

영상 기반 Semantic Segmentation 알고리즘을 이용한 도로 추출 Road Extraction from Images Using Semantic Segmentation Algorithm

오행열¹⁾ · 전승배²⁾ · 김 건³⁾ · 정명훈⁴⁾

Oh, Haeng Yeol · Jeon, Seung Bae · Kim, Geon · Jeong, Myeong-Hun

Abstract

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 road information is necessary for various purposes, such as High Definition Map for autonomous car driving. In the case of the Republic of Korea, accurate spatial information can be generated by making a map through the existing map production process. However, targeting a large area is limited due to 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. Therefore, it is essential to update road information accurately and quickly. This study uses Semantic Segmentation algorithms Such as 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. Subsequent studies should focus on comparing 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 can be applied to improve the speed of the existing map update process.

Keywords : Drone Image, Semantic Segmentation, Remote Sensing, Road Extraction

초 록

현대에는 급속한 산업화와 인구 증가로 인해 도시들이 더욱 복잡해지고 있다. 특히 도심은 택지개발, 재건축, 철거 등으로 인해 빠르게 변화하는 지역에 해당한다. 따라서 자율주행에 필요한 정밀도로지도와 같은 다양한 목적을 위해 빠른 정보 갱신이 필요하다. 우리나라의 경우 기존 지도 제작 과정을 통해 지도를 제작하면 정확한 공간정보를 생성할 수 있으나 대상 지역이 넓은 경우 시간과 비용이 많이 든다는 한계가 있다. 지도 요소 중 하나인 도로는 인류 문명을 위한 많은 다양한 자원을 제공하는 중추이자 필수적인 수단에 해당한다. 따라서 도로 정보를 정확하고 신속하게 갱신하는 것이 중요하다. 이 목표를 달성하기 위해 본 연구는 Semantic Segmentation 알고리즘인 LinkNet, D-LinkNet 및 NL-LinkNet을 사용하여 광주광역시 도시철도 2호선 공사 현장을 촬영한 드론 정사영상에서 도로를 추출한 다음 성능이 가장 높은 모델에 하이퍼 파라미터 최적화를 적용하였다. 그 결과, 사전 훈련된 ResNet-34를 Encoder로 사용한 LinkNet 모델이 85.125 mIoU를 달성했다. 향후 연구 방향으로 최신 Semantic Segmentation 알고리즘 또는 준지도 학습 기반 Semantic Segmentation 기법을 사용하는 연구의 결과와의 비교 분석이 수행될 것이다. 본 연구의 결과는 기존 지도 갱신 프로세스의 속도를 개선하는 데 도움을 줄 수 있을 것으로 예상된다.

핵심어 : 드론 정사영상, Semantic Segmentation, 원격 탐사, 도로 추출

Received 2022. 06. 10, Revised 2022. 06. 15 Accepted 2022. 06. 21

1) Dept. of Civil Engineering, M.S. student, Chosun University (E-mail: 5matrix0427@hanmail.net)

2) Dept. of Civil Engineering, Ph.D. student, Chosun University (E-mail: zeon6779@gmail.com)

3) Dept. of Civil Engineering, M.S. student, Chosun University (E-mail: rlarjs0151@naver.com)

4) Corresponding Author, Dept. of Civil Engineering, Associate Professor, Chosun University (E-mail: mhjeong@chosun.ac.kr)

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. 서론

현대에 들어 급속한 산업화와 인구의 증가로 인하여 도시가 점점 복잡해지고 있다(Yun *et al.*, 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 증가함에 따라 국토교통부는 2018년 자율주행 지원을 위한 신속 갱신 기술 개발 및 실증화 연구사업을 통해 관련 국내 기술을 개발 중이고, TomTom과 Deepmap 과 같은 지도 관련 기업들은 차량 내 지도 정보를 업데이트하는 기술들을 개발 중이다(Park *et al.*, 2019).

현재 우리나라에서의 지도는 항공 또는 위성사진 촬영을 통해 지표면의 정보를 수집하고 도화 작업에 필요한 좌표를 현장 측량 작업(지상기준점 측량, 사진기준점 측량)으로 수집한 이후 수치 도화를 통해 지표면에 존재하는 지형, 지물에 대한 정보를 생성하고 지리 지명 조사, 정위치 편집 작업 후 지리 정보의 전산화를 거쳐 최종적으로 국가 지리정보 Data Base 에 구축된다. 기존 지도 제작 과정을 통해 지도를 구축하면 정밀한 공간정보를 생성할 수 있으나 대상 지역이 넓을 경우, 많은 인력과 비용이 소요되는 단점이 있다(Kim *et al.*, 2013). 또한 소규모 시설 공사에 대한 정보는 등재되지 않기에 빠르게 변화하는 도로 정보를 수집함에 한계가 존재하며 이에 따라 지도 갱신에 어려움이 있다.

이러한 문제점을 해결하기 위하여 원격 탐사 분야에선 고 해상도 정사영상에서의 풍부한 정보를 이용하여 도로와 같은 지표면의 인공구조물 혹은 피복의 상태를 분석하는 연구가 이루어지고 있다. 특히 도로의 경우 인간 문명을 위한 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 업데이트와 같은 다양한 분야에 활용되기에 매우 중요하다(Shi *et al.*, 2013). 이에 따라 정사영상에서 도로의 여러 특징을 활용하여 도로를 추출하는 연구가 이루어졌다. 대표적으로 도로 추출을 위한 texture progressive analysis(Mena and Malpica, 2005)와 mathematical morphology(Zhu *et al.*, 2005), 도로의 분광적 특성을 이용한 방법(Panboonyuen *et al.*, 2017)과 같은 분석 방법으로 도로 구간에 대한 특정 정보를 사용하여 원격 탐사 이미지에서의 도로를 추출하였다. 하지만 이러한 다양한 유형의 도로를 처리함에 있어 한계를 가진다. 이와 같은 문제점

을 해결한 방법이 데이터 기반의 분석 방법인 딥러닝 기법이다. 본 연구에선 도로 정보의 신속한 갱신을 위하여 빠른 수집이 가능한 드론으로 정사영상을 수집한 후 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 적용하여 도심에서의 도로를 추출하였다.

2. 선행 연구

원격 탐사 이미지에서의 도로는 폭이 좁고 교차로를 가지는 기하학적 특성, 비도로 영역과의 색상이 다른 분광적 특성, 끊기지 않고 서로 연결되는 위상적 특성, 실제세계의 제약 조건에 의한 특정한 조건을 가지는 functional 특성(Vosselman and Knecht, 1995), 흑백영역에서의 공간적 특성을 가지는 texture 특성을 (Wang *et al.*, 2014) 가진다. 딥러닝 분야가 빠르게 발전하기 이전엔 양질의 데이터가 충분하지 않아 도로의 일부 특성을 다룬 비지도 학습 기반의 연구가 활발하게 진행되었다. Miao *et al.*(2014)은 고해상도의 위성영상에서의 도로 추출을 위해 mean shift 기법의 적용을 제안하였다. 해당 연구는 사용자가 미리 설정한 지점들을 연결하기 위해 측지선 방법을 적용하여 초기 도로 윤곽을 추출하고 임계값을 적용하여 도로를 비도로 지역과 분리하였다. 이후 측지선 방법을 다시 적용하여 연구자가 설정한 지점을 연결하여 커널 밀도 추정 맵을 생성하였다. 이러한 방법은 빠르고 효율적이지만 사용자가 수동으로 지점을 설정해야 한다는 단점을 가진다.

도로 추출을 위해 그래프 이론을 적용한 연구로 Tao and Jin(2007)은 배경과 도로를 구분하기 위하여 graph division method를 사용하였다. 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 기반하여 도로의 중심선과 형상 및 네트워크를 생성할 수 있는 모듈을 제안하였다. 그래프 이론을 기반한 도로 추출 방법은 복잡한 데이터 유형을 처리하고 도로의 전역적 특징을 제대로 추출할 수 있는 장점이 있지만 대부분 스펙트럼 행렬 개념을 채택하고 유사 행렬의 곱셈과 고유 벡터를 사용하기 때문에 고유 벡터의 수가 증가함에 따라 모델이 복잡해지는 단점이 있다.

Mathematical morphology는 도로의 기하학적 특성을 다

룬다. Zhang *et al.*(1999)은 이미지 전처리 단계에서 도로망을 분할 한 후 morphological trivial opening 연산을 적용하였으며 이를 통해 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 FCM (Fuzzy c-Means)과 SOM (Self Organizing Map)을 결합한 FCMs-SOM 알고리즘으로 이미지에서의 도로를 분류하고 분류 결과에서의 이질성을 줄이기 위해 morphological 방법인 dilation과 erosion을 적용하였다. 영상 분할 결과는 모양 및 크기와 같은 요소들의 선택에 따라 결과에 큰 영향을 받기 때문에 높은 정확도와 좋은 추출 결과를 얻기 위하여 mathematical morphology만을 사용하기 어렵다는 단점이 있다.

앞서 설명한 연구들의 경우 도로의 일부 특성만을 고려하기 때문에 폭, 길이, 형태 등 다양한 유형의 도로를 제대로 추출하지 못한다는 한계를 가진다. 특히 고해상도의 원격 탐사 이미지는 일반적인 이미지보다 포함된 정보가 풍부하기 때문에 기존 방법들은 이러한 복잡성을 제대로 반영하지 못하는 문제를 가진다(Abdollahi *et al.*, 2020). 따라서 기존의 접근 방식은 데이터 기반 접근법에 비하여 그 성능이 상대적으로 떨어지게 된다. 최근 원격 탐사 이미지에서의 도로 추출을 위해 컴퓨터 비전 분야에서 영상 분할을 목적으로 하는 딥러닝 기반의 semantic 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 CNN (Patch-based Convolutional Neural Network)(Mnih and Hinton, 2010)을 활용한 방법으로 이는 다양한 공간 특징을 추출하기 위하여 이미지를 여러 Patch로 나누어 합성곱 연산과 Pooling과 같은 연산을 통해 각 픽셀의 범주를 분류하는 구조를 가진다. Alshehhi *et al.*(2017)는 원격 탐사 이미지에서 도로와 건물을 동시에 추출하기 위하여 Patch-based CNN 구조를 적용하였으며 이를 통해 도로와 건물을 추출하였고 Wei *et al.*(2017)는 항공영상에서 도로를 추출하기 위해 VGG(Visual Geometry Group) 네트워크와 deconvolutional layer 및 Fusion Layer가 결합된 구조의 모델을 제안하였다. Patch-based CNN의 경우 하나의 이미지를 여러 Patch로 분할함으로써 인해 인접한 Pixel 간의 공간적 상관성을 잃어버리며 이에 따라 객체들의 경계가 선형적이지 않거나 모호해짐으로 인해 잘못 분류된 영역들을 줄이기 위한 별도의

의 후처리 과정이 필요하다는 한계를 가진다.

다음은 FCN (Fully Convolutional Network)(Long *et al.*, 2015)를 기반한 도로 추출 방법으로 기존 이미지 분류에서 높은 성능을 보인 CNN 기반 모델들로 이미지를 학습한 후 up-sampling을 통해 입력 이미지와 동일한 크기의 출력 이미지를 생성하여 객체를 추출하는 구조를 가진다. Varia *et al.*(2018) 초고해상도의 UAV (Unmanned Aerial Vehicle) 이미지에서의 FCN-32 모델을 통해 도로를 추출하였으나 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 공간정보의 손실을 메우기 위하여 보간법을 활용하기 때문에 모델의 정밀도가 낮아진다는 단점을 가진다.

Deconvolution 구조는 일반적으로 encoder와 decoder의 네트워크로 구성된 구조를 가지고 있으며 encoder에서 이미지 내 존재하는 객체의 특징을 뽑아내고 이를 decoder가 받아들여 입력 이미지와 동일한 크기의 출력 이미지를 생성한다. Deconvolution 구조는 Patch-based CNN과 FCN 모델들보다 높은 성능을 보이지만 학습해야 할 파라미터의 수가 많아 높은 컴퓨팅 프로세스가 필요하다는 단점이 있다(Abdollahi *et al.*, 2020). Deconvolution 구조의 단점인 모델 파라미터 수를 줄이면서 성능을 높이기 위해 개발된 딥러닝 모델이 LinkNet이다. LinkNet의 효율적인 구조를 활용하여 여러 semantic 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 추출에 사용하였다.

3. 방법론

3.1 Semantic Segmentation 모델

본 연구에서 사용한 LinkNet은 Fig. 1과 같이 왼쪽의 encoder block과 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 이미지

를 생성하게 된다. 객체의 공간정보를 유지하기 위해 encoder Block에서 생성되는 특징 맵이 decoder Block으로 전달되며 이를 통해 decoder block에서의 up-sampling으로 생기는 공간정보의 손실을 복구하게 된다.

본 연구에 LinkNet 모델과 함께 사용된 D-LinkNet은 deepglobe road extraction challenge(Demir *et al.*, 2018)에서 우수한 모델로 높은 해상도의 위성영상에서 도로 검출을 목

적으로 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 이미지 전체를 효율적으로 학습할 수 있다고 한다. 따라서 D-LinkNet은 이러한 도로의 특성을 잘 반영하기 위하여 사

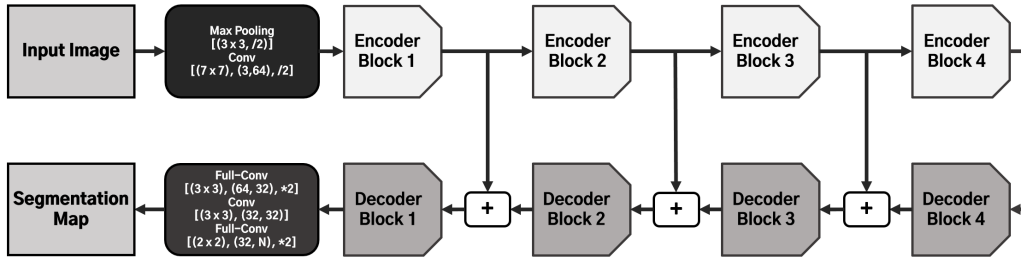


Fig. 1. LinkNet architecture (Chaurasia and Culurciello, 2017)

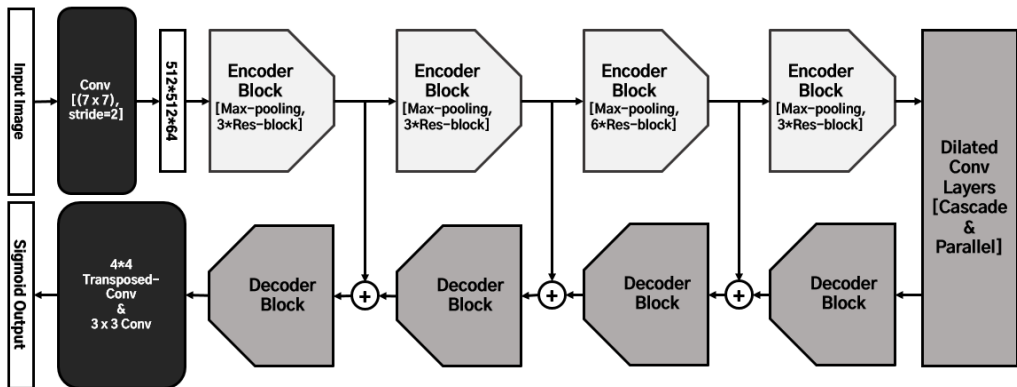


Fig. 2. D-LinkNet architecture (Zhou *et al.*, 2018)

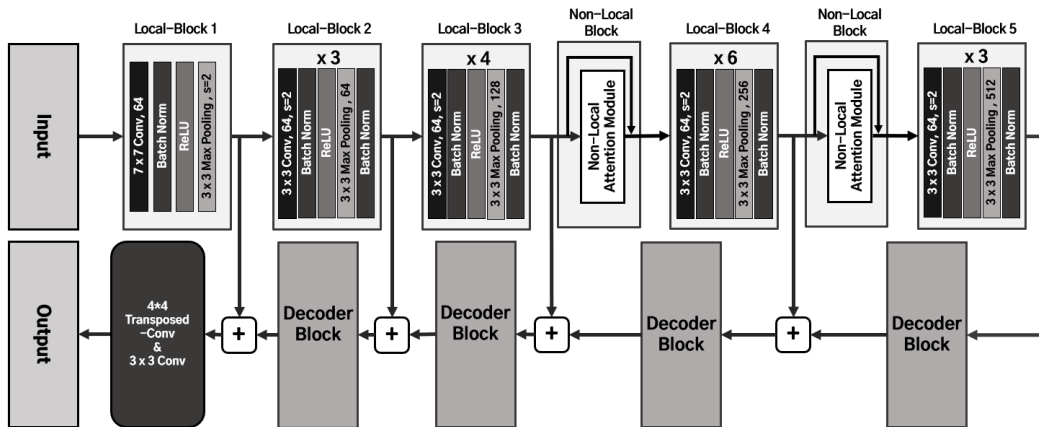


Fig. 3. NL-LinkNet architecture (Wang *et al.*, 2021)

전 학습된 encoder로 구성된 LinkNet에 dilated convolution layer 들이 중앙에 추가된 구조를 가진다(Fig. 2).

NL-LinkNet은 신경망의 매개변수를 줄이는 동시에 고해상도의 위성영상에서 도로를 효율적으로 추출하기 위해 기존 LinkNet의 구조에 non-local attention module을 삽입한 모델에 해당한다(Fig. 3). Computer vision 분야에서의 convolution은 kernel의 크기만큼 이미지를 학습하기 때문에 local operator에 해당한다. 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 방법은 layer의 수가 늘어나 연산량이 많아지기 때문에 모델을 학습시키기 어려워진다. 특히 고해상도 정사영상에선 도로 간의 간격이 넓고 긴 범위에 걸쳐 이미지에 존재하기에 local operator의 한정된 receptive field로 객체를 추출할시 올바르게 검출하지 못할 수도 있다. 따라서 NL-LinkNet은 도로의 특성을 잘 반영하기 위하여 non-local block(Wang et al., 2018)을 LinkNet 구조에 삽입한 구조를 가진다.

3.2 데이터

본 연구에서 사용한 데이터는 AIHub 위성영상 객체 판독 데이터와 광주광역시 드론 정사영상이다. 먼저 한국지능정보사회진흥원 AIHub에서 제공하는 AIHub 위성영상 객체 판독 데이터 중 도로 윤곽 추출 데이터는 총 1,271장의 위성영상 이미지와 각 이미지에 대칭되는 도로 annotation으로 구성된 데이터이다. 위성영상 이미지는 아리랑 3호와 아리랑 3A호에서 촬영된 위성영상을 기반으로 구축되었으며 모든 이미지는 1024×1024 pixel의 크기를 가진다. 아리랑 3호와 아리랑 3A호의 경우 각각 70cm, 55cm의 해상도를 가지므로 아리랑 3호와 아리랑 3A호의 위성영상 이미지는 각각 716.8m×716.8m와 563.2m×563.2m의 공간 범위를 가진다. AIHub 위성영상 객체 판독 데이터는 광주광역시 드론 정사영상과 다르게 모델이 인식할 수 있는 형태로 가공되어 제공되는 데이터이기 때문에 별도의 전처리 작업을 수행할 필요가 없다. 본 연구의 목적인 드론 정사영상에서의 도로 추출을 위하여 AIHub 위성영상 객체 판독 데이터는 모델의 학습에만 사용되었다.

현재 광주광역시는 장기적인 도시 여건 변화에 따른 교통 수요 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 구역이 공사 중에 있다. 해당 구역들은 도시철도 공사로 인해 도로의 상태가 빠르게 변화하기 때문에 본 연

구의 목적인 도로 추출에 적합하다 판단하였고 이에 따라 (주)공간정보에서 제공한 1단계 구역의 공사 현상이 촬영된 드론 정사영상을 모델의 학습 및 평가에 사용하였다. 광주광역시 정사영상의 좌표계와 전체 크기, 용량 및 해상도에 대한 정보는 Table 1과 같다.

Table 1. Drone image details

Component	2019 Drone image	2021 Drone image
Period	2019. 12.	2021. 10.
Coordinate system	EPSG:5186-Korea 2000 / Central Belt 2010	
Size	9 km × 6.3km	
Volume	5.13 GB	6.54 GB
Resolution	4.211cm	7.205 cm

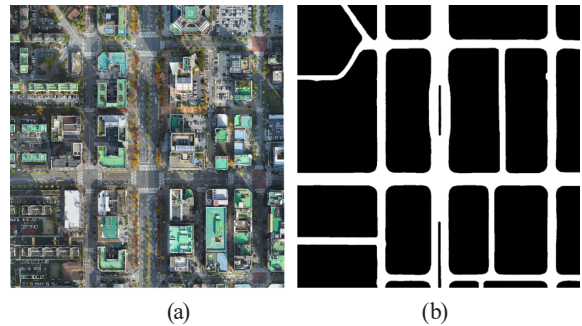


Fig. 4. Example of drone image data generation
(a) visualization of pre-processed drone image, (b) visualization of road coordinate information in image

원시데이터의 광주광역시 드론 정사영상을 모델이 인식할 수 있는 형태로 가공하기 위해 공간 범위를 정의하였다. 공간 범위가 서로 다르다면 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 QGIS 소프트웨어를 활용하여 드론 정사영상의 정보가 최대한 포함될 수 있도록 그리드의 범위를 2019년 드론 정사영상 해상도의 100배인 421.1m×421.1m로 설정하여 그리드를 생성한 후 그리드 범위 안에 포함되는 지역을 선택하였다. 그리드로 선택된 드론 정사영상을 도로 추출에 활용하기 위해선 semantic segmentation 모델이 데이터를 인식할 수 있는 해상

도를 낮춰주어야 한다. 본 연구에서 사용한 AIHub 위성영상 도로 데이터의 위성영상 이미지는 1024×1024 pixel이기 때문에 잘라낸 드론 정사영상을 이와 동일한 해상도로 낮추었다. semantic segmentation 모델 학습에 필요한 도로 좌표 정보는 label studio를 통해 생성하였다(Fig. 4).

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 학습 데이터로 도시철도 1단계 구역에서 유촌동과 백운교차로 사이에 해당하는 지역의 정사영상을 사용하였으며 평가 데이터로는 백운교차로부터 광주역 사이에 해당하는 지역의 정사영상을 사용하였다. 이는 데이터의 시간대가 달라도 동일한 지역이 학습에 사용된다면 평가 결과에 영향을 줄 수 있기 때문이다. 따라서 지역이 겹치지 않게 드론 정사영상 데이터를 훈련 데이터 60%, 평가 데이터 40% 비율로 구분하였다.

모델 학습 및 평가는 Window OS 기반의 12th Gen intel(R) Core(TM) i9-12900K, NVIDIA GeForce RTX 3090, RAM

64GB 환경에서 이루어졌다. GPU 사용하기 위해 NVIDIA® CUDA® Toolkit version 11.1을 사용하였으며 Python 3.6.13, Pytorch 1.8.0 환경을 기반으로 deep learning을 구현하였으며 모델 평가 기준은 일반적으로 사용되어지는 mIoU (mean Intersection over Union)를 사용하였다.

4. 결과

동일한 조건에서 모델을 학습하기 위해 이미지 크기는 768x768, learning rate 0.001, epoch 600, weight decay factor 5.0, batch size 15로 설정하였으며 모델의 성능은 Table 2에서 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. Fig. 5는 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.

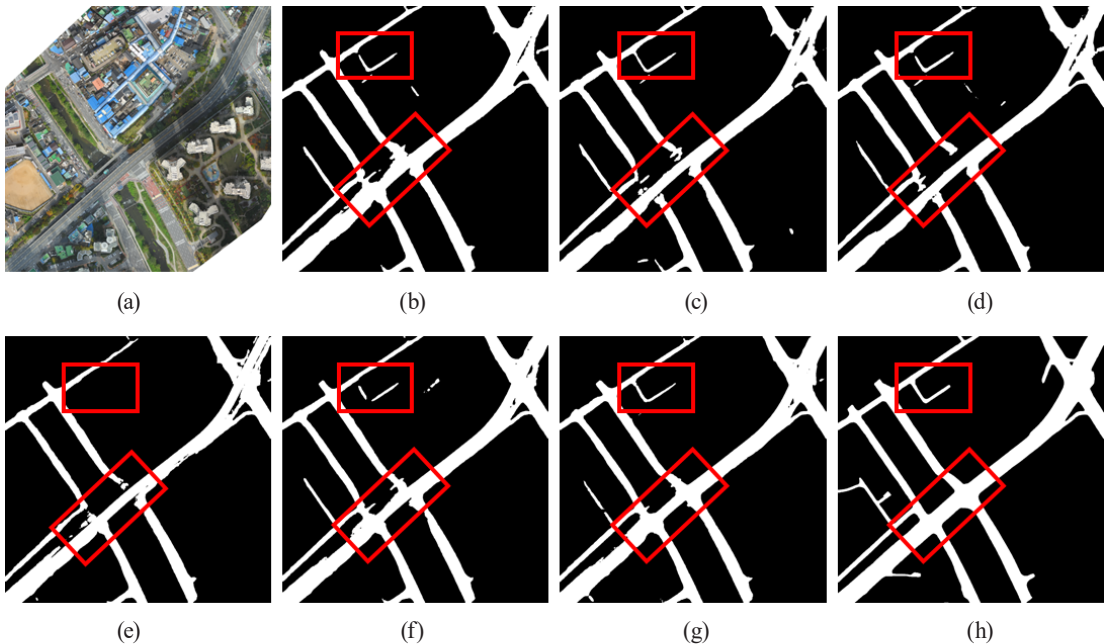


Fig. 5. Comparison of each model inference result (a) Drone image, (b) NL-LinkNet(Resnet-18), (c) D-LinkNet(ResNet-18), (d) LinkNet(ResNet-18), (e) NL-LinkNet(ResNet-34), (f) D-LinkNet(ResNet-34), (g) LinkNet(ResNet-34), (h) Ground Truth

Table 2. Model performance assessment results

Models	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%

본 연구에선 가장 높은 성능을 나타낸 LinkNet-34의 성능을 더 높이기 위해 하이퍼파라미터 최적화를 적용하였다. 모델에 가장 최적화된 하이퍼파라미터의 수치는 이미지 크기 1024×1024, learning rate 0.0005, epoch 800, weight decay factor 5.0, batch size 24이며 하이퍼파라미터 최적화를 적용하지 않은 이전 모델 보다 0.204% 더 높은 85.125%의 성능을 나타냈다. Fig. 6은 LinkNet-34 모델에 하이퍼파라미터 최적화를 했을 경우와 하지 않았을 경우의 도로 추출 결과를 보여준다.

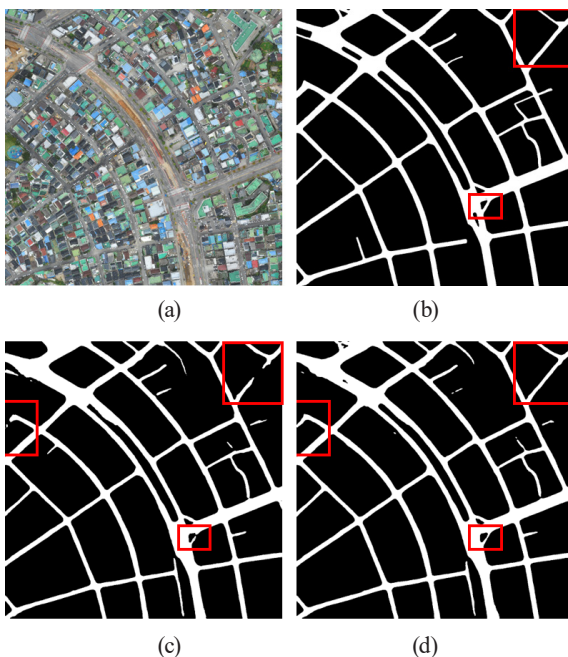


Fig. 6. Comparison of inference results (a) drone image source, (b) ground truth, (c) LinkNet(ResNet-34) before hyperparameter optimization, and (d) LinkNet(ResNet-34) after hyperparameter optimization

5. 요약 및 결론

본 연구의 목적은 기존 지도 구축의 효율화를 위해 고해상도 정사영상을 기반으로 신속한 도로 추출을 수행하는 것이다. 이러한 연구 목적을 달성하기 위하여 도로의 상태가 빠르게 변화하는 광주광역시 도시철도 2호선 공사 현장을 촬영한 드론 정사영상에서 도로 윤곽을 추출하기 위해 동일한 환경에서 LinkNet, D-LinkNet, NL-LinkNet 알고리즘에 pretrained encoder로 ResNet-18과 ResNet-34를 달리 적용하였다. 그 결과, ResNet-34를 pretrained encoder로 사용한 LinkNet의 추론 결과의 mIoU가 84.921%로 동일한 encoder를 사용한 D-LinkNet과 NL-LinkNet보다 0.435%, 0.659% 더 높은 성능을 보였다. 이후 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. 향후 연구로 hybrid 구조의 모델 혹은 pseudo-labeling 기법과 같은 semi-supervised Learning 기반의 semantic segmentation 알고리즘과의 비교 분석 연구를 진행할 예정이다.

감사의 글

이 논문은 2021년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2021 R1C1C1012785). 또한 본 논문은 행정안전부 “극한재난대응기반기술개발사업(20017423)”의 지원을 받아 작성되었음.

References

- 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*, Vol. 12, No. 9, pp. 1444-1465. <https://doi.org/10.3390/rs12091444>
- 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*, Vol. 130, No. 2017, pp. 139-149.

- <https://doi.org/10.1016/j.isprsjprs.2017.05.002>
- Awad, M. M., (2013), A morphological model for extracting road networks from high-resolution satellite images, *Journal of Engineering*, Vol. 2013, pp. 243021 1-9.
<https://doi.org/10.1155/2013/243021>
- Chaurasia, A. and Culurciello, E., (2017), Linknet: Exploiting encoder representations for efficient semantic segmentation, *IEEE Visual Communications and Image Processing-2017*, 10-13 December, St. Petersburg, FL, USA, pp. 1-4.
<https://doi.org/10.1109/VCIP.2017.8305148>
- 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-2018*, 18-22 June, Salt Lake City, UT, USA, pp. 172-181.
<https://doi.org/10.1109/CVPRW.2018.00031>
- 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-2016*, 27-30 June, Las Vegas, NV, USA, pp. 770-778.
<https://doi.org/10.48550/arXiv.1512.03385>
- Jang, Y.J., Oh, J.H., and Lee, C.N., (2020), Urban Building Change Detection Using nDSM and Road Extraction, *Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography*, Vol. 38, No. 3, pp. 237-246.
<https://doi.org/10.7848/ksgpc.2020.38.3.237>
- Kim, J.Y., Huh, Y, Yu, K.Y., and Kim, J.O., (2013), Automatic Change Detection Based on Areal Feature Matching in Different Network Data-sets, *Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography*, Vol. 31, No. 6, pp. 483-491.
<https://doi.org/10.7848/ksgpc.2013.31.6-1.483>
- Liu, W. and Wang, H., (2008), An interactive image segmentation method based on graph theory, *J. Electron. Inf. Technol*, Vol. 8, No. 30, pp. 1973-1976.
- 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-2015*, Boston, 7-12 June, MA, USA, pp. 3431-3440.
<https://doi.org/10.48550/arXiv.1411.4038>
- 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*, Vol. 6, No. 1, pp. 063610 1-14. (in Korean with English abstract)
<https://doi.org/10.1117/1.JRS.6.063610>
- 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*, Vol. 26, No. 9, pp. 1201-1220.
<https://doi.org/10.1016/j.patrec.2004.11.005>
- 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*, Vol. 11, No. 11, pp. 1856-1860.
- Mnih, V. and Hinton, G. E., (2010), Learning to detect roads in high-resolution aerial images, *European conference on computer vision-2010*, 5-11 September, Heraklion, Crete, Greece, pp. 210-223.
https://doi.org/10.1007/978-3-642-15567-3_16
- 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-2017*, 27-29 December, Singapore, Singapore, pp. 191-201.
https://doi.org/10.1007/978-3-319-60663-7_18
- Park, Y.K., Kang, W.P., Choi, J.E., and Kim, B.J., (2019), A study on the Evaluation of Real-Time Map Update Technology for Automated Driving, *Journal of the Korean Association of Geographic Information Studies*, Vol. 22, No. 3, pp. 146-154. (in Korean with English abstract)
<https://doi.org/10.11108/kagis.2019.22.3.146>
- 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*, Vol. 52, No. 6, pp. 3359-3372.
<https://doi.org/10.1109/TGRS.2013.2272593>
- Tao, W. B. and Jin, H., (2007), A novel method of image

- threshold segmentation based on graph theory, *Chinese Journal of Computer*, Vol. 1, pp. 110-119.
- Unsalan, C. and Sirmacek, B., (2012), Road network detection using probabilistic and graph theoretical methods, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 50, No. 11, pp. 4441-4453.
<https://doi.org/10.1109/TGRS.2012.2190078>
- Varia, N., Dokania, A. & Senthilnath, J., (2018), DeepExt: A convolution neural network for road extraction using RGB images captured by UAV, *IEEE Symposium Series on Computational Intelligence-2018*, 18-21 November, Bengaluru, India, pp. 1890-1895.
<https://doi.org/10.1109/SSCI.2018.8628717>
- Vosselman, G. and Knecht, J. D., (1995), Road tracing by profile matching and Kaiman filtering, *Automatic extraction of man-made objects from aerial and space images*, Springer, pp. 265-274.
https://doi.org/10.1007/978-3-0348-9242-1_25
- 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, *IEEE Geoscience and Remote Sensing Symposium-2014*, 13-18 July, Quebec City, QC, Canada, pp. 533-536.
<https://doi.org/10.1109/IGARSS.2014.6946477>
- 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.
<https://doi.org/10.1080/01431161.2015.1054049>
- 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.
<https://doi.org/10.1109/CVPR.2018.00813>
- 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*, Vol. 19, pp. 1-5.
<https://doi.org/10.1109/LGRS.2021.3050477>
- Wei, Y., Wang, Z., and Xu, M., (2017), Road structure refined CNN for road extraction in aerial image, *IEEE Geoscience and Remote Sensing Letters*, Vol. 14, No. 5, pp. 709-713.
<https://doi.org/10.1109/LGRS.2017.2672734>
- Yun, B.Y., Moon, D.Y., and Hong, S.H., (2006), A Study on Updating of Digital Map using Beacon GPS, *Journal of the Korean Geophysical Society*, Vol. 9, No. 4, pp. 387-395. (in Korean with English abstract)
- Zhang, C., Murai, S., and Baltsavias, E. P., (1999), Road network detection by mathematical morphology, *ISPRS Workshop 3D Geospatial Data Production: Meeting Application Requirements*, 7-9 April, Paris, France, pp. 185-200.
<https://doi.org/10.3929/ethz-a-004334280>
- 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-2018*, 18-22 June, Salt Lake City, UT, USA, pp. 182-186.
<https://doi.org/10.1109/CVPRW.2018.00034>
- 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*, Vol. 26, No. 24, pp. 5493-5508.
<https://doi.org/10.1080/01431160500300354>