Myeong-Hun, Ph.D

Department of Civil Engineering

Graduate School of Chosun University

Cities are becoming more complex due to rapid industrialization and population growth in modern times. In particular, urban areas are rapidly changing due to housing site development, reconstruction, and demolition. Thus accurate change information is necessary for various purposes, such as High Definition Map for autonomous car driving. In the case of Korea, accurate spatial information can be generated by making a map through the existing map production process. However, when targeting a large area, it has a limitation in that it takes a lot of time and money. Road, one of the map elements, is a hub and essential means of transportation that provides many different resources for human civilization. Road information is used for traffic management, urban planning, road monitoring, GPS navigation, and map updates. Therefore, it is important to update road change information accurately and quickly. To achieve this goal, this study uses Semantic Segmentation algorithms LinkNet, D-LinkNet, and NL-LinkNet to extract roads from drone images and then apply hyperparameter optimization to models with the highest performance. As a result, the LinkNet model using pre-trained ResNet-34 as the encoder achieved 85.125 mIoU. After detecting roads, change detection was

performed by comparing the original road lines in 2019 with the results of road extraction inference in the drone image in 2021. Window and threshold were applied for precise change detection, and as a result, F1-score showed the highest performance with 0.75966 when a window size of 16×16 and threshold 25% was applied. Subsequent studies will be conducted to compare the results of this study with those of studies using state-of-the-art object detection algorithms or semi-supervised learning-based Semantic Segmentation techniques. The results of this study are expected to be used to improve the speed of the existing map update process, which was carried out manually for the changed area, through speedy detection of the changed area.

제 1장. 서 론

제1절 연구 배경 및 목적

현대에 들어 급속한 산업화와 인구 증가로 인하여 도시가 점점 복잡해지고, 도로의 신설과 개발도 빠르게 확장되고 있다(운부열 등, 2006). 특히 도심지역의 경우 택지 개발과 재건축, 철거 등으로 인해 빠르게 변화하는 지역에 해당하기 때문에 빠르고 정확한 변화 발생 정보가 요구되고 있다(장영재 et al., 2020). 또한 최근 자율주행 분야가 부상하며 자율주행에 필수적인 정밀도로지도의 구축에 대한 관련 기술 연구 및 개발 지원 필요성이 강조됨에 따라(NGII, 2019) 이를 위해 국토부는 2018년 자율주행 지원을 위한 도로 변화 신속 탐지, 갱신 기술 개발 및 실증화 연구사업을 통해 관련 국내 기술 개발 중이고, TomTom과 Deepmap 과 같은 지도 관련 기업들은 차량 내 지도 정보를 업데이트하는 기술들을 개발 중이다(박유경 등, 2019).

현재 우리나라에서의 지도 갱신은 그림 1.1과 같이 항공 또는 위성사진 촬영을 통해 지표면의 정보를 수집한 후 도화 작업에 필요한 좌표를 현장 측량 작업(지상기준점 측량, 사진기준점 측량)을 통해 얻는다. 이후 수치 도화 작업으로 지표면에 존재하는 지형, 지물에 대한 정보를 생성하고 지리 지명 조사, 정위치 편집 작업 후 지리정보의 전산화를 거쳐 최종적으로 국가 지리정보 Data Base(DB)에 구축된다.

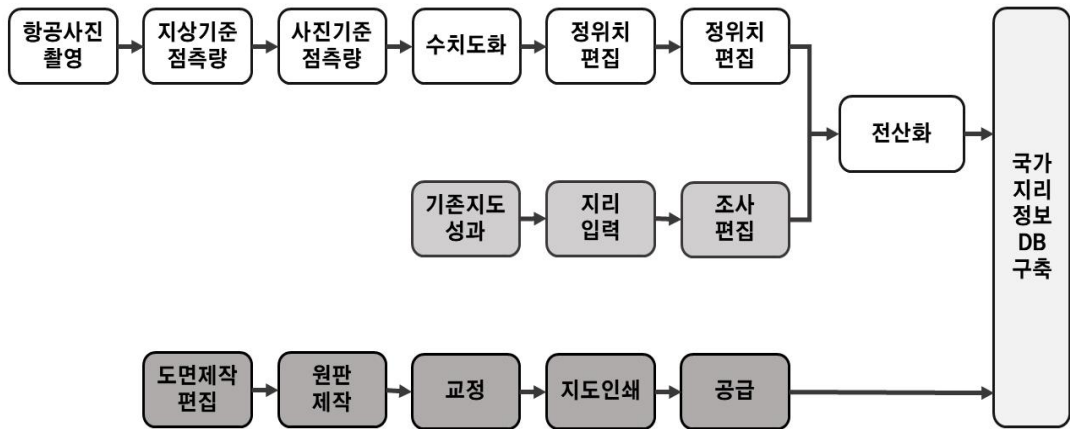


그림 1.1 우리나라 지도 제작 과정 (국토교통부, 2022)

기존 지도 제작 과정을 통해 지도를 구축하면 정밀한 공간정보를 생성할 수 있으나 대상 지역이 넓을 경우, 많은 인력과 비용이 소요되는 한계를 가진다(김지영 등, 2013). 또한 소규모 시설 공사에 대한 정보는 등재되지 않기에 빠르게 변화하는 도로 정보를 수집함에 한계가 존재하며 이에 따라 지도 갱신에 어려움이 있다(NGII, 2018). 현재 지도의 변화된 부분을 반영하기 위해 수시 수정이 이루어지지만 대부분 현장 조사 및 측량과 같은 수작업을 통해 이루어지므로 수정 작업의 효율성이 다소 떨어진다(주영은 and 이형석, 2005).

이러한 문제점을 해결하기 위하여 원격 탐사 분야에선 위성영상, 항공영상과 같은 고해상도의 정사영상에서의 풍부한 정보를 이용하여 도로와 같은 지표면의 인공구조물 혹은 피복의 변화를 탐지하는 연구가 이루어지고 있다(Bruzzone and Bovolo, 2012, Coppin et al., 2004, Jin et al., 2013, Singh, 1989, 오행열 등, 2022). 이중 특히 도로의 경우 인간 문명을 위한 많은 다양한 자원을 제공하는 교통의 중추이자 필수적인 수단에 해당하며(Abdollahi et al., 2020) 도로에 대한 정보는 교통 관리, 도시 계획, 도로 모니터링, GPS 내비게이션 및 지도 업데이트 등에 활용되기에 매우 중요하다(Shi et al., 2013). 이에 따라 정사영상에서 도로의 여러 특징을 활용하여 도로를 추출 혹은 변화를 탐지하는 연구가 이루어졌다. 대표적으로 도로 추출을 위해 texture progressive analysis(Mena and Malpica, 2005)와 mathematical morphology(Zhu et al., 2005)같은 분석 방법으로 도로 구간에 대한 특정 정보를 사용하여 원격 탐사 이미지에서의 도로를 추출하였다. 하지만 이러한

접근 방식은 딥러닝과 같은 데이터 기반 분석 방법에 비해 다양한 유형의 도로를 처리함에 있어 한계를 가진다. 또한 도로의 분광적 특성을 이용한 방법 (Panboonyuen et al., 2017)의 경우 원격 탐사 이미지의 도로 색상이 둘 이상의 색상으로 이루어질 때 도로 추출 및 분류가 제대로 수행되지 않는다. 따라서 도로를 정확히 추출하기 위해선 도로의 다양한 특성들을 함께 고려하여 분석하여야 한다. 이와 같은 문제점을 해결한 방법이 데이터 기반의 분석 방법인 딥러닝 기법이다. 본 연구에선 딥러닝 기법 중 이미지의 모든 화소를 분류하는 Semantic Segmentation 모델인 LinkNet(Chaurasia and Culurciello, 2017), D-LinkNet(Zhou et al., 2018), NL-LinkNet(Wang et al., 2021)을 사용하여 도로를 추출한 후 모델의 도로 추론 결과를 활용하여 도로의 변화를 탐지하였다.

본 연구는 우리나라 위성인 아리랑 3호, 3A호의 위성영상으로 제작된 AIHub 위성영상 객체 판독 데이터와 2019년 및 2021년 광주광역시 드론 정사영상 데이터를 기반으로 정사영상 이미지에서의 도로를 Semantic Segmentation 모델들로 도로를 추출한 후 가장 높은 성능의 모델 추론 결과에 차분 기법을 적용하여 도로의 변화를 탐지하는 것을 목적으로 한다. 연구의 결과는 변화된 지역들의 자동 탐지와 클러스터링을 통해 현장 측량 작업 할당 구역을 지정할 수 있고 이를 실제 지도 데이터를 신속하게 갱신할 수 있는 프로세스에 도움을 줄 수 있을 것으로 기대된다.

제2절 논문의 구성

본 연구는 기존 정사영상을 이용하여 도로 추출 및 변화 탐지를 위해 제시되었던 연구들의 문헌을 검토하며, Semantic Segmentation 알고리즘인 LinkNet, D-LinkNet, NL-LinkNet 중 드론 정사영상에서의 도로 추출 성능이 가장 뛰어난 모델의 추론 결과를 활용하여 도심에서의 도로 변화 탐지 방법을 제시하고자 한다. 본 논문은 총 5장으로 구성된다.

위 1장의 서론에선 신속한 지도 갱신에 대한 중요성에 관하여 서술하였다. 이후 2장의 문헌 검토에서 기존 연구들의 조사와 이론적 고찰을 통해 본 연구의 중요성과 필요성에 대하여 논의하고, 3장에선 도로 추출에 사용된 알고리즘과 도로 변화탐지의 전체적인 프로세스 및 각 방법론의 평가 방법을 제시한다. 4장 실험에선 본 연구에 사용된 데이터에 대한 설명과 함께 드론 정사영상의 전처리 과정 및

라벨링된 데이터의 시각화를 통해 데이터에 대한 전반적인 설명을 한다. 또한 본 연구의 실험 환경과 도로 추출 및 변화 탐지 결과의 시각화와 평가 결과를 제시한다. 마지막으로 5장 결론에선 본 연구의 전체적 요약 및 결과들과 함께 연구가 가지는 한계점 및 향후 연구 과제에 대하여 제시한다.

제 2장. 선행연구

정사영상에서의 도로 추출 분야는 크게 머신러닝 방식과 딥러닝 방식으로 나뉜다. 본 연구에서는 도로 추출 분야와 변화 탐지 분야에서의 각 분석 방식에 관한 선행 연구들을 조사하였다.

제1절 도로 추출

도로의 경우 폭이 좁고 교차로를 가지는 기하학적 특성, 비도로 지역과의 색깔이 다른 분광적 특성, 끊기지 않고 서로 연결되는 위상적 특성, 현실 세계에서 제약 조건에 의한 특정한 조건을 가지는 functional 특성(Vosselman and Knecht, 1995), 흑백영역에서의 공간적 특성을 가지는 texture 특성을 (Wang et al., 2014) 가진다. 기존 머신러닝 방식의 경우 이러한 특성 중 몇 가지를 조합하여 정사영상에서의 도로를 추출하였으며, 먼저 머신러닝 방식을 활용한 정사영상에서의 도로 추출 연구를 살펴본다.

1. 머신러닝 방식

도로 추출 방식은 크기 지도 학습(supervised learning)과 비지도 학습(unsupervised learning)으로 구분한다. 지도 학습은 일반적으로 객체의 범주가 지정된 대용량의 이미지 데이터를 통해 모델을 학습하는 방법이다. 딥러닝 분야가 빠르게 발전하기 이전엔 양질의 데이터가 충분하지 않아 비지도 학습 기반의 연구가 활발하게 진행되었다.

Miao et al.(2014)은 고해상도의 위성영상에서의 도로 추출을 위해 semi-automatic method에 해당하는 mean shift 기법의 적용을 제안하였다. 해당 연구는 사용자가 미리 정의한 지점들을 연결하기 위해 측지선 방법을 적용하여 초기 도로 윤곽을 추출하고 임계값을 적용하여 도로를 비도로 지역과 분리하였다. 이후 측지선 방법을 다시 적용하여 연구자가 정의한 지점을 연결하여 커널 밀도 추정 맵을 생성하였다. 이러한 방법은 빠르고 효율적이지만 사용자가 수동으로 지점을 정의해야 한다는 단점을 가진다.

도로 추출을 위해 그래프 이론을 적용한 연구로 Tao and Jin(2007)은 배경과 도

로를 구분하기 위한 기준으로 graph division method를 제안하였다. 해당 연구는 이미지 픽셀 간의 관계를 설명하기 위해 회색 영역을 기반한 weight matrix를 사용했다. Liu and Wang(2008)은 그래프 이론을 기반한 interactive image Segmentation method를 제안하였다. 이는 이미지 texture, color, edge의 확률 모델을 빠르게 얻을 수 있음을 보여주었다. Unsalan and Sirmacek(2012)은 확률론과 그래프 이론을 기반으로 도로 중심선 추출 모듈, 도로 형상 추출 모듈, 도로 네트워크 형성 모듈을 제안하였다. 그래프 이론을 기반한 도로 추출 방법은 복잡한 데이터 유형을 처리하고 도로의 전역적 특징을 제대로 추출할 수 있는 장점이 있다. 하지만 대부분 스펙트럼 행렬 개념을 채택하고 유사 행렬의 고윳값과 고유 벡터를 사용하여 이미지 분할을 수행하기 때문에 고유 벡터의 수가 증가함에 따라 모델이 복잡해지는 단점이 있다.

Mathematical morphology는 도로의 기하학적 특성을 다룬다. Zhang et al.(1999)은 이미지 전처리 단계에서 이미지에서 도로망을 분할 한 후 morphological trivial opening 연산을 적용하였다. 이 방법은 도로의 선형적 요소를 유지하고 거의 모든 영역에서 Noise를 제거하였다. Ma et al.(2012)은 canny edge detection operator를 통해 도로의 윤곽을 추출하고 hough line transform과 morphological operator를 통해 도로의 중심선과 분기점 그리고 도로의 끝점을 탐지하였다. Awad(2013)는 Fuzzy c-Means(FCM)과 Self Organizing Map(SOM)을 결합한 FCMs-SOM 알고리즘으로 이미지에서의 도로를 분류하고 분류 결과에서의 이질성을 줄이기 위해 morphological 방법인 dilation과 erosion을 적용하였다. 영상 분할 결과는 구조 요소(모양 및 크기)의 선택에 따라 결과에 큰 영향을 주기 때문에 구조 요소의 불확실성으로 인해 높은 정확도와 좋은 추출 결과를 얻기 위하여 Mathematical morphology만을 사용하기 어렵다는 단점이 있다.

2. 딥러닝 방식

앞서 설명한 머신러닝 방식으로 도로를 추출한 연구들의 경우 도로의 일부 특성만을 고려하기 때문에 폭, 길이, 형태 등 다양한 유형의 도로를 추출하는데 한계를 가진다. 특히 고해상도의 원격 탐사 이미지는 일반적인 이미지에 비하여 이미지에 포함된 정보가 풍부하기 때문에 모델이 이러한 복잡성을 제대로 반영하지 못하는 문제가 있다(Abdollahi et al., 2020). 따라서 기존의 접근 방식은 데이터 기반

접근법에 비하여 그 성능이 상대적으로 떨어지게 된다.

최근 원격 탐사 이미지에서의 도로 추출을 위해 딥러닝 컴퓨터 비전 분야에서 영상 분할을 목적으로 하는 Segmantic Segmentation 방법론이 많이 활용되고 있다. Semantic Segmentation이란 이미지를 픽셀 단위로 어떤 범주에 해당하는지 분류함을 의미한다. 도로 추출에 사용된 Semantic Segmentation 모델들은 크게 다음과 같이 분류된다.

먼저 Patch-based Convolutional Neural Network(Patch-based CNN)(Mnih and Hinton, 2010)을 활용한 도로 추출 방법이다. 이는 다양한 공간 특징을 추출하기 위하여 이미지를 여러 Patch로 나누어 합성곱(convolution) 연산과 Pooling 연산 등을 통해 각 픽셀들의 범주를 분류하는 구조를 가진다(그림 2.1). Alshehhi et al.(2017)는 원격 탐사 이미지에서 도로와 건물을 동시에 추출하기 위하여 Patch-based CNN 구조를 적용하였으며 이를 통해 도로와 건물을 추출하였다. Wei et al.(2017)는 항공 영상에서 도로를 추출하기 위해 Visual Geometry Group(VGG) 네트워크와 deconvolutional layer 및 Fusion Layer가 결합된 구조의 모델을 제안하였다. Patch-based CNN의 경우 하나의 이미지를 여러 Patch로 나눔으로 인해 인접한 Pixel 간의 공간적 상관성을 잃어버린다. 이에 따라 객체들의 경계가 선형적이지 않거나 모호해짐으로 인해 잘못 분류된 영역들을 줄이기 위한 별도의 후처리 과정이 필요하다는 한계를 가진다.

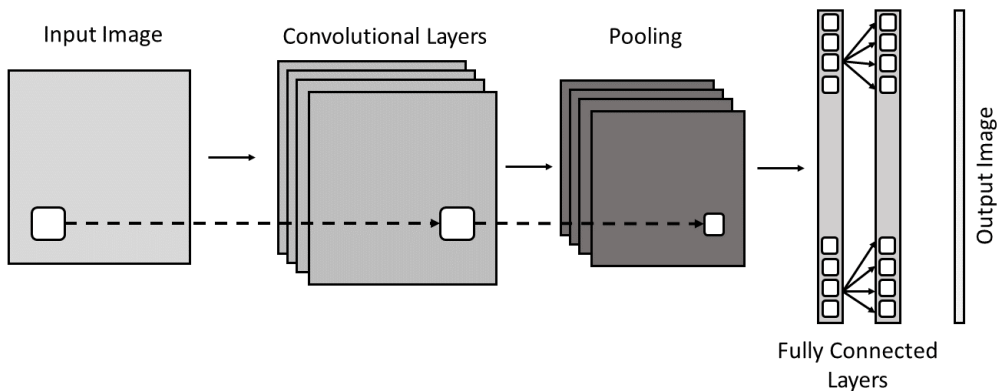


그림 2.1 Patch-level CNN 기반 모델들의 일반적인 아키텍처(Abdollahi et al., 2020)

다음은 Fully Convolutional Network(FCN)(Long et al., 2015)를 기반한 도로 추출 방법이다. FCN은 기존 이미지 분류에서 높은 성능을 보인 CNN 기반 모델들로 이미지를 학습한 후 Up-sampling을 통해 입력 이미지와 동일한 크기의 출력 이미지를 생성하여 Segmantic Segmentation을 수행하는 방법이다. 그림 2.2는 FCN의 구조를 보여준다. Varia et al.(2018) 초고해상도의 Unmanned Aerial Vehicle(UAV) 이미지에서의 FCN-32 모델을 통해 도로를 추출하였다. 이 연구의 경우 도로가 아닌 지역 중 복잡성이 높은 특정 영역을 도로로 잘못 분류하는 결과를 보였다. Wang et al.(2015)은 SAR 이미지에서의 도로 추출을 위해 FCN 구조가 효과적임을 입증하기 위해 FCN-8 모델을 사용하였지만 교차로와 같은 연결 부위를 제대로 추출하지 못하였다. FCN의 경우 Up-sampling 과정에서 발생하는 공간 정보의 손실을 매우기 위하여 보간법을 활용하기 때문에 모델의 정밀도가 낮아진다는 단점을 가진다.

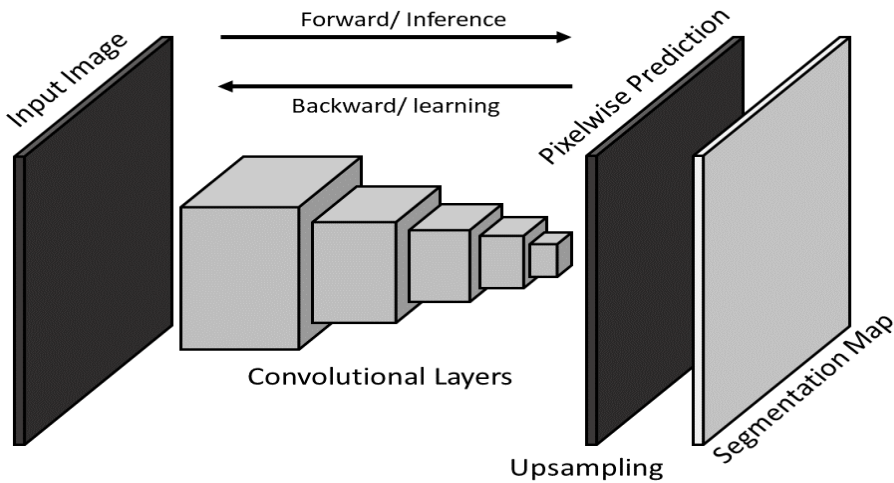


그림 2.2 FCN 기반 모델들의 일반적인 아키텍처(Abdollahi et al., 2020)

Deconvolution 구조는 일반적으로 Encoder와 Decoder의 네트워크로 구성된 구조를 가지고 있으며 Encoder에서 이미지의 특징을 뽑아내고 이를 Decoder가 받아들여 FCN과 같이 입력 이미지와 동일한 크기의 출력 이미지를 생성한다. 그림 2.3은 일반적인 Deconvolution 모델의 구조를 보여준다. Deconvolution 구조는 Patch-based CNN과 FCN 모델들보다 높은 성능을 보이지만 많은 용량의 메모리

와 높은 컴퓨팅 프로세스가 필요하다는 단점이 있다(Abdollahi et al., 2020).

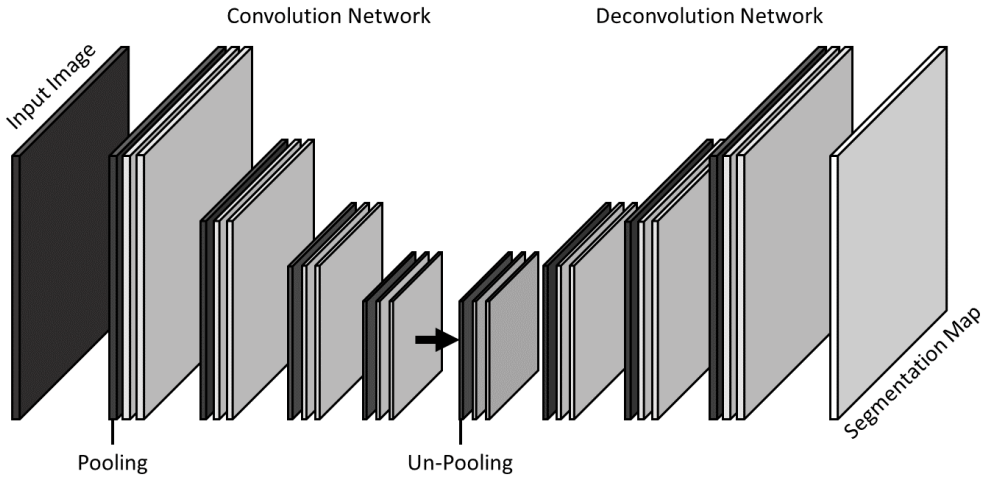


그림 2.3 Deconvolutional network 기반 모델의 일반적인 아키텍처(Abdollahi et al., 2020)

Deconvolution 구조의 단점인 모델 파라미터 수를 줄이면서 성능을 높이기 위해 개발된 딥러닝 모델이 LinkNet이다. LinkNet의 효율적인 구조를 활용하여 여러 Segmantic Segmentation 모델들이 개발되었으며 이중 도로 추출을 목적으로 만들어진 딥러닝 모델이 D-LinkNet과 NL-LinkNet이다. 본 연구에서는 Deconvolution 구조의 단점인 파라미터의 수를 줄이면서 도로를 높은 정확도로 추출할 수 있는 LinkNet, D-LinkNet 및 NL-LinkNet 모델을 광주광역시 드론 정사영상에서의 도로 추출 및 변화 탐지에 사용하였다.

제2절 변화 탐지

변화 탐지란 동일한 지역의 여러 시기의 이미지를 사용하여 일정 기간에 발생하는 지구 표면의 변화를 식별하는 프로세스를 의미한다. 변화 탐지는 탐지 단위에 따라 크게 화소 기반 변화 탐지와 객체 기반 변화 탐지로 구분한다(Cleve et al., 2008). 화소 기반의 변화 탐지는 서로 다른 시기의 동일한 위치의 화소를 비교하여 변화를 탐지하는 방법이며 대표적으로 차분(difference) 기법, 비율(ratio) 기법, 분

광 변환(image transformation), Principal component analysis(PCA), Multivariate Alteration Detection(MAD), Iteratively Reweighted-Multivariate Alteration Detection(IR-MAD) 등이 있다. 고해상도의 정사영상에선 건물, 도로와 같은 고주파 영역이 산발적으로 분포하기에 변화 탐지를 화소 기반으로 수행할 경우, 미변화 지역들이 변화 지역으로 잘못 탐지될 수 있는 문제가 존재한다(송창우 등, 2020). 따라서 작은 객체 혹은 저해상도의 영상에선 화소 기반 변화 탐지가 주로 적용되며 고해상도 영상에는 화소들을 객체로 분류하여 변화를 탐지하는 객체 기반 변화 탐지 기법이 주로 사용되었다(Hussain et al., 2013). 하지만 객체 기반 변화 탐지 또한 영상의 해상도가 높아질수록 영상 전처리 결과에 민감하게 반응하여 탐지 정확도가 떨어지는 한계를 가진다(Im et al., 2008). 이러한 문제점들을 해결하기 위해 FCM(Shi et al., 2016), Support Vector Machine(SVM)(Habib et al., 2009)과 같은 머신러닝 알고리즘이 변화 탐지에 적용되었지만 미변화 지역에 과적합되거나 변화 특징을 탐지하기 위해 모델을 정밀하게 설계해야 하는 한계를 가진다.

최근 딥러닝 분야의 발전으로 원격 탐사 분야에서 정사영상에 Semantic Segmentation 알고리즘을 적용한 변화 탐지 연구들이 이루어지고 있다. Wiratama et al. (2020)은 Feature-level U-Net 모델을 통해 여러 시기의 고해상도 다분광 이미지에서 피복의 변화를 탐지하였으며 기하학적 왜곡과 노이즈가 있는 이미지에서도 기존 변화 탐지 알고리즘을 능가함을 보였다. Moustafa et al.(2021)은 U-Net과 변형된 U-Net 모델들의 초분광 원격 탐지 이미지에서의 변화 탐지 결과 비교를 통해 Attention Recurrent residual UNet 모델이 가장 높은 성능을 보임을 검증하였다. Liu et al.(2020)은 가중치를 공유하는 Siamese network와 Encoder-Decoder 구조가 결합된 Dual-Task Constrained Deep Siamese Convolutional Network(DTCDCSCN)을 통해 이미지에서의 건물 윤곽 추출과 변화 탐지를 동시에 수행하였다. 이처럼 딥러닝 기반의 변화 탐지를 위해 여러 연구에서 encoder-decoder 구조의 Semantic Segmentation 모델을 이용하였다.

이에 따라 앞 절에서 설명한 Deconvolution 구조인 LinkNet, D-LinkNet, NL-LinkNet 모델을 통해 여러 시기의 드론 정사영상에서의 도로를 화소 단위로 추출하여 이진영상을 생성한 후 두 이미지에 차분 기법을 적용하여 도로의 변화를 탐지하였다.

제 3장. 방법론

제1절 Semantic Segmentation

Semantic Segmentation은 이미지의 모든 픽셀에 범주를 지정하는 것을 포함하므로 공간 정보를 유지하는 것이 가장 중요하며 Deconvolution 구조의 모델은 공간 정보를 유지함과 동시에 이미지의 객체를 높은 정확도로 추출할 수 있다. 본 절에선 Deconvolution 구조의 단점인 모델 파라미터의 수를 줄이면서 성능을 높인 Semantic Segmentation 알고리즘에 해당하는 LinkNet과 LinkNet을 기반으로 효율적인 도로 추출을 위해 개발된 D-LinkNet 및 NL-LinkNet에 대하여 설명한다.

1. LinkNet

드론 정사영상에서 도로의 변화를 탐지하기 위해선 서로 다른 시간대의 영상에서의 도로를 추출하여 비교해야 한다. 이를 위해 Chaurasia and Culurciello(2017)이 제안한 딥러닝 기반 Semantic Segmentation 기법인 LinkNet을 사용하였다. LinkNet은 기존 Segmantic Segmentation 알고리즘들의 단점인 느린 속도를 개선하기 위해 신경망의 매개변수를 크게 늘리지 않는 동시에 정확도를 유지함으로써 제한된 GPU 메모리를 효율적으로 사용할 수 있고 데이터를 빠르게 학습할 수 있는 기법으로 알려져 LinkNet의 구조를 기반으로 다양한 모델들이 개발되고 있다. 본 논문에서 사용한 LinkNet은 그림 3.1과 같이 그림 왼쪽의 Encoder Block과 그림 오른쪽의 Decoder Block을 통해 이미지를 학습하고 각 픽셀의 범주를 분류한다. LinkNet에서의 Encoder Block은 이미지의 Feature Map 정보를 얻기 위한 단계로 사전 학습된 ResNet을 통해 이미지를 학습하게 된다. Decoder Block은 Encoder Block의 연산으로 줄어든 이미지의 크기를 Up-sampling을 통해 Input 이미지의 크기와 동일한 이미지를 생성하게 된다. 객체의 공간정보를 유지하기 위해 Encoder Block에서 생성되는 Feature map이 Skip Connection을 통해 Decoder Block으로 전달되며 이를 통해 Decoder Block에서의 Up-sampling으로 생기는 공간 정보의 손실을 복구하게 된다.

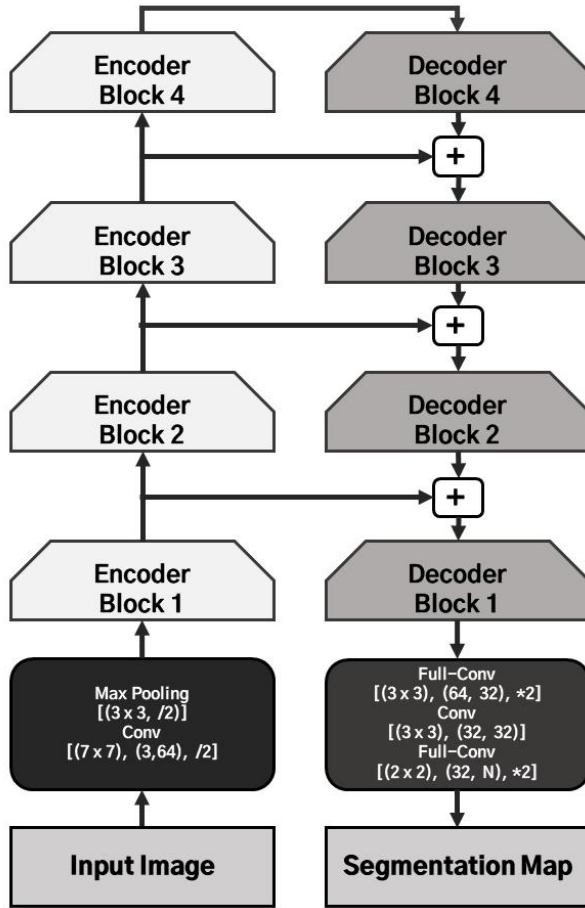


그림 3.1 LinkNet Architecture (Chaurasia and Culurciello, 2017)

2. D-LinkNet

본 연구에 LinkNet 모델과 함께 사용된 D-LinkNet은 Deepglobe Road Extraction Challenge(Demir et al., 2018)에서 우승한 모델로 높은 해상도의 위성영상에서 도로 검출을 목적으로 설계된 모델에 해당한다. Zhou et al.(2018)은 위성영상에서의 도로는 서로 연결되어 있고 얇고 복잡한 형상을 가지지만 이미지 전체에서 작은 부분만을 차지하며 특히 고해상도의 이미지에선 이러한 특징이 도드라지게 나타나기 때문에 모델의 network가 큰 Receptive Field를 가질 수 있어야만 이

이미지 전체를 효율적으로 학습할 수 있다고 한다. 따라서 D-LinkNet은 이러한 도로의 특성을 잘 반영하기 위하여 사전 학습된 Encoder로 구성된 LinkNet에 Dilated Convolution Layer들이 중앙에 추가된 구조를 가진다(그림 3.2).

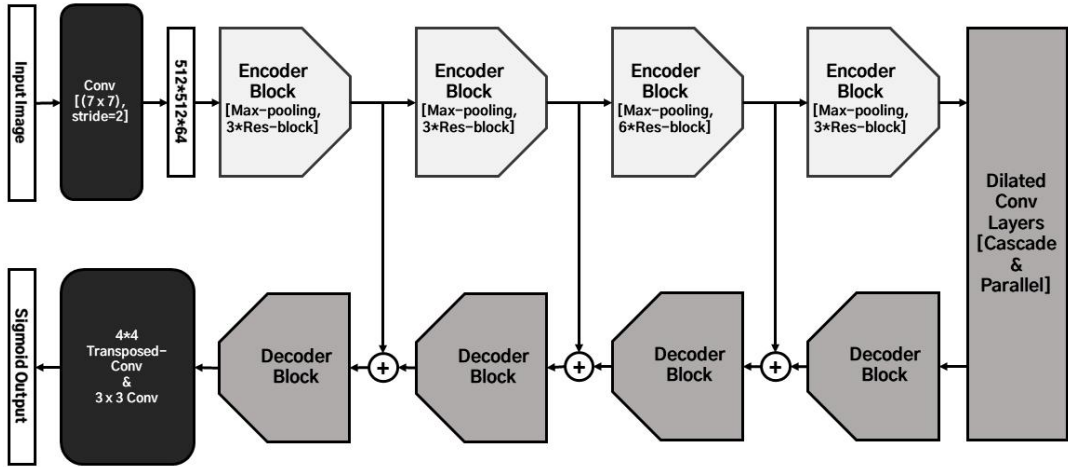


그림 3.2 D-LinkNet Architecture (Zhou et al., 2018)

기존 Convolution은 그림 3.3(a)와 같이 6×6 크기의 2D 이미지를 학습하기 위한 짙은 회색 부분에 해당하는 Kernel의 크기가 3×3일 경우, Stride를 1로 설정하게 되면 Kernel이 한 칸씩 이동하며 이미지를 처리하게 된다. 이는 연산에 있어 인접 화소와의 공간적 관계만 고려되기에 고해상도 정사영상 전반에 걸쳐 존재하는 도로 특성을 고려하기 어렵다. 따라서 도로를 제대로 추출하기 위해선 Convolution의 Kernel의 크기를 키워 Receptive Field의 크기를 늘리거나 여러 개의 Convolution을 사용하여 layer를 추가해야 한다. 하지만 이러한 방법을 고해상도의 정사영상에 적용하게 되면 학습에 필요한 연산량이 증가하게 되는 단점이 생긴다. 특히 Segmentation 시 단순히 Pooling 및 Convolution 연산 후 Up-sampling을 하게 된다면 Output 이미지에서의 객체 해상도가 떨어지게 되어 공간적 정보의 손실이 생긴다(Chen et al., 2017). 고해상도의 정사영상에서 도로 추출 시 이러한 문제점이 생길 수 있기 때문에 D-LinkNet에선 Dilated Convolution이 사용된다. Dilated Convolution은 그림 3.3(b)와 같이 기존 Convolution filter가 수용하는 픽셀 사이에 zero padding을 추가하여 강제로 Receptive Field를 늘리는 방법이다. 이는 일반적인 Convolution과 동일한 연산 비용으로 더 넓은 크기의 Receptive Field로

데이터를 처리할 수 있기 때문에 고해상도의 정사영상에 적용하기 적합하다.

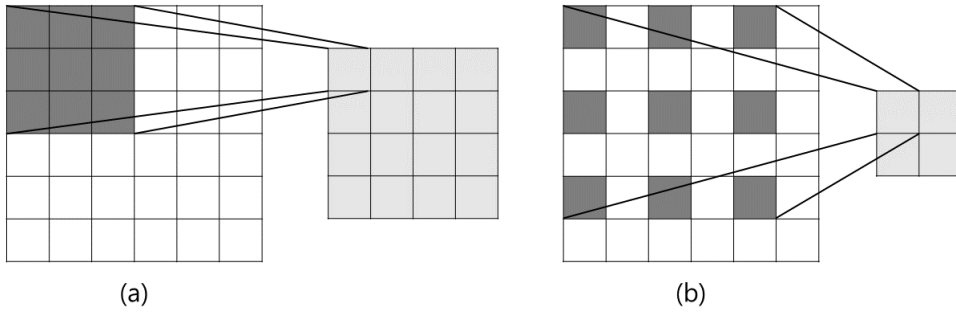


그림 3.3 Convolution과 Dilated Convolution

Dilated Convolution은 크게 Cascade 방식(Yu and Koltun, 2015)과 Parallel 방식(Chen et al., 2018)으로 사용된다. D-LinkNet에선 그림 3.4와 같이 1, 2, 4, 8의 dilation을 가지는 Dilated Convolution을 그림 오른쪽 방향으로 계단식의 cascade 구조를 통해 학습이 되며 multi-resolution feature의 이점을 활용하기 위하여 cascade 구조에 4개의 parallel 학습 결과를 더하는 구조를 통해 Dilated Convolution을 사용한다.

D-LinkNet의 Decoder 부분은 계산의 효율을 위해 LinkNet의 방법과 동일하게 적용된다. Decoder 및 Dilation의 결과물인 32×32 크기의 Feature map을 Input Image의 크기와 동일한 1024×1024 크기로 복원하기 위하여 Transposed Convolution Layer를 통해 Up-sampling이 이루어진다.

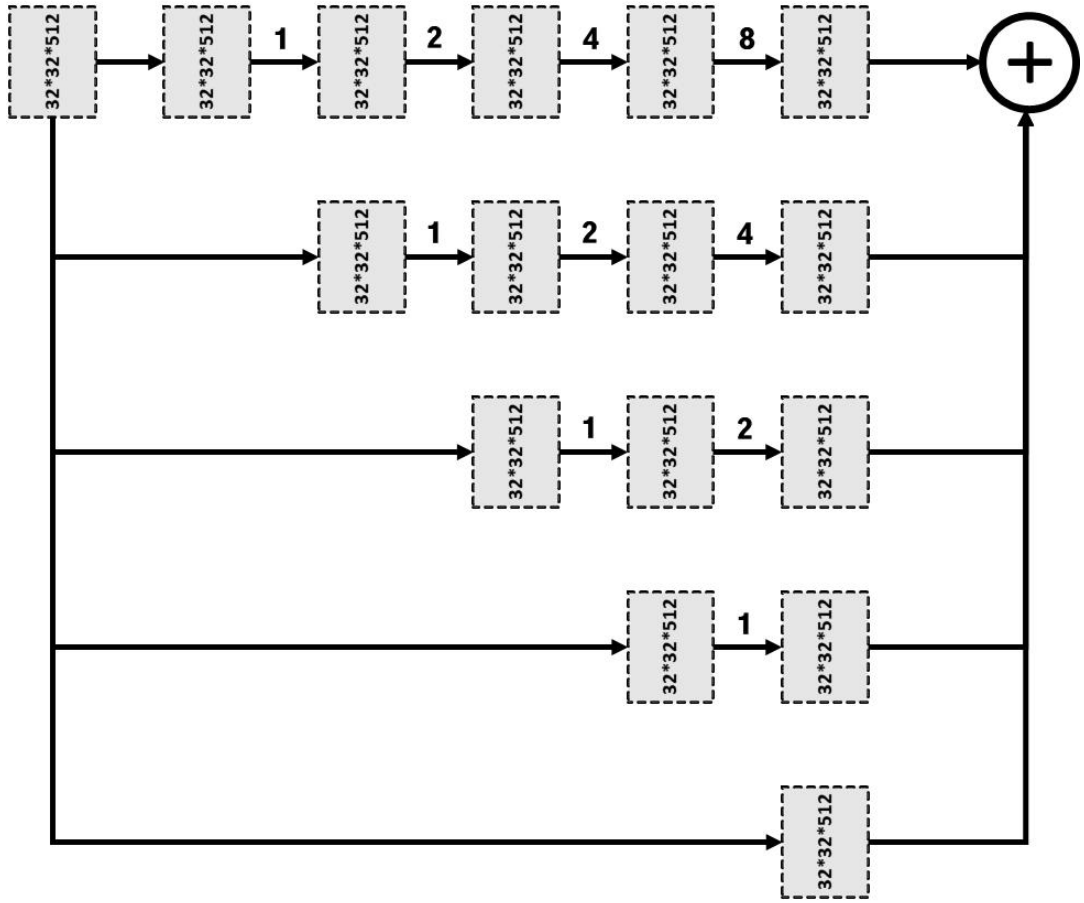


그림 3.4 Dilated Convolutional Architecture in D-LinkNet (Zhou et al., 2018)

3. NL-LinkNet

NL-LinkNet은 신경망의 매개변수를 줄이는 동시에 고해상도의 위성영상에서 도로를 효율적으로 추출하기 위해 기존 LinkNet의 구조에 Non-Local Attention Module을 삽입한 모델에 해당한다. Computer vision 분야에서의 Convolution은 Kernel의 크기만큼 이미지를 학습하기 때문에 Local Operator에 해당한다. 이러한 Local Operator의 경우 큰 Receptive Field를 가지기 위하여 여러 개의 Convolution Layer를 쌓거나 혹은 Pooling Layer로 이미지의 크기를 줄여 한 번에 볼 수 있는 영역을 넓히는 방법을 사용해야 한다. 이러한 방법을 사용하게 되면 Layer의 수가 늘어나 연산 비용이 늘어나며 모델을 학습시키기 어려워진다. 특히 고해상도 정사

영상에서의 도로는 서로 멀리 떨어져 있거나 긴 범위에 걸쳐 이미지에 존재하기 때문에 Local Operator로 도로 추출에 사용한다면 객체를 올바르게 검출하지 못할 수도 있다. 따라서 도로의 특성을 잘 반영하기 위하여 Non-Local Block(Wang et al., 2018)을 LinkNet 구조에 삽입한 NL-LinkNet을 제안하였다(Wang et al., 2021).

NL-LinkNet에 사용된 Non-Local Attention Module은 Non-Local Operation을 기반으로 연산된다. Non-Local Operation은 식 1과 같이 계산되며 y 는 출력값, x 는 입력값, i 는 출력 위치, j 는 모든 위치를 열거하는 index, 함수 f 는 pairwise function, 함수 C 는 normalize 함수를 의미한다. 해당 수식은 출력값 y 의 위치 i 는 입력값 x 의 위치 i 와 다른 모든 위치인 j 들에 대하여 계산된 결과를 의미한다. 이를 통해 CNN과 같이 근처 몇 개의 화소 정보나, RNN과 같이 시간 축으로 몇 개만을 보는 것이 아닌 모든 위치를 고려할 수 있는 연산이다. 이를 통해 다른 위치 간의 관계를 학습할 수 있게 된다.

$$y_i = \frac{1}{C(x)} \sum_j f(x_i, y_i)g(x_j) \quad (1)$$

Non-Local Attention Module의 경우 다른 구조에도 쉽게 더해질 수 있도록 Non-Local Operation을 Block 형태로 감싼 구조로 되어있으며 그림 3.5에서 해당 구조를 보여준다.

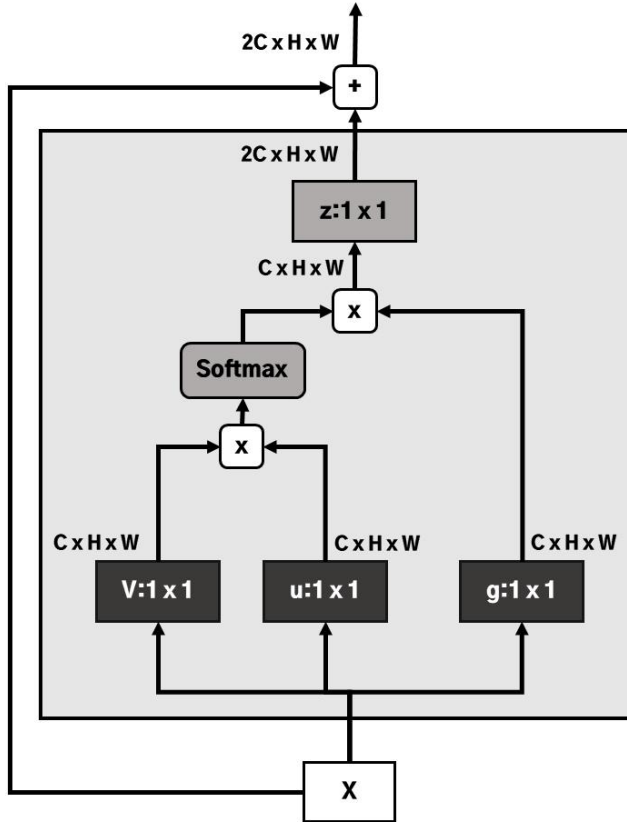


그림 3.5 Non-Local Attention Module (Wang et al., 2021)

Non-Local Attention Module은 식 2와 같이 계산되는데 식에서의 z 는 출력값, y 는 식 1에서의 y , x 는 residual connection을 의미한다. 이 residual connection이 추가됨으로 인해 Non-Local Attention Module이 이미 학습된 모델이라도 중간에 추가하여 사용될 수 있게 된다.

$$z_i = W_z y_i + x_i \quad (2)$$

NL-LinkNet은 앞서 설명한 Non-Local Attention Module을 3번째 Encoder와 4번째 Encoder의 뒤에 삽입한 구조를 가진다(그림 3.6). NL-LinkNet은 D-LinkNet과 동일한 Pretrained Encoder Block과 Decoder Block을 사용하며 각 Encoder Block에서 생성된 Feature Map은 Skip Connection을 통해 각각의 Decoder Block

연산 결과와 합쳐져 최종 결과물이 생성된다.

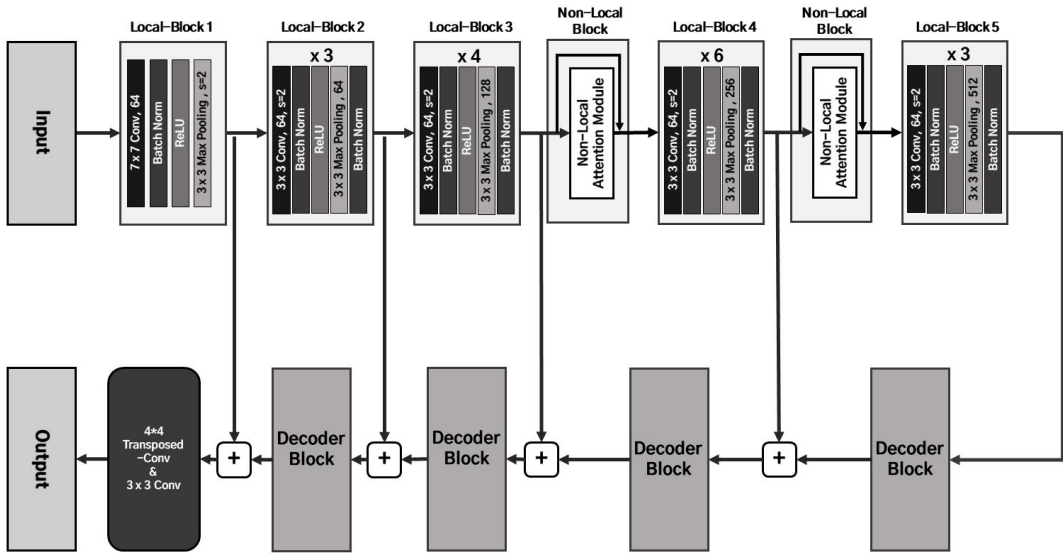


그림 3.6 NL-LinkNet Architecture (Wang et al., 2021)

4. Pretrained Encoder

앞서 설명한 세 가지 모델들의 경우 이미지에서의 특징을 뽑아내기 위하여 Encoder Block으로 사전 학습된 ResNet 모델을 사용한다. 딥러닝 분야에선 기울기 소실 또는 증폭과 같은 문제가 존재하며 이는 모델의 성능에 치명적인 영향을 끼친다. 이러한 문제를 해결하기 위하여 ResNet은 딥러닝 네트워크의 깊이가 깊어짐에 따라 발생하는 기울기 소실 문제를 완화하기 위해 제시된 모델이다(He et al., 2016). Skip connection 혹은 shortcut connection을 사용하지 않은 일반적인 CNN 신경망(Krizhevsky et al., 2012, Simonyan and Zisserman, 2014)들은 모델 네트워크가 점점 깊어질수록 기울기 소실 문제가 일어났는데 이는 모델 파라미터의 가중치를 업데이트하기 위한 오차역전파 과정 중 미분 횟수가 많아져 Input Layer에 가까워질수록 미분 값이 작아져 모델 전체에 전달되지 않기 때문이다. 이러한 문제를 해결하기 위해 ResNet은 신경망에 Shortcut 구조를 추가함으로써 보다 깊은 모델의 학습을 가능하게 하였다. 그림 3.7(a)와 같이 기존의 신경망은 입력값 x 를 타겟값 y 로 매핑하는 함수 $H(x)$ 를 얻는 것이 목적이었다면 ResNet은 그림 3.7(b)처럼 네트워크의 출력값이 x 가 되도록 $F(x)+x$ 를 최소화하는 것을 목적으로 한다. 입

력값 x 는 상수에 해당하고 $F(x)$ 를 0에 수렴되도록 학습이 이루어지며 만약 $F(x)$ 가 0에 수렴하게 되면 출력과 입력 모두 x 로 같아지게 된다. 이러한 $F(x)$ 를 잔차(residual)라 정의하며 이 잔차를 최소화하는 것이 ResNet의 목적이다.

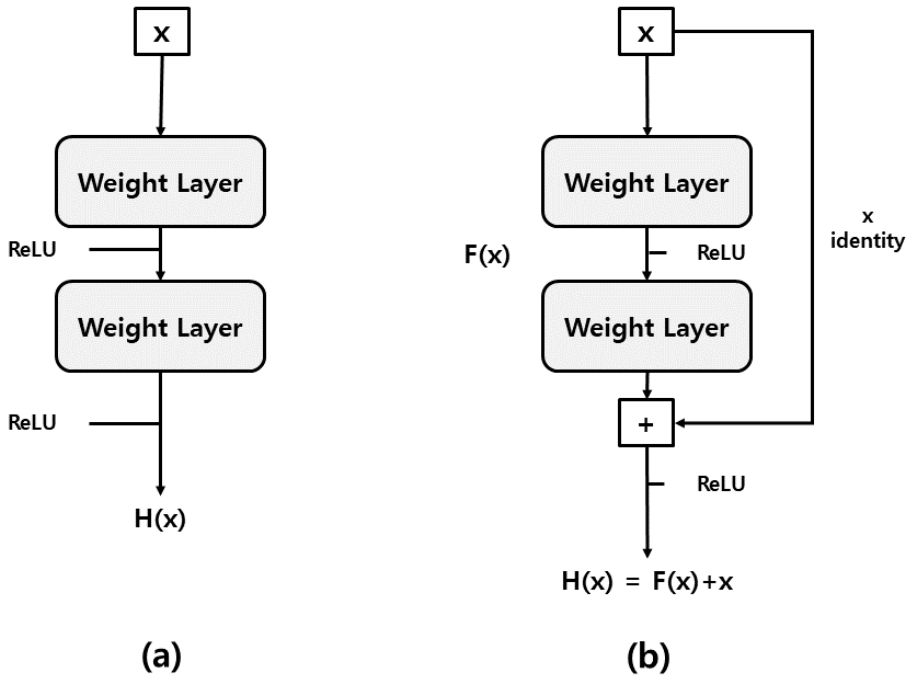


그림 3.7 신경망 구조 비교 (a) 기존 CNN의 구조, (b) ResNet의 구조
 (He et al., 2016)

본 연구에선 앞서 설명한 세 가지 모델의 Encoder로 잔차 학습으로 구성된 ResNet 중 ImageNet Dataset(Deng et al., 2009)을 통해 사전 학습된 ResNet-18과 ResNet-34를 각 모델에 적용하였다. ResNet-18과 ResNet-34는 표 3.1와 같이 5 단계의 구조를 가진다. ResNet-18과 ResNet-34의 차이점으로 conv2의 Basic Block을 2개에서 3개로, conv3의 basic block을 2개에서 4개로, conv4의 Basic Block을 2개에서 6개로 늘렸다는 점에 있다. ResNet은 Input 이미지를 7×7 과 3×3 크기의 Convolution filter를 통해 합성곱과 Max Pooling 연산을 한다. ResNet에서의 stride는 2로 설정되기에 이미지의 크기가 절반으로 줄어들게 되며, 단계마다 필터의 수가 늘어나 이미지의 채널 수는 두 배가 된다. 본 연구에선 고해상도 정사영

상에서의 도로를 추출하기 위해 사전 학습된 ResNet의 레이어 중 conv1, conv2, conv3, conv4, conv5을 불러와 Semantic Segmentation 알고리즘에 활용된다.

표 3.1 ResNet 구조 (He et al., 2016)

Layer name	Output	ResNet-18 Layer	ResNet-34 Layer
conv1	112×112	7×7, 64, stride 2	7×7, 64, stride 2
conv2	56×56	3×3, max pooling, stride 2	3×3, max pooling, stride 2
		$\begin{pmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{pmatrix} \times 2$	$\begin{pmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{pmatrix} \times 3$
conv3	28×28	$\begin{pmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{pmatrix} \times 2$	$\begin{pmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{pmatrix} \times 4$
conv4	14×14	$\begin{pmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{pmatrix} \times 2$	$\begin{pmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{pmatrix} \times 6$
conv5	7×7	$\begin{pmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{pmatrix} \times 2$	$\begin{pmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{pmatrix} \times 3$
	1×1	Average Pooling	
Flatten			
Fully-Connected			

5. IoU(Intersection over Unit)

Semantic Segmentation 분야에서는 모델의 성능을 평가하기 위하여 IoU 라는 지표를 사용하여 예측 결과와 실체가 얼마나 정확하게 겹치는지 수치화하게 된다. 이는 Semantic Segmentation 모델을 통해 예측된 bounding box와 실제 bounding box의 교집합 면적을 전체 bounding box 합집합 면적으로 나눈 값으로 그림 3.8에서 보여주고 있다. 도출된 IoU 결괏값은 0과 1 사이의 수치를 가지고 되고 해당 값이 1에 가까울수록 모델의 성능이 좋음을 의미한다.

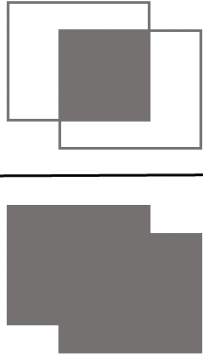
$$IoU = \frac{\text{Area of Intersection}}{\text{Area of Union}} = \frac{\text{Intersection Area}}{\text{Union Area}}$$


그림 3.8 IoU

제2절 변화 탐지

1. 변화 탐지 프로세스

본 연구에선 도로 추출 모델의 추론 결과를 변화 탐지에 활용하기 위하여 원격 탐사의 변화 탐지 분야에서 가장 대표적인 방법인 차분 기법을 사용하였다. 차분 기법이란 서로 다른 두 시기의 영상의 화소 값 차이를 이용하여 변화 분석 경향을 정량화하는 방식에 해당한다(최철웅 등, 2020). 변화 탐지의 프로세스는 그림 3.9와 같다. 먼저 이전 시점에 해당하는 정사영상에서의 도로 Ground Truth 정보와 이후 시점의 정사영상에 대한 도로 추출 모델의 추론 결과를 차분하게 되다. 시계열의 정사영상에서의 도로는 날씨, 계절, 등록오차 등 이미지에 포함된 정보의 성질이 서로 상이하므로 동일한 도로임에도 차분을 하게 되면 변화하지 않은 지역

들도 변화지역으로 나오는 Noise 정보가 차분 결과에 포함될 수 있다. 이러한 정보들을 제거하기 위하여 본 연구에선 차분 결과에 Window와 Threshold 파라미터를 추가하였다. Window는 차분 결과를 Window 크기에 따라 나누는 파라미터이고, Threshold는 Window 크기로 나누어진 차분 결과에 변화 정보가 얼마나 있는지 판단하는 파라미터에 해당한다. 예를 들어 그림 3.9와 같이 6×6의 차분 결과 이미지에 window를 3×3으로 적용하면 총 4개의 이미지로 분할된다. 이후 Threshold를 0.5로 지정하면 각 3×3의 이미지 안에 변화지역 Pixel의 개수가 5개 미만일 시 해당 이미지의 모든 Pixel값을 0으로 치환되며 분할된 이미지들을 합쳐 원래 크기로 복구한다. 이를 통해 차분 이미지에 존재하는 Noise를 제거하여 실질적으로 변화된 지역만을 남기게 된다.

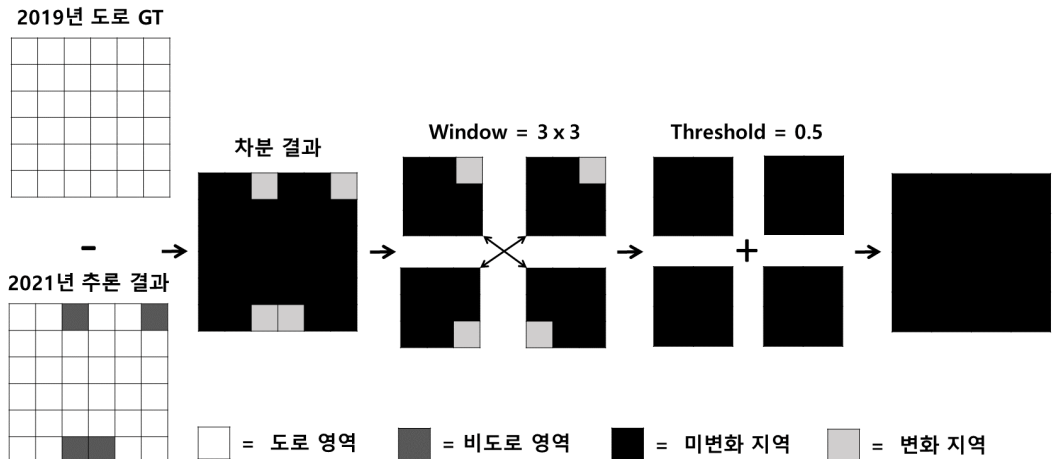


그림 3.9 변화 탐지 프로세스

2. Confusion Matrix

화소 기반의 변화 탐지의 성능은 대표적으로 Overall Accuracy, Recall, Precision, F1-score를 통해 평가된다. 언급한 네 가지의 평가 방법은 Confusion Matrix의 요소를 통해 계산할 수 있다. 본 연구에서는 변화 지역은 P(Positive), 미변화 지역은 N(Negative)으로 정의하였으며 이 두 가지를 조합하여 그림 3.10과 같이 Confusion Matrix를 구성할 수 있다. 먼저 TP는 True Positive의 약자로 실제

변화 지역을 변화지역으로 옳게 탐지했을 경우를 의미하고, FN은 False Negative의 약자로 실제 변화 지역을 미변화 지역으로 잘못 탐지했을 경우를 뜻한다. 반대로, TN은 True Negative의 약자로 실제 미변화 지역을 미변화 지역으로 옳게 탐지했을 경우를 의미하고, FP은 False Positive의 약자로 실제 미변화 지역을 변화 지역으로 잘못 탐지했을 경우를 말한다.

		Predicted value	
		Positive	Negative
Ground truth	Positive	TP (True Positive)	FN (False Negative)
	Negative	FP (False Positive)	TN (True Negative)

그림 3.10 Confusion Matrix

앞서 설명한 Confusion Matrix를 통해 여러 평가 방법을 사용할 수 있다. 먼저 Overall Accuracy는 변화탐지 분야에서 가장 대표적으로 사용되는 평가 지표로서 전체 추론 결과 중 변화 지역과 미변화지역을 얼마나 정확하게 분류되었는지를 나타내며 식 3과 같이 계산된다. Overall Accuracy의 경우 분자에 TN 값 즉, 미변화 지역을 미변화 지역으로 옳게 판단한 값이 들어간다. 미변화 지역은 변화지역 보다 비하여 월등히 많기 때문에 Overall Accuracy로 변화 탐지 결과를 평가할 경우 미변화 지역을 과하게 수용하게 되어 모든 경우의 Overall Accuracy가 높은 값으로 나타나는 문제점을 가지고 있다.

$$OA = \frac{TP + TN}{TP + FN + TN + FP} \tag{3}$$

Overall Accuracy의 단점인 미변화 지역을 과하게 수용하는 부분은 Recall을

통해 보완할 수 있다. 식 4와 같이 Recall은 실제 변화된 지역(FN+TP)이 변화 탐지 결과에서도 변화 지역이라 옳게 탐지(TP)한 비율을 나타낸 지표에 해당한다. Recall에는 TN이 고려되지 않기 때문에 미변화 지역에 대하여 Overall Accuracy보다 더 정확히 성능을 판단할 수 있다.

$$Recall = \frac{TP}{FN+TP} \quad (4)$$

하지만 Recall의 경우 변화 탐지의 결과가 대부분 지역을 변화된 지역(TP+FP)으로 나타날 경우, Recall에 고려되는 FN 값이 0이 되고 TP 값이 1이 되어 Recall의 1이 될 수 있다는 한계를 가진다. Precision을 통해 변화 탐지 결과를 함께 평가한다면 Recall의 단점을 보완할 수 있다. 식 5과 같이 Precision은 변화 탐지 결과에서 변화된 지역이라 탐지된 지역(FP+TP) 중 실제로 변화지역을 옳게 탐지(TP)한 비율을 나타낸 지표이다. 이를 통해 Recall이 높은 값으로 나타나더라도 변화된 지역이라 판단한 지역 중 실제 미변화된 지역이 포함되기 때문에 낮은 값의 Precision을 가질 수 있다.

$$Precision = \frac{TP}{FP+TP} \quad (5)$$

F1-score는 Precision과 Recall을 조합하여 하나의 통계치로 반환한 지표에 해당한다. F1-score에선 Precision과 Recall에 단순히 평균을 적용하지 않고 조화평균을 사용한다. 평균을 적용하게 된다면 둘 중 하나가 낮은 값을 가지더라도 나머지가 높은 값을 가지면 모델의 성능이 높게 나올 수 있다. 이를 보완하기 위해 F1-score는 식 6과 같이 조화평균에 Precision과 Recall을 대입하여 두 값 중 어느 한 가지 값이라도 극단적으로 낮을 경우 F1-score의 수치가 낮게 나오게 되며 두 값이 모두 커야만 F1-score이 높은 값으로 나타날 수 있다. 이를 통해 F1-score 지표는 Precision과 Recall 간의 상호보완적인 요소를 한 번에 고려할 수 있다.

$$F1-score = 2 \times \frac{Recall \times Precision}{Recall + Precision} \quad (6)$$

제 4장. 실험

제1절 데이터 처리 및 분석 아키텍처

연구를 진행하기 위해 한국지능정보사회진흥원 AIHub에서 제공하는 AIHub 위성영상 객체 판독 데이터와 주식회사 공간정보에서 제공한 광주광역시 드론 정사영상을 사용하였다. 광주광역시 드론 정사영상의 전처리는 QGIS 환경에서 이루어졌고 도로 윤곽 라벨링은 Label-Studio를 이용하여 작업되었으며 모델 학습 및 평가는 파이썬(python) 환경에서 이루어졌다. 이후 R 환경에서 도로 추출 모델의 추론 결과를 활용하기 위한 변화 탐지가 이루어졌으며, 연구의 전반적인 아키텍처는 그림 4.1과 같다.

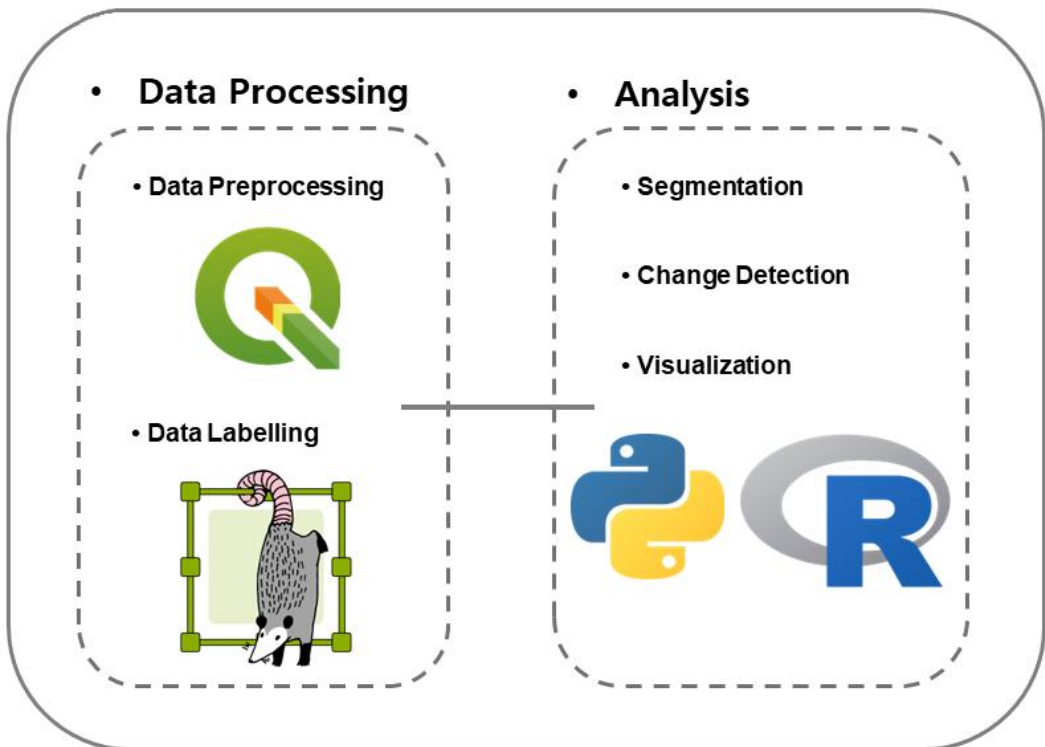


그림 4.1 시스템 아키텍처

1. AIHub 위성영상 객체 판독 데이터

AIHub 위성영상 객체 판독 데이터는 본 연구에서 사용된 데이터 중 하나로 아리랑 위성 영상을 활용한 5종(관심객체, 건물, 도로, 구름, 데이터영상 수계)으로 이루어져 있다. 이 중 본 연구에 사용된 데이터인 도로 윤곽 추출 데이터 총 1,271장의 위성영상 이미지로 구성된 데이터이다. 각각의 위성영상 이미지는 이미지 내에 존재하는 도로 윤곽 정보를 별도의 JSON 형태로 정리한 설명 파일과 1:1쌍을 이루고 있으며, 이를 annotation이라고 한다. 이는 txt, csv, xml 파일과 같은 형식으로 주로 사용되며 그림 4.2는 위성영상 예시 이미지, 해당 이미지의 annotation 파일에 포함된 도로 윤곽 좌표정보를 시각화하여 보여주고 있다.

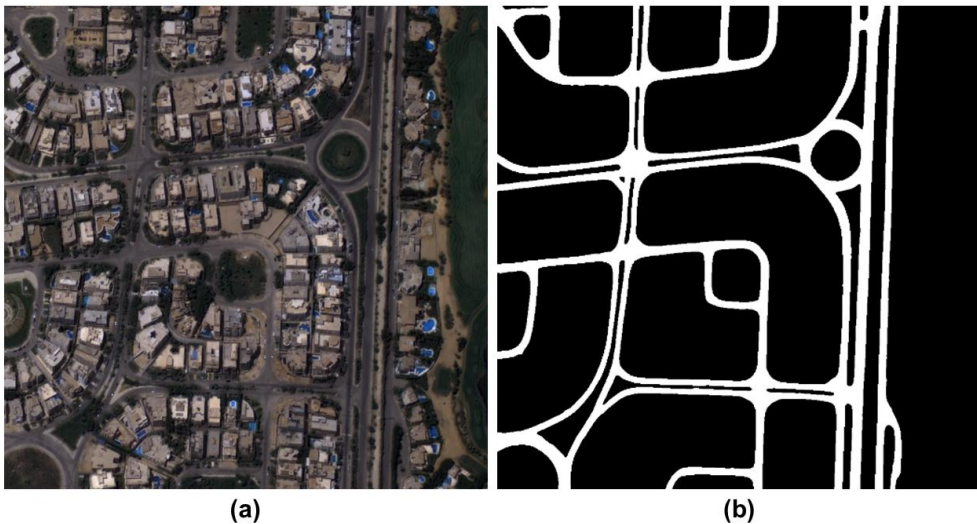


그림4.2 AIHub 데이터 (a) 위성영상 원본 이미지, (b) annotation 시각화

위성영상 이미지는 아리랑 3호와 아리랑 3A호에서 촬영된 위성영상을 기반으로 구축되었으며 각 위성의 특성은 표 4.1과 같다. 모든 이미지는 1024×1024 pixel의 크기를 가지고 있으며 아리랑 3호와 아리랑 3A호의 경우 각각 70cm, 55cm의 해상도를 가지기 때문에 아리랑 3호와 아리랑 3A호의 위성영상 이미지는 각각 716.8m×716.8m와 563.2m×563.2m의 공간 범위를 가지게 된다.

표 4.1 이리랑 위성 정보

특성	아리랑3호	아리랑3호
센서	광학 카메라	광학 카메라
궤도	태양동기	태양동기
고도	685km	528km
해상도	70cm	55cm
밴드 구성	PAN + 4MS	PAN + 4MS
촬영각	-30° ~ +30°	-30° ~ +30°
촬영시각	13:30경	13:30경
영상크기	약 24,000×24,000	약 24,000×24,000

본 연구의 목적인 도로의 변화를 탐지하기 위해선 동일 지역에 대한 여러 시기의 영상이 필요하다. 하지만 AIHub 위성영상 객체 판독 도로 윤곽 추출 데이터는 이미지에서의 도로 추출만을 목적으로 만들어졌기 때문에 동일 지역에 대한 여러 시점의 데이터가 존재하지 않는다. 이에 따라 AIHub 위성영상 객체 판독 도로 데이터의 경우 1271장 모두 도로 추출 모델의 학습에만 사용되었다.

2. 광주광역시 드론 정사영상 데이터

본 연구에서 사용된 드론 정사영상 데이터는 주식회사 공간정보로부터 제공받았으며 정사영상에서의 도로 추출과 변화 탐지에 사용되었다. 현재 광주광역시에선 장기적인 도시 여건 변화에 따른 교통수요 반영과 선진 대중교통 토대를 마련하기 위해 도시철도 2호선의 건설을 추진 중이다. 도시철도 2호선 공사는 1단계, 2단계, 3단계 구역으로 나뉘어 순차적으로 공사를 진행하게 되며 현재 1단계 구역(차량기지-시청-상무역-월드컵경기장-백운광장-조선대학교-광주역)이 공사 중에 있다. 해당 구역들은 도시철도 공사로 인해 도로의 상태가 빠르게 변화하기 때문에 본 연구의 목적인 도로 추출 및 변화 탐지에 적합하다고 판단하였고 이에 따라 1단계 구역의 공사 현장이 촬영된 서로 다른 시기의 드론 정사영상을 연구에 사용하였다.

그림 4.3은 2019년 드론 정사영상과 2021년 드론 정사영상의 전경을 보여주고 있다. 광주광역시 정사영상의 좌표계와 전체 크기, 용량 및 해상도에 대한 정보는 표 4.2와 같다.



그림4.3 드론 정사영상 전경 (a) 2019년 데이터, (b) 2021년 데이터

표 4.2 드론 정사영상 정보

특성	2019년 드론 정사영상	2021년 드론 정사영상
시기	2019년 12월	2021년 10월
좌표계	EPSG:5186-Korea 2000 / Central Belt 2010	
전체 크기	약 9 km × 6.3km	
용량	5.13 GB	6.54 GB
해상도	4.211cm	7.205 cm

가. 드론 정사영상 전처리

주식회사 공간정보에서 제공한 광주광역시 드론 정사영상은 원시데이터로 이루어져 있으므로 데이터를 Semantic Segmentation 모델이 인식할 수 있는 형태로

가공할 필요가 있다. 먼저 드론 정사영상에서의 공간 범위를 정의하여야 한다. 공간 범위가 서로 다르면 도로의 폭, 길이, 범위 등 여러 특성이 서로 상이할 수 있기 때문에 데이터의 일관성을 유지하기 위해선 드론 정사영상 또한 AIHub 위성영상 객체 판독 도로 데이터와 비슷한 공간 범위로 잘라낼 필요가 있다. 이를 위해 본 연구에선 GIS 소프트웨어 중 하나인 QGIS를 활용하여 일정 범위를 가진 그리드를 생성한 후 그리드 범위 안에 포함되는 지역을 추출하였다. 2019년 드론 정사영상의 경우 2021년 드론 정사영상에 비하여 더 높은 해상도를 가지기 때문에 상대적으로 더 작은 공간 범위를 가진다. 이에 따라 2019년 드론 정사영상의 공간 범위 보다 아리랑 위성영상 데이터의 공간 범위가 훨씬 크게 되어 그리드 내에 상당한 영상 정보가 포함되지 않는 문제점이 있다(그림 4.4).



그림 4.4 그리드 생성 예시

따라서 그리드 안에 드론 정사영상의 정보가 최대한 포함될 수 있도록 그리드의 범위를 2019년 드론 정사영상 해상도의 100배인 421.1m×421.1m로 설정하여 생

성하였고 이를 기준으로 공사 현장이 포함된 지역을 원시데이터에서 추출하였다. 그림 4.5는 2019년 드론 정사영상에서 그리드를 활용한 공사 지역 추출을 보여준다.

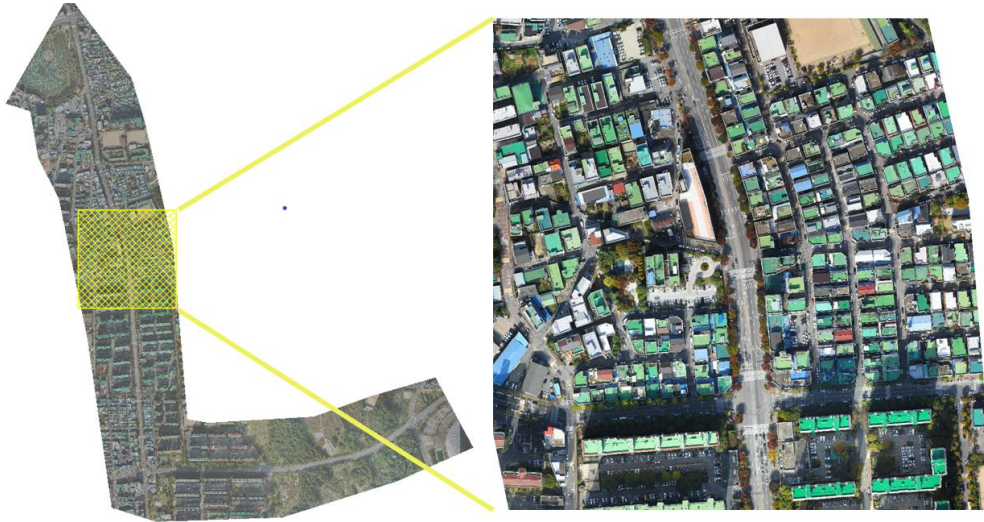


그림4.5 그리드를 활용한 도시철도 공사현장 추출 예시

잘라낸 드론 정사영상을 도로 추출에 활용하기 위해선 Semantic Segmentation 모델이 데이터를 인식할 수 있는 크기로 이미지 해상도를 낮춰주어야 한다. 본 연구에서 사용한 AIHub 위성영상 도로 데이터의 위성영상 이미지는 1024×1024 pixel 이기 때문에 잘라낸 드론 정사영상을 이와 동일한 해상도로 낮추었으며, 그 결과를 그림 4.6에서 보여준다.



(a)



(b)

그림 4.6 해상도 비교 (a) 드론 정사영상 원본 일부, (b) 해상도를 낮춘 드론 정사영상 일부

나. 데이터 라벨링

Semantic Segmentation 모델이 전처리 된 드론 정사영상을 통해 네트워크를 학습하고 학습된 모델을 평가하기 위해선 이미지 내에 존재하는 도로 윤곽에 대한 좌표정보가 필요하다. 이를 위해 이미지 라벨링 툴인 Label Studio를 활용하여 전처리 된 드론 정사영상에서의 도로 윤곽 좌표 정보를 생성하였으며, 그 예시를 그림 4.7에서 보여주고 있다. 그림 4.8에서는 전처리 된 드론 정사영상과 해당 영상을 기반으로 생성한 도로 윤곽 좌표정보를 시각화하여 보여준다.

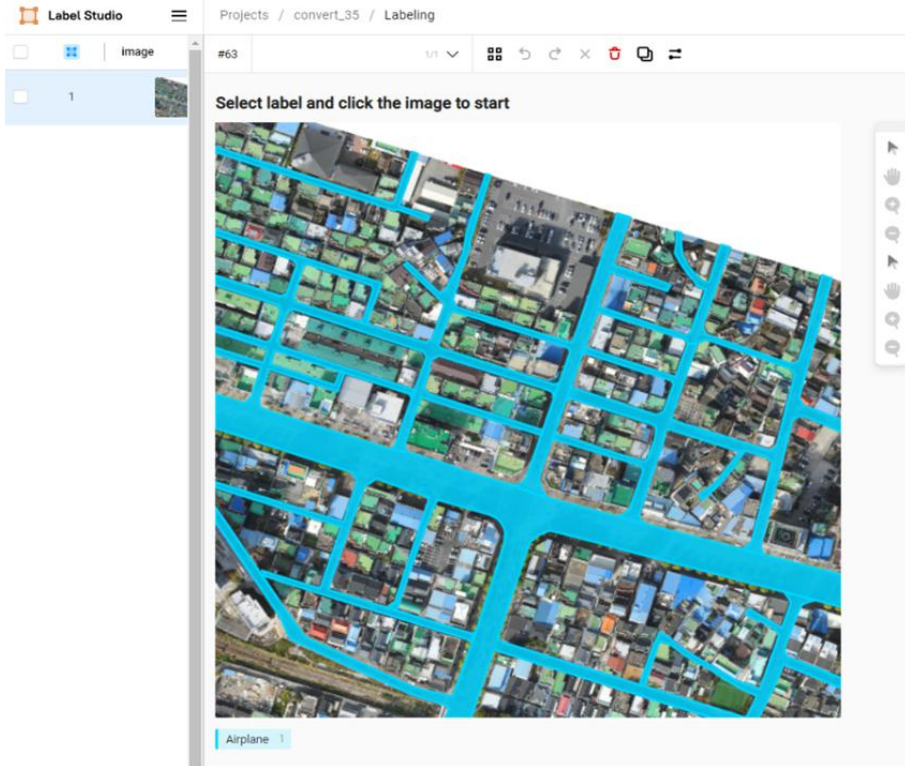
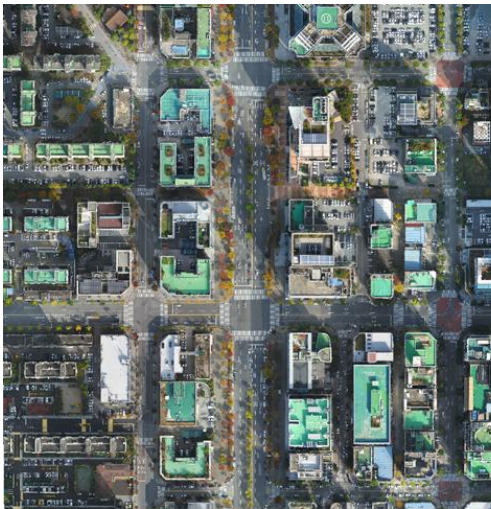
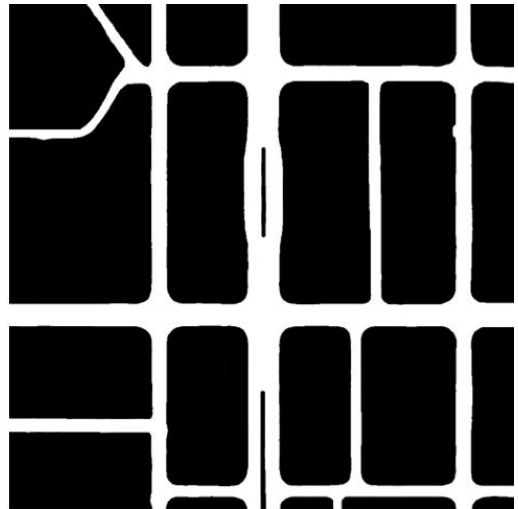


그림4.7 Label Studio 라벨링 예시



(a)



(b)

그림 4.8 드론 정사영상 데이터 생성 예시 (a) 전처리된 드론 정사영상, (b) 드론 정사영상에서의 도로 윤곽 좌표 정보 시각화

3. 최종 데이터 구성

최종적으로 생성한 광주광역시 드론 정사영상 데이터는 2019년 36장, 2021년 36장, 총 72장에 해당한다. Semantic Segmentation 모델의 학습 및 평가를 위하여 학습 데이터로 AIHub 위성영상 객체 판독 데이터 1271장과 광주광역시 드론 정사영상 데이터 42장을 사용하였으며 평가 데이터로 광주광역시 드론 정사영상 데이터 30장을 사용하였다. 광주광역시 드론 정사영상의 경우 동일 지역이 훈련과 평가에 사용되지 않게 하도록 그림 4.9와 같이 학습 데이터로 도시철도 1단계 구역에서 유촌동과 백운교차로 사이에 해당하는 지역의 정사영상을 사용하였으며 평가 데이터로는 그림 4.10와 같이 백운교차로부터 광주역 사이에 해당하는 지역의 정사영상을 사용하였다. 이는 데이터의 시간대가 달라도 동일한 지역이 학습과 평가에 사용된다면 도로의 형태가 크게 변화하지 않는 이상 기존 도로의 형태를 모델이 학습하게 되어 평가 결과에 영향을 줄 수 있기 때문이다. 따라서 드론 정사영상 데이터를 지역이 겹치지 않게 훈련 데이터 60%, 평가 데이터 40% 비율로 구분하여 연구를 진행하였다.

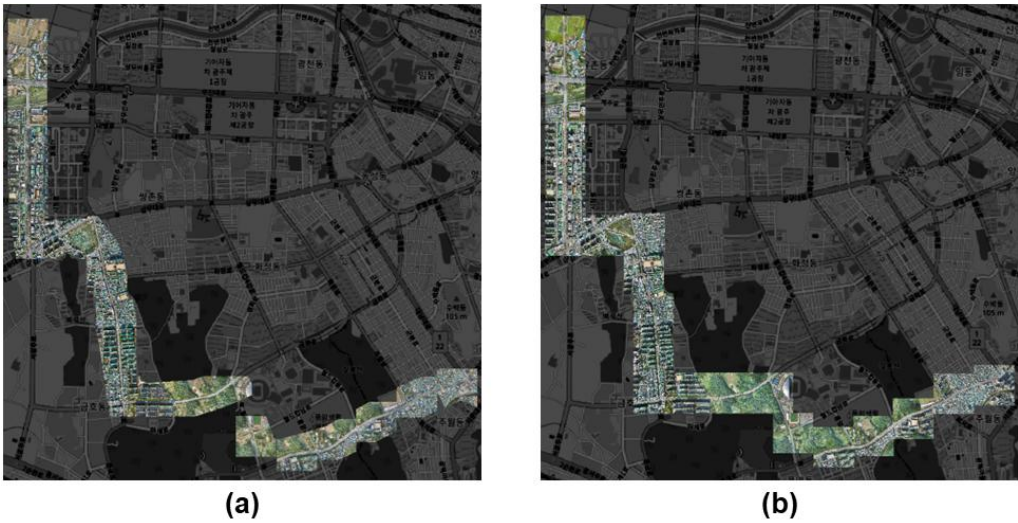


그림4.9 드론 정사영상 학습데이터 : 유촌동-백운교차로 구간 (a) 2019년 드론 정사영상 21장, (b) 2021년 드론 정사영상 21장

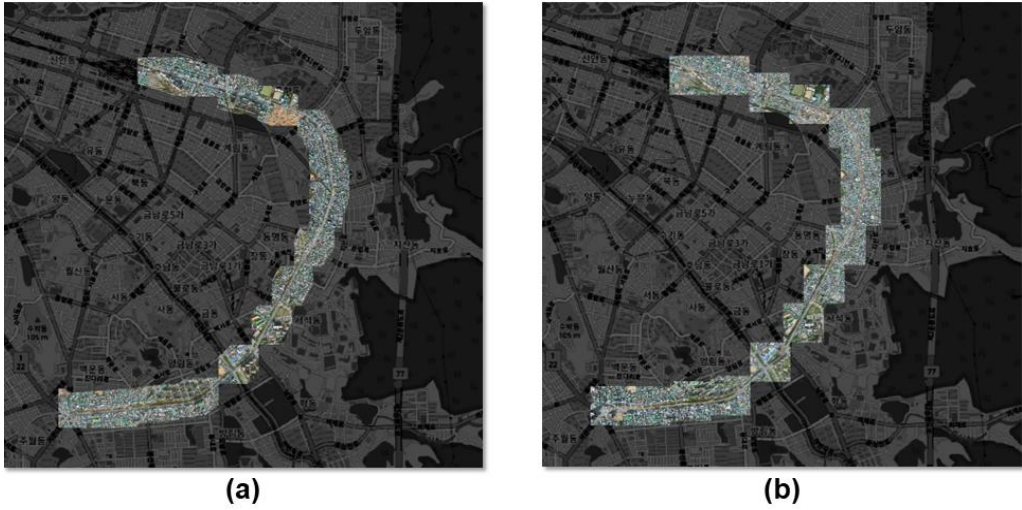


그림4.10 드론 정사영상 평데이터 : 백운교차로-광주역 구간 (a) 2019년 드론 정사영상 15장, (b) 2021년 드론 정사영상 15장

제2절 실험 환경

연구를 진행한 로컬 환경은 표 4.3과 같다. GPU 사용하기 위해 NVIDIA® CUDA® Toolkit version 11.1을 사용하였으며 Python 3.6.13, Pytorch 1.8.0 환경을 기반으로 Deep Learning을 구현하였다.

표 4.3 실험 환경

구성	사양
CPU	12 th Gen intel(R) Core(TM) i9-12900K
GPU	NVIDIA GeForce RTX 3090
RAM	64GB
OS	Window

제3절 실험 결과

1. 도로 추출 모델 평가 결과

광주광역시 드론 정사영상에 대한 도로 추출은 Semantic Segmentation 모델인 LinkNet, D-LinkNet, NL-LinkNet에 의해 수행되었다. 세 모델의 Pretrained Encoder로 ResNet-18과 ResNet-34를 각각 달리 적용하였으며 동일한 환경에서 모델의 성능을 비교하기 위하여 표 4.4와 같이 하이퍼파라미터를 설정하고 모델을 학습하였다.

표 4.4 하이퍼파라미터

Hyperparameter	Input
Image Size	768 x 768
Learning Rate	0.001
Epoch	500
Weight Decay factor	5.0
Batch Size	15

동일한 하이퍼파라미터를 적용한 각 모델의 성능은 표 4.5에서 확인할 수 있다. 가장 높은 성능을 보인 모델은 사전 학습된 ResNet-34를 Encoder로 사용한 LinkNet(이하 LinkNet-34)이 84.921%의 성능을 보였다. LinkNet과 D-LinkNet의 경우 ResNet-18을 사용하였을 때보다, 더 깊은 신경망 구조를 가진 ResNet-34를 사용하였을 때 더 높은 성능을 보였지만 이와 반대로 NL-LinkNet의 경우 ResNet-18을 Encoder로 사용하였을 때 더 높은 성능을 보였다. 그림 4.11은 각 모델의 도로 추출 추론 결과와 실제 도로 annotation의 시각화를 보여주며 이를 통해 LinkNet-34의 모델이 도로의 특징인 도로의 연결성과 복잡성, 얇고 긴 특성을 다른 모델보다 더 잘 파악하였음을 알 수 있다.

표 4.5 모델 성능 평가 결과

Model	Encoder	mIoU
LinkNet	ResNet-18	83.697%
	ResNet-34	84.921%
D-LinkNet	ResNet-18	83.662%
	ResNet-34	84.516%
NL-LinkNet	ResNet-18	84.499%
	ResNet-34	84.262%

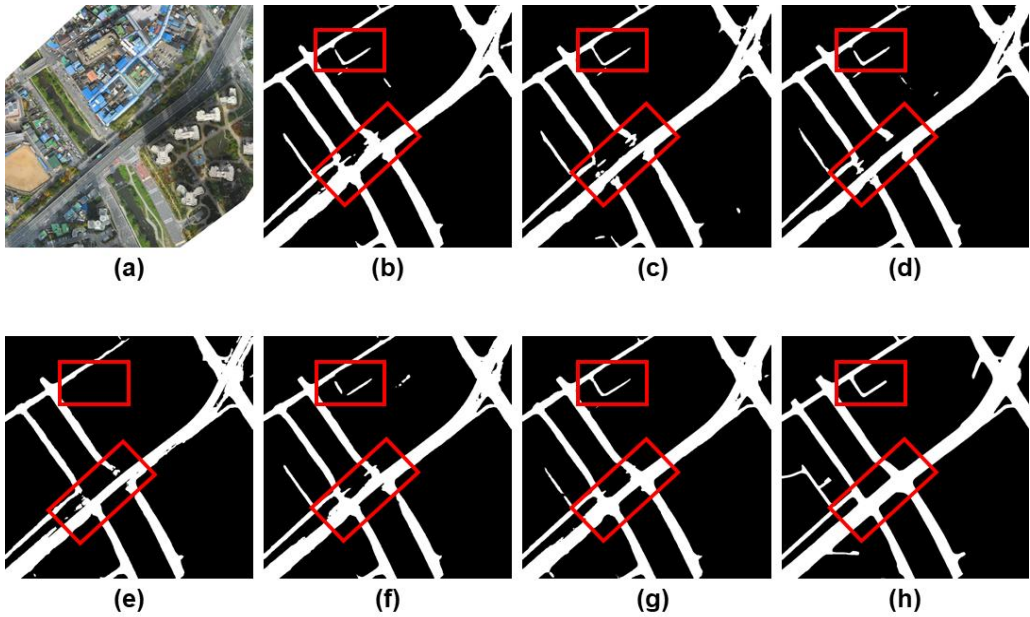


그림 4.11 각 모델 추론 결과 비교 (a) 드론 정사영상 원본, (b) NL-LinkNet(Resnet-18), (c) D-LinkNet(ResNet-18), (d) LinkNet(ResNet-18), (e) NL-LinkNet(ResNet-34), (f) D-LinkNet(ResNet-34), (g) LinkNet(ResNet-34), (h) Ground Truth

2. 도로 추출 모델 하이퍼파라미터 최적화

본 연구에선 가장 높은 성능을 나타낸 LinkNet-34의 성능을 더 높이기 위해 하이퍼파라미터 최적화를 적용하였다. Grid Search를 통해 찾은 모델에 가장 최적화된 하이퍼파라미터의 수치는 표 4.6과 같으며 하이퍼파라미터 최적화를 적용하지 않은 이전 모델 보다 0.204% 더 높은 85.125%의 성능을 나타냈다(표 4.7). 그림 4.12와 그림 4.13은 LinkNet-34 모델에 하이퍼파라미터 최적화를 했을 경우와 하지 않았을 경우의 도로 추출 결과를 보여준다.

표 4.6 LinkNet-34 하이퍼파라미터

Hyperparameter	Input
Image Size	1024 x 1024
Learning Rate	0.0005
Epoch	800
Weight Decay factor	5.0
Batch Size	24

표 4.7 하이퍼파라미터 최적화 적용 결과

Model	하이퍼파라미터 최적화 적용 여부	mIoU
LinkNet-34	X	84.921%
LinkNet-34	O	85.125%

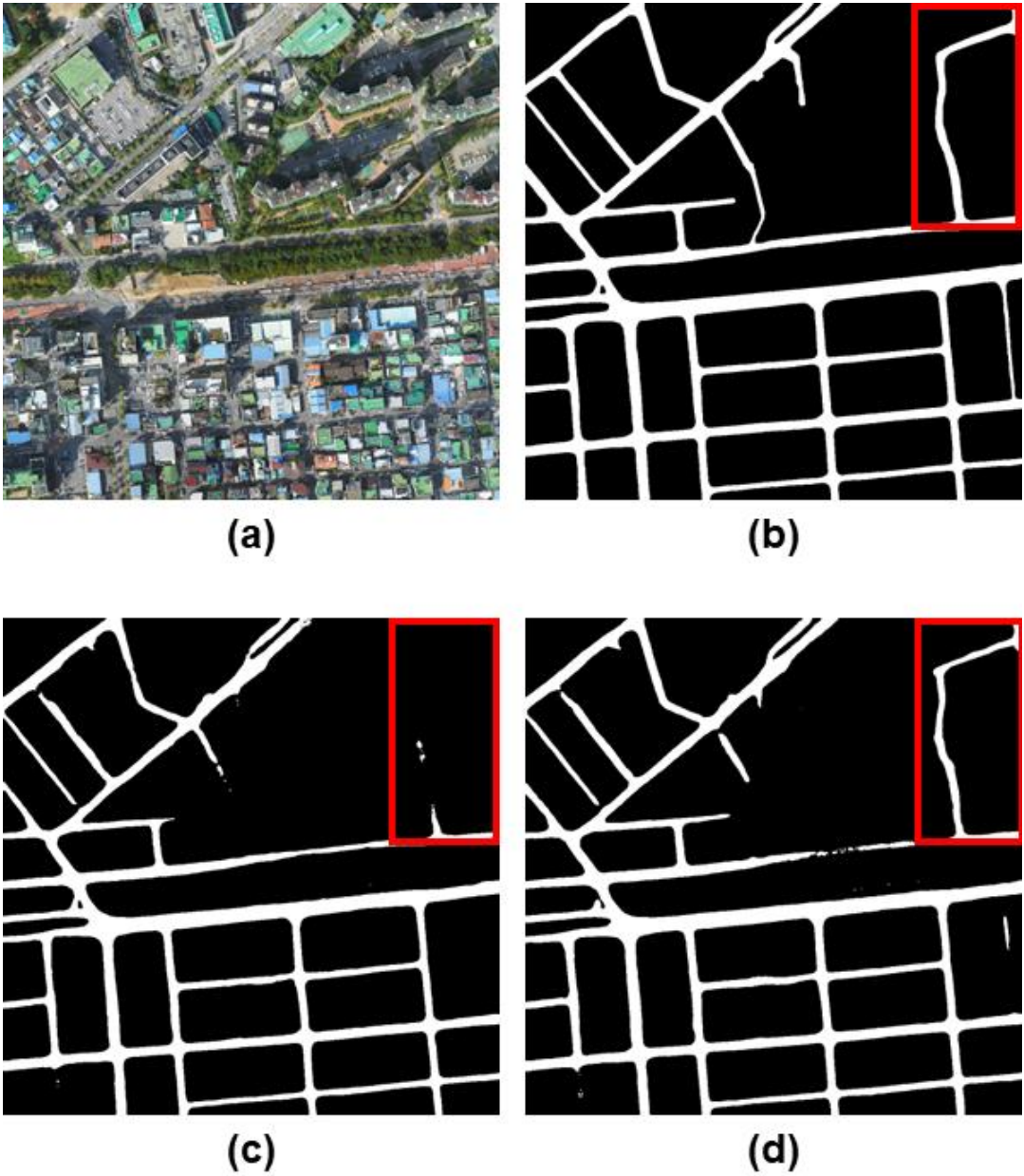
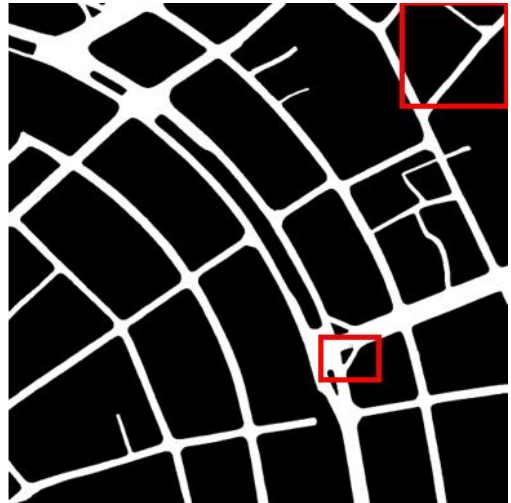


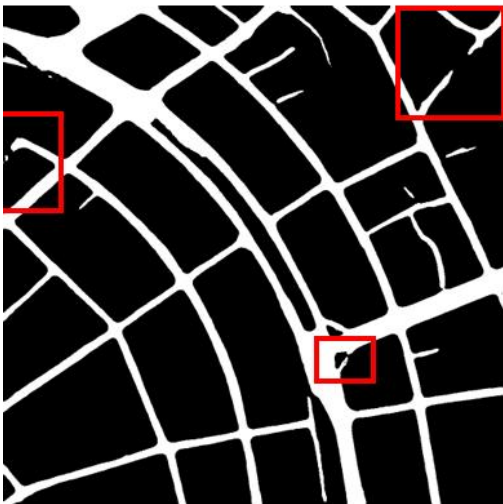
그림 4.12 추론 결과 비교1 (a) 드론 정사영상 원본, (b) 정사영상 Ground Truth, (c) 파라미터 최적화 전 LinkNet(ResNet-34), (d) 파라미터 최적화 후 LinkNet(ResNet-34)



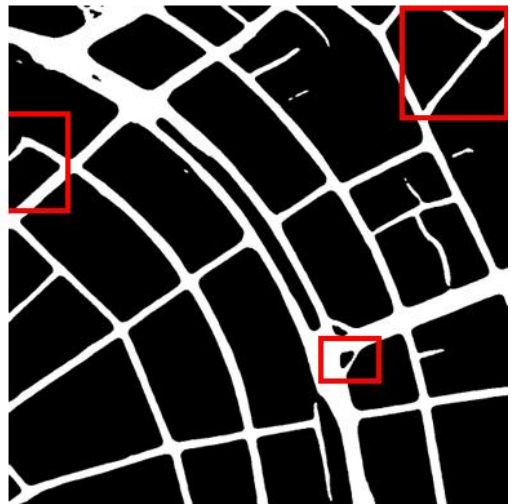
(a)



(b)



(c)



(d)

그림 4.13 추론 결과 비교2 (a) 드론 정사영상 원본, (b) 정사영상 Ground Truth, (c) 파라미터 최적화 전 LinkNet(ResNet-34), (d) 파라미터 최적화 후 LinkNet(ResNet-34)

3. 변화 탐지 결과

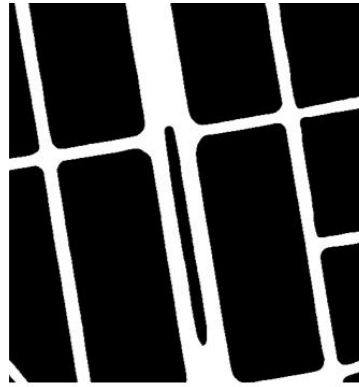
앞 절에서 Semantic Segmentation 모델들을 이용하여 드론 정사영상에서의 도로 윤곽을 추출한 후 가장 높은 성능을 보인 모델이 LinkNet-34인 것과 하이퍼파라미터 최적화를 통해 LinkNet-34의 성능을 높인 것을 확인하였다. 본 절에서는 하이퍼파라미터 최적화가 이루어진 LinkNet-34의 도로 추출 추론 결과를 활용하여 변화 탐지의 대표적인 방법인 차분 기법을 통해 변화 탐지 결과를 비교했다.

변화 탐지하기 위해선 서로 다른 두 시기의 정사영상에서의 도로 간의 화소 차 정보를 생성해야 한다. 변화 탐지를 위한 드론 정사영상 차분 정보는 2019년 드론 정사영상인 그림 4.14(a)의 도로 annotation 정보인 그림 4.14(b)와 2021년 드론 정사영상인 그림 4.14(c)에서의 도로 추출 추론 결과인 그림 4.14(d) 간의 화소 차를 구하여 그림 4.14(e)와 같이 생성하였다. 도로 추출에 사용된 드론 정사영상은 도로 추출 평가데이터인 2019년 15장, 2021년 15장을 사용하였으며 이에 따라 차분 결과는 총 15장의 이미지를 얻게 된다.

단순 차분만을 적용한다면 그림 4.14(e)와 같이 같은 도로를 검출하여도 Noise가 존재하게 되며 이를 제거하기 위해 본 연구에서는 Window와 Threshold 파라미터를 차분 결과에 적용하였다. Window 크기는 2×2, 4×4, 8×8, 16×16, 32×32, 64×64, 128×128을 적용하였고 Threshold는 2×2의 픽셀 개수를 고려하여 25%, 50%, 75%를 적용하였다. 표 4.8은 차분 결과에 Window와 Threshold를 적용한 평가 결과를 보여준다. Threshold의 경우 치환 기준을 %의 미만으로 설정하였으며 Window 크기 2×2에 Threshold 25%로 설정했을 경우 제거할 화소 개수가 총 4개 중 1개 미만이 되어 차분 원본 결과와 똑같아지므로 결괏값을 표 4.8에 넣지 않았다.



(a)



(b)



(c)



(d)



(e)

그림 4.14 차분 예시 (a) 2019년 드론 정사영상 일부, (b) 2019년 드론 정사영상 도로 Ground Truth 일부, (c) 2021년 드론 정사영상 일부, (d) 2021년 드론 정사영상 도로 추출 추론 결과 정보 일부, (e) 차분 결과 일부

표 4.8 변화 탐지 성능 평가 결과

Window	Threshold	Overall Accuracy	Precision	Recall	F1-score
차분 원본	-	0.97008	0.55397	0.84288	0.66855
2×2	50%	0.97213	0.57622	0.83715	0.68260
	75%	0.97697	0.64048	0.81333	0.71663
4×4	25%	0.97306	0.58689	0.83596	0.68962
	50%	0.97845	0.66249	0.81142	0.72943
	75%	0.98145	0.73015	0.76445	0.74691
8×8	25%	0.97833	0.65833	0.82076	0.73063
	50%	0.98203	0.74404	0.75929	0.75158
	75%	0.98215	0.80572	0.66065	0.72601
16×16	25%	0.98203	0.72891	0.79311	0.75966
	50%	0.98248	0.81341	0.66272	0.73037
	75%	0.97873	0.88895	0.46375	0.60953
32×32	25%	0.98336	0.79069	0.72798	0.75804
	50%	0.97944	0.90618	0.47503	0.62332
	75%	0.96947	0.93355	0.15842	0.27087
64×64	25%	0.98222	0.87777	0.58469	0.70187
	50%	0.96771	0.93840	0.10512	0.18907
	75%	0.96457	0.94942	0.01083	0.02142
128×128	25%	0.97014	0.89556	0.18773	0.31040
	50%	0.96420	NA	0.00000	NA
	75%	0.96420	NA	0.00000	NA

표 4.8에서 Overall Accuracy의 평가 지표가 가장 높은 경우는 Window 크기를 32×32 , Threshold를 25%로 설정하였을 때이며 0.98336의 성능을 나타냈다. 하지만 도로 변화 Ground Truth인 4.15(a)와 Window 크기를 32×32 , Threshold를 25%로 설정하였을 때의 결과인 4.16(c)를 비교해보면 Noise 정보들이 상당히 제거되었지만, 실질적으로 변화된 지역들의 정보가 4.16(b)에 비하여 많이 소실되었음을 확인할 수 있다. 3장에서 설명하였듯이 Overall Accuracy 평가 지표를 화소기반 변화 탐지에 적용하게 된다면 미변화 지역이 과하게 수용된다는 단점이 있다고 언급하였다. 표 4.8에서 Overall Accuracy 평가 결과를 보면 알 수 있듯이 차분 원본을 포함하여 모든 평가 결과가 95% 이상임을 알 수 있다. 따라서 Overall Accuracy 만으로 변화 탐지 성능을 제대로 판단할 수 없으므로 Precision과 Recall 지표로 평가한 결과를 비교해보아야 한다.

표 4.8에서 Recall의 평가 지표의 수치가 높은 경우 Window와 Threshold가 적용되지 않은 차분 원본에 해당하며 0.84288의 성능을 보였다. 하지만 같은 조건에서의 Precision의 경우 0.55397로 상당히 낮은 값을 가진다. 이러한 결과가 나온 이유로 그림 4.15(b)와 같이 실질적으로 변화된 지역이 Window와 Threshold 파라미터에 의해 제거되지 않고 보존되어 Recall 지표 값이 크지만, Noise 정보들이 제거되지 않았기 때문에 반대로 낮은 Precision 지표 값을 가지게 된다. 이와 반대로 Window 크기가 64×64 , Threshold가 75%인 상태에서 변화 탐지를 하게 되면 Precision 값이 0.94942로 가장 큰 값을 가지지만 대부분의 실제 변화된 지역들이 제거되어 극단적으로 낮은 0.01083의 Recall 지표 값을 가지게 된다.

Recall과 Precision의 한계를 보완할 수 있는 F1-score의 경우 Window 크기 16×16 , Threshold 25%로 설정하였을 때의 변화 탐지 결과가 0.75966로 가장 높게 나타났다. 해당 조건에서의 Recall과 Precision 지표 값의 경우 각각 0.79311, 0.72891로 다른 조건의 변화 탐지에 비하여 두 값이 상대적으로 차이가 크게 나지 않음을 알 수 있다. 또한 그림 4.16(b)과 같이 동일한 Threshold가 적용된 다른 변화 탐지 결과인 4.15(c), 4.15(d), 4.16(a)보다 Noise에 해당하는 정보들을 잘 제거함을 알 수 있다. 또한 Window 크기 32×32 , Threshold 25%인 변화 탐지 결과의 F1-score가 0.74944로 Window 크기 16×16 , Threshold 25%인 변화 탐지 결과의 F1-score와 0.00162의 차이가 나며 큰 차이를 보이지 않지만 Window 크기 32×32 의 적용 결과인 그림 4.16(c)의 경우 실제 변화 지역들이 일부 제거되면서 형상이 깨짐을 알 수 있다. 따라서 Window 크기가 16×16 , Threshold 25%일 때 Noise를 적절히 제거함과 동시에 변화 지역의 형상을 잘 유지함을 알 수 있다.

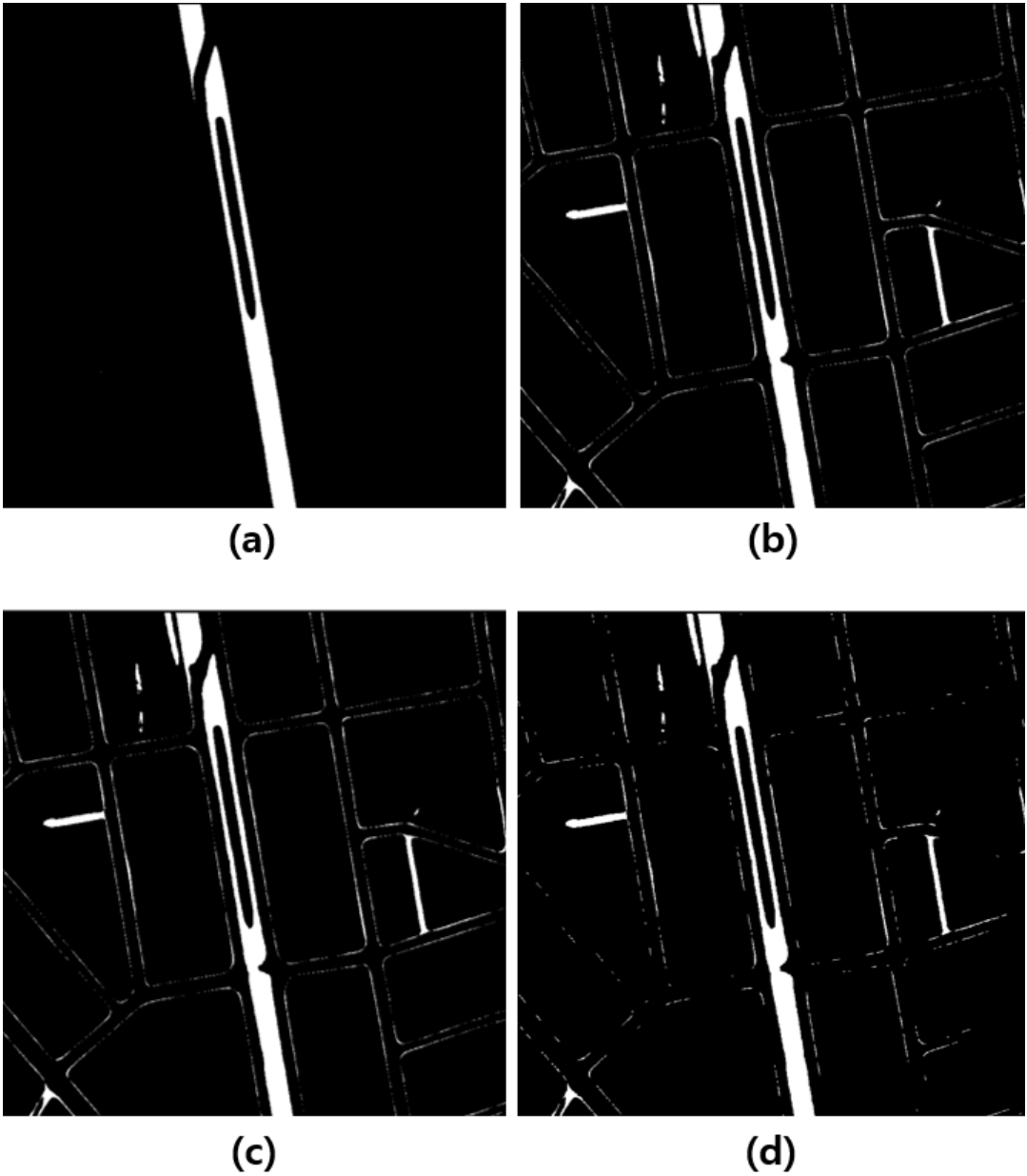


그림 4.15 Threshold 0.25에서의 Window 크기 별 변화 탐지 결과1 (a) Window Ground Truth. (b) 차분 원본, (c) Window 2x2, (d) Window 4x4

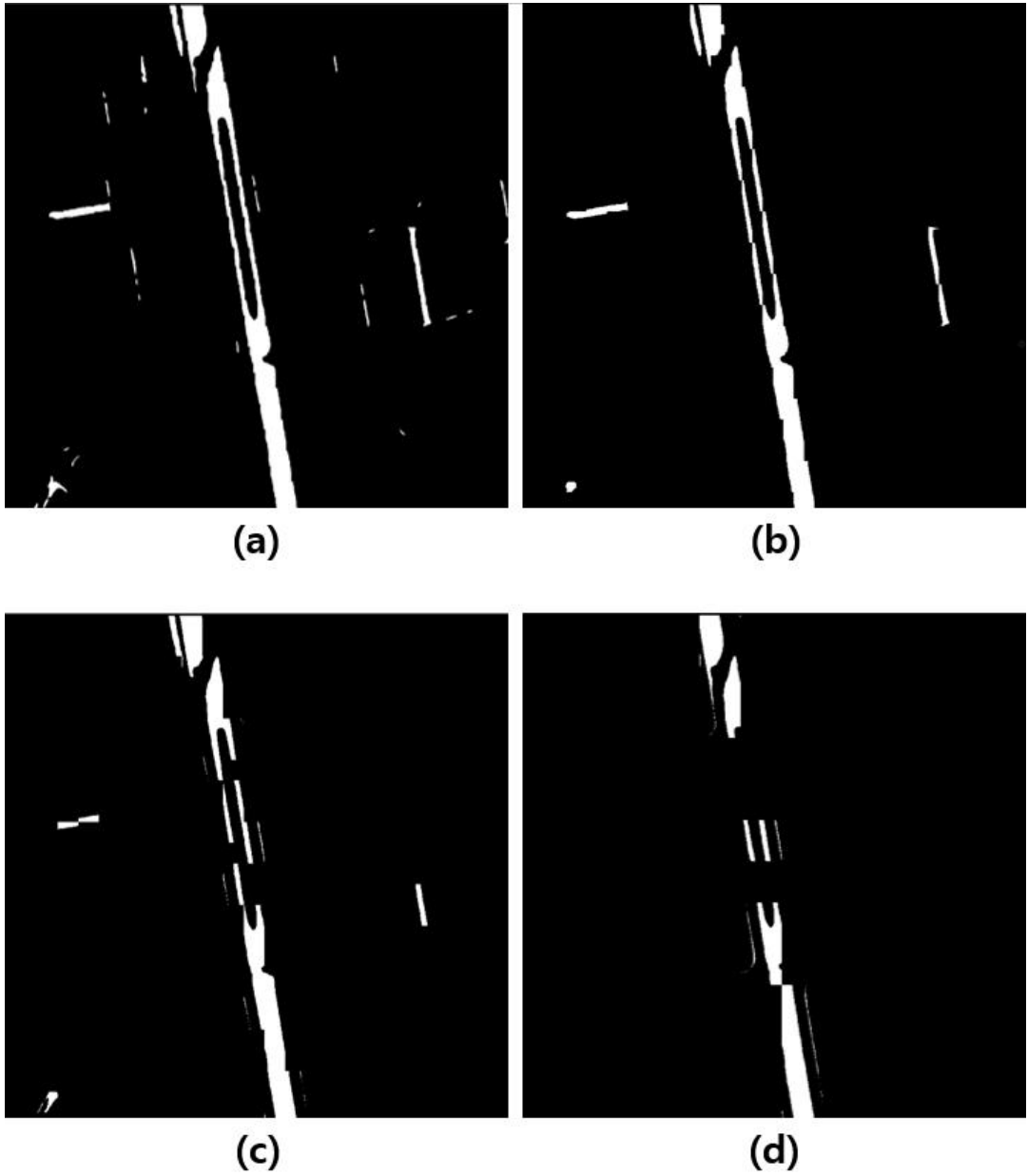


그림 4.16 Threshold 0.25에서의 Window 크기 별 변화 탐지 결과2 (a) Window 8×8, (b) Window 16×16, (c) Window 32×32, (d) Window 64×64

제 5장. 결론

본 연구의 목적은 기존 지도 구축을 효율적으로 갱신을 하고자 정사영상 이미지를 통해 빠르게 도로를 추출하고 변화 탐지를 수행하는 것이다. 이러한 연구 목적을 달성하기 위해 빠르게 데이터를 수집할 수 있는 드론 정사영상에 존재하는 도로 윤곽을 추출 할 수 있도록 딥러닝 기반 Semantic Segmentation 모델인 LinkNet, D-LinkNet, NL-LinkNet을 통해 도로 추출 모델을 구성했다. 이후 세 가지 모델 중 가장 높은 IoU 값을 가진 모델로 도로 추출 추론 결과를 뽑아내었고 이를 기반으로 도로가 기존과 달라진 점이 있는지 파악할 수 있게 하였다.

본 연구의 결과는 다음과 같다.

1. 드론 정사영상 이미지에서 도로 윤곽을 추출하기 위해 동일할 환경에서 LinkNet, D-LinkNet, NL-LinkNet 알고리즘에 Pretrained Encoder로 ResNet-18과 ResNet-34를 달리 적용해본 결과, ResNet-34를 Pretrained Encoder로 사용한 LinkNet의 추론 결과의 mIoU가 84.921%로 동일한 Encoder를 사용한 D-LinkNet 과 NL-LinkNet보다 0.435%, 0.659% 더 높은 성능을 보였다. 가장 높은 성능을 보인 ResNet-34를 사용한 LinkNet의 모델의 성능을 더 높이기 위해 하이퍼파라미터 최적화를 하였으며 Image Size를 1024 x 1024, Learning Rate를 0.0005, Epoch을 800, Weight Decay factor를 5.0 그리고 Batch Size를 24로 설정하였을 때 기존 LinkNet 모델 보다 0.204% 더 높은 85.125%의 성능을 보였다.

2. 도로 변화 탐지를 하기 위해 광주광역시 도시철도 2호선 공사 현장이 촬영된 서로 다른 시기의 드론 정사영상에서 ResNet-34를 Pretrained Encoder로 사용한 LinkNet 모델로 추론한 도로 추출 결과를 활용하였다. 2019년 드론 정사영상의 도로 annotation과 모델로 추출한 도로 정보의 화소 차를 구하여 차분 결과를 도출한 후 다양한 값의 Window 크기와 Threshold를 적용한 결과, Window 크기를 16×16, Threshold를 25%로 설정하였을 때 F1-score 평가 결과가 0.75966으로 가장 높았으며 변화 지역의 형상을 잘 유지하면서 차분 결과에 생기는 Noise 정보들을 제거함을 확인할 수 있었다.

본 연구는 기존 지도 제작 과정 중 수작업으로 진행되는 변화유무 파악 과정을 자동으로 탐지하여 지도를 효율적으로 생산하고자 빠르게 수집할 수 있는 드론 정사영상 이미지를 활용하여 빠르게 변화 탐지를 수행하는 것이다. 이를 위해 Deconvolution 구조의 Semantic Segmentation 알고리즘들을 드론 정사영상에 적용하였으며 각 모델 성능을 비교한 결과, ResNet-34를 Encoder로 사용한 LinkNet이 가장 높은 성능을 보임을 식별하였다. 이후 하이퍼파라미터 최적화가 이루어진 LinkNet-34 모델의 도로 추출 추론 결과에 차분 기법을 적용하였으며 Noise를 제거하기 위해 F1-score를 기준으로 Window와 Threshold 파라미터의 최적값을 탐색하였다. 이를 통해 실질적으로 도로의 변화된 부분만을 탐지한 결과를 보였지만 Segmantic Segmentation 모델의 도로 추출 결과가 100%의 정확도를 갖지 않기 때문에 도로를 제대로 추출하지 못하였을 경우 미변화 지역이 변화 지역으로 오탐지되어 변화 탐지 성능을 저하하는 한계점을 가진다. 또한 차분 결과의 Noise를 제거하기 위해 Window와 Threshold 파라미터를 적용하였지만 Window 크기에 따라 Noise가 제거되기에 실제 변화된 지역의 정보 또한 함께 소실되는 단점을 가진다.

향후 연구로 기존 지도 제작 과정에서 현장 측량 작업의 효율성을 위하여 탐지된 변화 지역 간의 클러스터링을 통해 작업 할당 구역을 지정하고 이를 실제 정사영상에 갱신하는 연구를 하고자 한다. 더 나아가 차분 기법뿐만이 아닌 새롭게 개발되고 있는 변화 탐지 알고리즘을 사용함으로써 도로의 변화 지역을 추출하여 성능을 비교하는 연구를 진행할 예정이며, hybrid 구조의 모델 혹은 Pseudo-Labeling 기법과 같은 Semi-Supervised Learning 기반의 Semantic Segmentation 기법을 적용하여 정사영상에 존재하는 객체를 추출하는 연구 또한 진행할 예정이다.

참고 문헌

1. 국토교통부, 2022, <http://www.nsd.go.kr/lxportal/?menuno=4065#3>
2. 박유경, 강원평, 최지은 and 김병주, 2019, 자율주행 지원을 위한 정밀도로지도 갱신기술 평가를 위한 기준 도출 연구, *한국지리정보학회지*, 22, 146-154.
3. 송창우, 정지훈, 홍성재, 김대회 and 강주형, 2020, SPADE 기반 U-Net 을 이용한 고해상도 위성영상에서의 도시 변화탐지, *대한원격탐사학회지*, 36, 1579-1590.
4. 오행열, 전승배, 김건, and 정명훈, 2022, 영상 기반 Semantic Segmentation 알고리즘을 이용한 도로 추출 및 변화 탐지, *한국측량학회*, under review.
5. 윤부열, 문두열 and 홍순현, 2006, Beacon GPS 를 이용한 수치지도 갱신에 관한 연구. *지구물리*, 9, 387-395.
6. 장영재, 오재홍 and 이창노, 2020, nDSM 및 도로망 추출 기법을 적용한 도심지 건물 변화탐지, *한국측량학회지*, 38, 237-246.
7. 주영은 and 이형석, 2005, 도로기반 수치지도의 수정 및 갱신을 위한 GPS-Van 적용에 관한 연구, *한국지리정보학회지*, 8, 129-141.
8. 최철웅, 이창현, 서용철 and 김지용, 2009, Landast 영상을 이용한 변화탐지 분석 기법 연구. *한국지리정보학회지*, 12, 13-20.
9. NGII, 2018, 정밀 도로지도 연계 효율화 연구 및 구축·갱신 연구 보고서.
10. NGII, 2019, 정밀 도로지도 연계 효율화 연구 및 구축·갱신 연구 보고서.
11. Abdollahi, A., Pradhan, B., Shukla, N., Chakraborty, S. and Alamri, A., 2020, Deep learning approaches applied to remote sensing datasets for road extraction: A state-of-the-art review, *Remote Sensing*, 12, 1444.
12. Alshehhi, R., Marpu, P. R., Woon, W. L. and Dalla mura, M., 2017, Simultaneous extraction of roads and buildings in remote sensing imagery with convolutional neural networks, *ISPRS Journal of Photogrammetry and Remote Sensing*, 130, 139-149.
13. Awad, M. M., 2013, A morphological model for extracting road networks from high-resolution satellite images, *Journal of Engineering*, 2013, 243021.
14. Bruzzone, L. and Bovolo, F., 2012, A novel framework for the design of

- change-detection systems for very-high-resolution remote sensing images. *Proceedings of the IEEE*, 101, 609-630.
15. Chaurasia, A. and Culurciello, E., 2017, Linknet: Exploiting encoder representations for efficient semantic segmentation. *2017 IEEE Visual Communications and Image Processing (VCIP)*, 1-4.
 16. Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K. and Yuille, A. L., 2017, Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs, *IEEE transactions on pattern analysis and machine intelligence*, 40, 834-848.
 17. Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F. and Adam, H., , 2018, Encoder-decoder with atrous separable convolution for semantic image segmentation, *Proceedings of the European conference on computer vision (ECCV)*, 801-818.
 18. Cleve, C., Kelly, M., Kearns, F. R. and Moritz, M., 2008, Classification of the wildland - urban interface: A comparison of pixel-and object-based classifications using high-resolution aerial photography, *Computers, Environment and Urban Systems*, 32, 317-326.
 19. Coppin, P., Jonckheere, I., Nackaerts, K., Muys, B. and Lambin, E., 2004, Review Article Digital change detection methods in ecosystem monitoring: a review, *International journal of remote sensing*, 25, 1565-1596.
 20. Demir, I., Koperski, K., Lindenbaum, D., Pang, G., Huang, J., Basu, S., Hughes, F., Tuia, D. and Raskar, R., 2018, Deepglobe 2018: A challenge to parse the earth through satellite images. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 172-181.
 21. Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K. and Fei-fei, L., 2019, Imagenet: A large-scale hierarchical image database, *2009 IEEE conference on computer vision and pattern recognition*, 248-255.
 22. Habib, T., Inglada, J., Mercier, G. and Chanussot, J., 2009, Support vector reduction in SVM algorithm for abrupt change detection in remote sensing, *IEEE Geoscience and Remote Sensing Letters*, 6, 606-610.

23. He, K., Zhang, X., Ren, S. and Sun, J., 2016, Deep residual learning for image recognition, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770-778.
24. Hussain, M., Chen, D., Cheng, A., Wei, H. and Stanley, D., 2013, Change detection from remotely sensed images: From pixel-based to object-based approaches, *ISPRS Journal of photogrammetry and remote sensing*, 80, 91-106.
25. Im, J., Jensen, J. and Tullis, J., 2008, Object based change detection using correlation image analysis and image segmentation, *International journal of remote sensing*, 29, 399-423.
26. Jin, S., Yang, L., Danielson, P., Homer, C., Fry, J. and Xian, G., 2013, A comprehensive change detection method for updating the National Land Cover Database to circa 2011, *Remote Sensing of Environment*, 132, 159-175.
27. Krizhevsky, A., Sutskever, I. and Hinton, G. E., 2017, Imagenet classification with deep convolutional neural networks, *Commun. ACM*, 60, 6, 84-90.
28. Liu, W. and Wang, H., 2008, An interactive image segmentation method based on graph theory, *J. Electron. Inf. Technol*, 8, 1973-1976.
29. Liu, Y., Pang, C., Zhan, Z., Zhang, X., and Yang, X., 2020, Building change detection for remote sensing images using a dual-task constrained deep siamese convolutional network model, *IEEE Geoscience and Remote Sensing Letters*, 18, 5, 811-815.
30. Long, J., Shelhamer, E. and Darrell, T., 2015, Fully convolutional networks for semantic segmentation, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 3431-3440.
31. Ma, R., Wang, W. and Liu, S., 2012, Extracting roads based on Retinex and improved Canny operator with shape criteria in vague and unevenly illuminated aerial images. *Journal of applied remote sensing*, 6, 1, 063610 1-14.
32. Mena, J. B. and Malpica, J. A., 2005, An automatic method for road extraction in rural and semi-urban areas starting from high resolution

- satellite imagery, *Pattern recognition letters*, 26, 1201-1220.
33. Miao, Z., Wang, B., Shi, W. and Zhang, H., 2014, A semi-automatic method for road centerline extraction from VHR images, *IEEE Geoscience and Remote Sensing Letters*, 11, 1856-1860.
 34. Mnih, V. and Hinton, G. E., 2010, Learning to detect roads in high-resolution aerial images. *European conference on computer vision*, 210-223.
 35. Moustafa, M. S., Mohamed, S. A., Ahmed, S. and Nasr, A. H., 2021, Hyperspectral change detection based on modification of UNet neural networks, *Journal of applied remote sensing*, 15, 2, 028505, 1-15.
 36. Panboonyuen, T., Vateekul, P., Jitkajornwanich, K. and Lawawirojwong, S., 2017, An enhanced deep convolutional encoder-decoder network for road segmentation on aerial imagery, *International conference on computing and information technology*, 191-201.
 37. Shi, A., Gao, G. and Shen, S., 2016, 2016, Change detection of bitemporal multispectral images based on FCM and DS theory, *EURASIP Journal on Advances in Signal Processing*, 1, 96-107.
 38. Shi, W., Miao, Z. and Debayle, J., 2013, An integrated method for urban main-road centerline extraction from optical remotely sensed imagery, *IEEE Transactions on Geoscience and Remote Sensing*, 52, 3359-3372.
 39. Simonyan, K. and Zisserman, A., 2014, Very deep convolutional networks for large-scale image recognition, *CoRR*, abs:1409.1556, 1-14
 40. Singh, A., 1989, Review article digital change detection techniques using remotely-sensed data, *International journal of remote sensing*, 10, 989-1003.
 41. Tao, W. B. and Jin, H., 2007, A novel method of image threshold segmentation based on graph theory, *Chinese Journal of Computer*, 1, 110-119.
 42. Unsalan, C. and Sirmacek, B., 2012, Road network detection using probabilistic and graph theoretical methods, *IEEE Transactions on Geoscience and Remote Sensing*, 50, 4441-4453.
 43. Varia, N., Dokania, A. & Senthilnath, J., 2018, DeepExt: A convolution

- neural network for road extraction using RGB images captured by UAV, *2018 IEEE Symposium Series on Computational Intelligence (SSCI)*, 1890-1895.
44. Vosselman, G. & Knecht, J. D., 1995, Road tracing by profile matching and Kaiman filtering, Automatic extraction of man-made objects from aerial and space images, Springer, 265-274.
 45. Wang, J., Qin, Q., Yang, X., Wang, J., Ye, X. and Qin, X., 2014, Automated road extraction from multi-resolution images using spectral information and texture, *2014 IEEE Geoscience and Remote Sensing Symposium*, 533-536.
 46. Wang, J., Song, J., Chen, M. and Yang, Z., 2015, Road network extraction: A neural-dynamic framework based on deep learning and a finite state machine. *International Journal of Remote Sensing*, 36, 3144-3169.
 47. Wang, X., Girshick, R., Gupta, A. and He, K., 2018, Non-local neural networks, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 7794-7803.
 48. Wang, Y., Seo, J. and Jeon, T., 2021, NL-LinkNet: Toward lighter but more accurate road extraction with nonlocal operations. *IEEE Geoscience and Remote Sensing Letters*, 19, 1-5.
 49. Wei, Y., Wang, Z. and Xu, M., 2017, Road structure refined CNN for road extraction in aerial image. *IEEE Geoscience and Remote Sensing Letters*, 14, 709-713.
 50. Wiratama, W., Lee, J. and Sim, D., 2020, Change detection on multi-spectral images based on feature-level U-Net. *IEEE access*, 8, 12279-12289.
 51. Yu, F. and Koltun, V., 2015, Multi-scale context aggregation by dilated convolutions, CoRR, abs/1511.07122, 1-13.
 52. Zhang, C., Murai, S. and Baltsavias, E. P. , 1999, Road network detection by mathematical morphology, *ISPRS Workshop 3D Geospatial Data Production: Meeting Application Requirements*, 185-200.
 53. Zhou, L., Zhang, C. and Wu, M., 2018, D-LinkNet: LinkNet with pretrained encoder and dilated convolution for high resolution satellite imagery road extraction, *Proceedings of the IEEE Conference on Computer Vision and*

Pattern Recognition Workshops, 182-186.

54. Zhu, C., Shi, W., Pesaresi, M., Liu, L., Chen, X. and King, B., 2005, The recognition of road network from high resolution satellite remotely sensed data using image morphological characteristics, *International Journal of Remote Sensing*, 26, 5493-5508.