

Real-time Segmentation of Black Ice Region in Infrared Road Images

Yu-Jie Li*, Sun-Kyoung Kang**, Sung-Tae Jung**

*Lecturer, School of Compute Science, Weifang University of Science and Technology, Shandong, China

Student, Dept. of Computer and Software Engineering, Wonkwang University, Jeonbuk, Korea

**Professor, Dept. of Computer and Software Engineering, Wonkwang University, Jeonbuk, Korea

**Professor, Dept. of Computer and Software Engineering, Wonkwang University, Jeonbuk, Korea

[Abstract]

In this paper, we proposed a deep learning model based on multi-scale dilated convolution feature fusion for the segmentation of black ice region in road image to send black ice warning to drivers in real time. In the proposed multi-scale dilated convolution feature fusion network, different dilated ratio convolutions are connected in parallel in the encoder blocks, and different dilated ratios are used in different resolution feature maps, and multi-layer feature information are fused together. The multi-scale dilated convolution feature fusion improves the performance by diversifying and expending the receptive field of the network and by preserving detailed space information and enhancing the effectiveness of diated convolutions. The performance of the proposed network model was gradually improved with the increase of the number of dilated convolution branch. The mIoU value of the proposed method is 96.46%, which was higher than the existing networks such as U-Net, FCN, PSPNet, ENet, LinkNet. The parameter was 1,858K, which was 6 times smaller than the existing LinkNet model. From the experimental results of Jetson Nano, the FPS of the proposed method was 3.63, which can realize segmentation of black ice field in real time.

▶ **Key words:** Black Ice, Image Segmentation, Dilated Convolution, Receptive Field

[요 약]

본 논문에서는 운전자한테 실시간으로 블랙 아이스 경고를 보내기 위해서 도로 영상에서 블랙 아이스 영역 분할을 위한 다중 척도 팽창 컨볼루션 특징 융합에 기반한 딥러닝 모델을 제안한다. 제안한 다중 척도 팽창 컨볼루션 특징 융합 네트워크는 인코더 블록에 서로 다른 팽창 비율 컨볼루션을 병렬로 추가하고, 서로 다른 해상도 특징 맵에서 서로 다른 팽창 비율을 설정하고, 다중 단계 특징 정보가 함께 융합된다. 다중 척도 팽창 컨볼루션 특징 융합은 수용 영역을 확장함과 동시에 공간의 세부 정보를 잘 보존하고 팽창 컨볼루션의 효과성을 높임으로써 기존 모델보다 성능을 향상시킨다. 실험 결과를 통해 본 논문 제안한 네트워크 모델은 병렬 팽창 컨볼루션 수가 증가함에 따라 성능이 향상되는 것을 알 수 있었다. 제안한 방법의 mIoU 값은 96.46%로 U-Net, FCN, PSPNet, ENet, LinkNet 등 기존 네트워크보다 높았다. 그리고 파라미터는 1,858K개로, 기존 LinkNet 모델보다 6배로 축소하였다. Jetson Nano에서 실험 결과 보면, 제안한 방법의 FPS는 3.63로 실시간으로 블랙 아이스 영역을 실시간으로 분할 할 수 있었다.

▶ **주제어:** 블랙 아이스, 이미지 분할, 팽창 컨볼루션, 수용 영역

-
- First Author: Yu-Jie Li, Corresponding Author: Sung-Tae Jung
 - *Yu-Jie Li (trista0211@wku.ac.kr), School of Computer Science, Weifang University of Science and Technology; Dept. of Computer and Software Engineering, Wonkwang University
 - **Sun-Kyoung Kang (doctor10@wku.ac.kr), Dept. of Computer and Software Engineering, Wonkwang University
 - **Sung-Tae Jung (stjung@wku.ac.kr), Dept. of Computer and Software Engineering, Wonkwang University
 - Received: 2022. 01. 19, Revised: 2022. 02. 14, Accepted: 2022. 02. 14.

I. Introduction

2021년의 교통사고 통계 분석 보고서에 따르면 한국에서 교통사고로 인한 사망자 수가 교통 선진국과 차이가 여전히 존재하며, 연간 수십만 건의 교통사고가 발생하는 현실을 고려할 때 사회적 관심과 노력이 필요한 실정이다[1]. 교통 안전 사고 발생에 영향을 미치는 요인은 매우 많다. 예를 들면 인위적인 요인, 차량 자체의 문제, 도로 상황, 기후 환경 등이다. 이 중 기후 환경으로 인한 도로상황 변화는 제 때 파악이 어려워 교통사고로 이어질 수가 있다. 특히 겨울철에는 노면 온도가 0°C 이하일 때 노면의 고인 물이나 습한 공기가 노면 표면에서 얇은 얼음 층으로 엉키는데, 이 얇은 얼음 층이 도로의 거친 표면 틈새에 매워져 아스팔트 등 어두운 노면에서 검은색으로 나타나기 때문에 블랙 아이스라는 이름이 붙었다[2]. 그림 1에서 보는 바와 같이 블랙 아이스는 젖은 노면과 시각적으로 접근하여 사람 육안으로 쉽게 인식되지 않으며, 제 때 감지되지 않으면 큰 교통사고를 유발하기 쉽다.



Fig. 1. Black Ice Road

블랙 아이스 교통사고는 기온이 영하로 떨어지는 12월부터 급증하기 시작해 평균 최저기온이 영하 5.6도로 제일 낮은 1월에 가장 많이 발생한다[3]. 도로교통공단의 실험 결과에 따르면 도로가 결빙된 상태에서 시속 100km로 주행했을 때 제동 거리는 203.9m였다. 도로가 마른 상태였을 때의 41.9m 보다 최대 5배까지 길어진 것이고 도로가 축축한 상태였을 때의 83.8m보다 최대 2.5배까지 길어진 것이다[4]. 블랙 아이스로 덮인 노면은 젖은 노면과 차이가 없어 보이지만 이 두 경우 제동거리가 120m 정도 차이가 나기 때문에 블랙 아이스 경보장치 없이 운전자가 블랙 아이스 노면을 젖은 노면이라고 잘못 판단하는 경우에 교통사고가 일어날 가능성이 높다.

한국교통연구원에서 2015년부터 2019년까지 4년간의 교통사고를 분석한 결과에 따르면 블랙 아이스 교통사고

발생 건수는 5,200건으로 눈길 교통사고 2,884건의 1.8배에 달했다. 이 중에서 블랙 아이스 교통사고 사망자는 170명으로 눈 교통사고로 인한 사망자 46명의 3.7배에 달했다. 또한 사망률을 비교하면 블랙 아이스 교통사고 치사율은 3.3%로 눈 교통사고의 1.6배에 달했다[5]. 눈은 사람이 쉽게 식별할 수 있어서 주의를 기울이지만 블랙 아이스는 잘 보이지 않아 대비가 쉽지 않기 때문에 대형 사고로 이어지기가 쉽다. 특히 겨울철 급격한 기온 강하는 밤늦게 발생하며, 많은 블랙 아이스 현상은 야간에 발생하기 때문에 인력에 의존해 24시간에 노면 모니터링 시스템을 통해 관측할 수 없다. 그리고 빙판길의 불균일성 때문에 제빙 작업을 광범위로 하면 인력 낭비가 발생할 수 있다.

따라서 주행도로에서 블랙 아이스 영역을 정확히 식별할 수 있는 노면 탐지 장치가 필요하며 블랙 아이스 발생 시 운전자에게 블랙 아이스 경고를 미리 보내 운전자가 미리 제동을 걸 수 있게 함으로써 교통사고 발생 확률을 낮출 필요가 있다.

최근 몇 년간 카메라 및 영상 분석 기술은 도로 조건 판단에 성공적으로 활용되었다[6]. 신경망을 이용해 도로 결빙 상태를 분석하는 검사 정밀도는 80~90%, 마른 노면과 젖은 노면을 구분은 45%~70%에 이른다[7]. 영상을 이용해서 도로 블랙 아이스 감지에 관한 연구는 다음과 같다.

Qi Li은 OpenCV+Python 기반 도로 결빙 감지 시스템을 설계해 에지 검출과 이진 형태학을 결합한 뒤 적응형 웨이블릿 알고리즘으로 노이즈를 제거한 뒤 노면 이미지의 특징 값을 SVM 분류기를 통해 건조, 고인 물, 적설, 결빙의 네 가지 노면 이미지를 인식했고 평균 분류 정확도는 80.4%였다[8].

Youngis E. Abdalla는 3차원 깊이 이미지 센서인 Kinect를 이용한 블랙 아이스 감지 시스템을 제안했다[9]. 제안된 시스템은 카메라로부터 82cm에서 1.52m 거리에서 나무, 유리, 세라믹, 플라스틱, 콘크리트의 5가지 재료에 부드러운 얼음, 젖은 눈, 딱딱한 얼음, 블랙 아이스, 글레이즈의 4가지 유형의 분류 실험을 실시하였으나 실험결과 인식정확률을 구체적으로 기술하지 못하였으며 도로에 대한 결빙 검지는 실시하지 않았다.

딥러닝 기술이 널리 적용됨에 따라 Hojun Lee는 구글 이미지 검색을 통해 블랙 아이스 검출 데이터셋을 만들고 CNN 딥러닝 기법을 이용해 마른 노면, 젖은 노면, 눈길, 블랙 아이스 노면의 네 가지 노면을 분류했으며, 블랙 아이스 인식률은 96%에 달했지만 블랙 아이스 영역을 분할하지 않았다[10].

영상에 기반한 블랙 아이스 감지 기술은 강력한 기능과 직관적인 이미지, 소프트웨어 제어로 업그레이드가 용이해

비용이 적게 들고, 설치와 수리 시 노면 파괴나 차선 폐쇄 없이 다양한 요구에 따라 위치를 재설정할 수 있다는 장점이 있다. 그러므로 본 논문에서는 영상에 기반한 블랙 아이스 검출에 관한 연구를 실시하였다. 그러나 대부분의 도로 감시 카메라는 가시광선 RGB 이미지를 획득하는 일반 카메라를 채택하고 있으며 가시광선의 스펙트럼 범위가 0.43~0.79 μm 로 빛의 조도가 충분하지 않은 상태에서는 블랙 아이스의 구분이 불가능하다. 반면에 열화상 카메라는 온도가 절대 영도보다 높은 모든 물질에서 나오는 열복사로 열화상을 만들어내 조도에 영향을 받지 않는 장점을 가지고 있다[11]. 열화상 영상은 목표 인식과 검출 분야에서 대체 불가한 역할을 하며 보안 감시, 원격 탐지 영상, 국방 등에 폭넓게 활용되고 있다. 따라서 본 논문에서는 열화상 카메라를 채용하여 블랙 아이스 영역에 대한 실시간 검출 방법을 제안한다.

블랙 아이스 영역에 대해 미리 경고해 주는 것은 주행 안전에 매우 중요한 역할을 하며, 원거리에서 블랙 아이스 영역의 위치와 면적의 크기를 어떻게 정확하게 제시하고 동적으로 경고를 할 수 있는지가 가장 중요한 문제이다. 기존의 블랙 아이스 영역 검지에 방법은 다음과 같이 문제점을 가지고 있다.

- 1) 블랙 아이스 검출의 정확도가 낮아 블랙 아이스 상태와 젖은 상태를 정확히 구별하기 어렵고 오판단을 하기 쉽다.
- 2) 블랙 아이스의 심각 정도에 대해서는 판단할 수 없으며 블랙 아이스의 유무만을 간단히 판단할 뿐 블랙 아이스 지역의 면적을 계량화하지 않는다.
- 3) 블랙 아이스 검출의 실시간성이 보장되지 않으며, 카메라를 일반카메라로 채택하고 있으며, 어두운 밤이나 극한의 날씨에 대해서는 영상 화질이 떨어지며, 검출의 정확도에 영향을 미친다.

본 논문에서는 열화상 카메라를 사용하여 임베디드 장치 장치에서 블랙 아이스 영역을 실시간으로 정확하게 분할하는 것을 주요 문제로 삼고, 열화 도로 블랙 아이스 데이터셋을 구축하며, 딥러닝 방법을 사용하여 작은 크기의 블랙 아이스 분할 네트워크 모델을 만든 다음에 Jetson Nano 보드에서 블랙 아이스 분할 네트워크를 배치하여 실시간으로 이미지 수집하여 블랙 아이스 영역 분할해서 블랙 아이스 등급 경고 신호를 제공하도록 하였다.

본 논문에서 제시한 방법으로는 그림 2와 같이 블랙 아이스 영역 분할 네트워크를 학습시키는 오프라인 모듈과 그림 3과 같이 실시간 블랙 아이스 영역 분할 온라인 모듈이 포함된다.

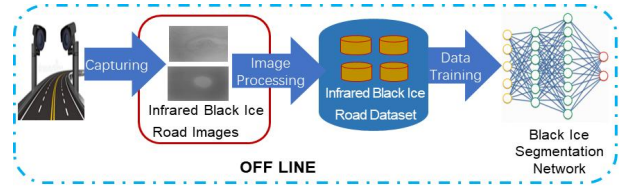


Fig. 2. The Module of Black Ice Semantic Segmentation Network Training

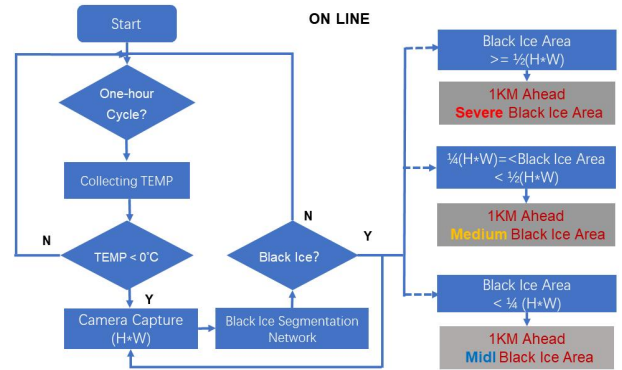


Fig. 3. The Module of Black Ice Semantic Segmentation and Warning

II. Related Work

2.1 Image segmentation with convolution network

이미지 분할은 컴퓨터 비전 분야에서 해결해야 할 기본적인 문제로서 많은 연구가 수행되고 있다. 최근에는 딥러닝 기반 이미지 분할 알고리즘들이 많이 제안되었다. Long은 이미지 분할을 위한 딥러닝 모델로 컨볼루션 레이어만으로 구성된 FCN(Fully Convolutional Neural Networks)을 제안하였다[12]. FCN은 기존에 이미지 분류에 사용되던 컨볼루션 레이어와 완전히 연결된 레이어로 구성된 딥러닝 모델에서 완전히 연결된 레이어를 제거하고, 업샘플링 작업을 통해 이미지 분할이 가능하게 만들었다. FCN은 풀링 레이어를 거치면서 이미지 해상도가 계속 줄어들어 이미지 정보가 많이 손실되고 업샘플링 레이어를 거치면서 원래 크기로 복원되는 과정에서 이미지의 컨텍스트 정보를 충분히 고려하지 않아 FCN을 이용해서 이미지를 분할할 때 에지의 분할이 세밀하지 못하다. 이 모델은 PASCALVOC 2012 데이터셋[13]에서 분할 정확도가 62.2%에 달했다.

Olaf Ronneberger는 의학적 이미지 분할 문제를 해결하기 위해 U-Net 모델을 제시하였다[14]. U-Net는 전형적인 이미지 분할을 위한 인코더-디코더 구조이다. U-Net 네트워크의 중요한 특징 중 하나는 완전 대칭 구조라는 것이며, 전체 네트워크 구조는 대문자 U자 형태이므로 U-Net라고

한다. U-Net에서는 인코더를 이용하여 저차원부터 고차원까지의 피라미드형 특징 맵을 생성하고, 디코더를 이용하여 특징 피라미드 대응 계층의 특징 맵을 생성하며, 다운 샘플링에 의해 얻어지는 저해상도의 특징 맵과 업샘플링에 의해 생성되는 특징 피라미드 대응 계층의 특징 맵을 연결하여 사용함으로써 이미지 컨텍스트 정보의 손실을 줄여 분할 성능을 향상시켰다. 이 모델은 PASCALVOC 2012 데이터셋에서 분할 정확도가 72.5%에 달했다.

Zhao는 PSPNet(Pyramid Scene Parsing Network)을 제안하였다[15]. 이 모델은 사전 훈련된 ResNet[16]과 팽창 네트워크를 결합하여 이미지의 특징을 추출하여 원래 이미지 1/8 크기의 특징 맵을 획득한 다음에 병렬로 연결된 4개의 풀링 레이어를 통하여 크기가 다른 4개 출력 특징 맵을 획득하며, 이 4개 특징 맵을 업샘플링을 통해서 원래 이미지 크기로 공간 정보를 복원하고 마지막으로 특징 맵을 연결해서 컨볼루션을 통해 이미지 분할을 수행하였다. 이 모델은 공간 피라미드 풀링 모듈과 직렬 모델 등을 통해 이미지의 척도별, 위치별 특징 정보를 포착하여 융합하고 분할 결과를 점차 미세화하여 큰 계산량과 높은 메모리 점유 등의 문제를 최적화하였다. 이 모델은 PASCALVOC 2012 데이터셋에서 85.4% 정확도를 보였다.

Chen은 DeepLabV3+ 모델을 제안하였다[17]. DeepLabV3+는 Xception[18]을 기반으로 하여 인코더-디코더 구조로 되어 있으며, 팽창 컨볼루션[19]을 통해 수용 영역을 확장시키고, ASPP(Atrous Spatial Pyramid Pooling)[17]을 통해 다중 척도 특징 융합을 통하여 이미지 컨텍스트 정보를 포착함으로써 분할 결과를 점진적으로 미세화하고, CRF(Conditional Random Field)이라는 후 처리 방법을 도입해서 이미지 지역 정보를 복원하여 분할 결과를 최적화하였다. 이 모델은 PASCALVOC 2012 데이터셋에서 89.0%의 분할 정확도를 보였다.

FCN, U-Net, PSPNet, DeeplabV3+는 좋은 분할 결과를 얻을 수 있지만, 네트워크 훈련에서 파라미터 가중치, 계산량 과다 등의 문제가 있어 실시간 분할 처리가 어렵다. 보다 빠른 모델로 Adam Paszke는 인코더-디코더 구조를 기반으로 효율적인 신경망 ENet을 제안하였다[20]. 이 모델은 ResNet을 참조하여 병목 모듈을 제안하였으며, 병목 모듈을 이용하여 팽창률에 따른 팽창 컨볼루션을 직렬로 실행해서 모델의 오버샘플링에 따른 특징 맵 해상도 저하 문제를 완화하였다. ENet은 분할 작업을 수행할 때 필터 분해 정책을 사용하여 작은 크기의 랭크 근사화를 사용하여 컨볼루션 작업을 보다 간단한 작업으로 분해하여 분할 정밀도를 보장하면서 계산량을 현저히 감소시킨 실

시간 이미지 분할 네트워크이다.

Chaurasia는 LinkNet 딥러닝 네트워크 모델을 제안하였다[21]. LinkNet은 ResNet의 잔차 모듈을 통해 네트워크 특징 추출 능력을 향상시키고 네트워크 역방향 전파 가능 네트워크의 얇은 부분으로 쉽게 전파되도록 하여 네트워크의 최적화를 보다 간단하게 하고, 네트워크 메모리의 소모를 줄임으로써 네트워크의 계산 효율을 높이고, 실시간으로 이미지 분할을 가능하게 한다.

기존 네트워크 모델들이 PASCALVOC 2012와 CityScapes 데이터셋에서 이미지 분할 정확도 비교 결과는 다음 표1과 같다.

Table 1. Experimental Comparison of Segmentation Accuracy for Classical Network Models

Model	Dataset	mIoU
FCN	PASCALVOC 2012	62.2%
U-Net	PASCALVOC 2012	72.5%
PSPNet	PASCALVOC 2012	85.4%
DeepLabV3+	PASCALVOC 2012	89.0%
ENet	CityScapes	58.3%
LinkNet	CityScapes	76.4%

III. The Proposed Method

본 논문에서 제시하는 네트워크 아키텍처는 인코더와 디코더의 두 부분으로 나뉘며, 그림 4에 나타나있는 바와 같이 인코더는 3단계의 인코더 블록으로 구성되고 디코더는 4단계의 디코더 블록으로 구성된다. 모델 크기를 작게 하기 위하여 인코더 블록 앞 단계에는 2개의 컨볼루션 레이어를 통하여 해상도를 4분의 1로 줄이고 디코더를 통과한 다음에는 컨볼루션 레이어를 통하여 원래 영상 크기로 복원한다. 인코더 블록에서 수용 영역을 확장하는 방식으로 특징 맵 정보를 증가시키고, 다중 척도 팽창 컨볼루션을 병렬로 연결해서 특징을 융합하는 방식으로 열화 도로 영상의 특징 정보를 추출한다. 디코더 블록에서는 업샘플링 대신에 전치된 컨볼루션을 사용하여 이미지 크기 및 세부 특징 정보를 복원하는 작업을 수행하였다. 각 컨볼루션 레이어는 Relu 레이어를 통해 활성화된 후 배치 정규화로 균일화 작업을 수행하였다. 그림 4에 나타나 있는 conv[(3x3), (3, 32), /2] 또는 conv[(3x3), (32, 3), *2]에서 conv는 컨볼루션을 나타내고, 첫 번째 값 (3x3)은 컨볼루션 커널 크기가 3x3이라는 것을 나타낸다. 두 번째 값 (3, 32)와 (32, 3)에서 첫 번째 값은 입력 채널수를 나타내

고 두 번째 값은 출력 채널수를 나타낸다. 세 번째 값에서 /2는 stride이 2인 다운샘플링을 나타내고, *2는 2배인 업샘플링을 나타낸다.

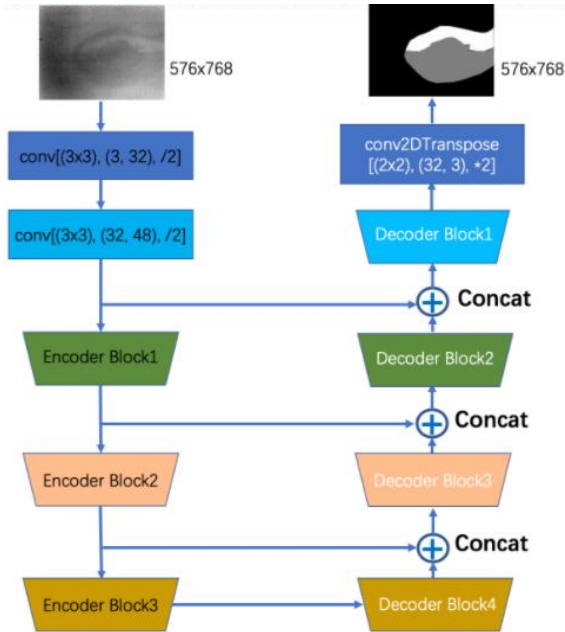


Fig. 4. The Proposed Network Architecture

3.1 Encoder

입력 영상의 크기는 576x768인데 먼저 출력 특징 채널이 각각 32와 48이고 2픽셀씩 건너뛰면서 수행하는 3*3 컨볼루션 2개 레이어를 통해 이미지의 크기를 144*192로 줄인다.

인코더 블록에서는 그림 5에 나타나 있는 바와 같은 팽창 컨볼루션을 사용한다. 팽창 컨볼루션은 컨볼루션에 사용되는 픽셀을 그림 5에서 색칠된 픽셀과 같이 팽창 비율에 따라 건너뛰면서 선택함으로써 표준 컨볼루션 연산에 비해 파라미터 수를 증가시키지 않고도 더 큰 수용 영역을 얻을 수 있으며, 팽창 컨볼루션에 의한 특징 맵은 입력 이미지 크기와 같을 수 있지만 출력 뉴런마다 더 큰 수용 영역을 가지기 때문에 보다 고차원적인 의미의 인코딩이 가능하다.

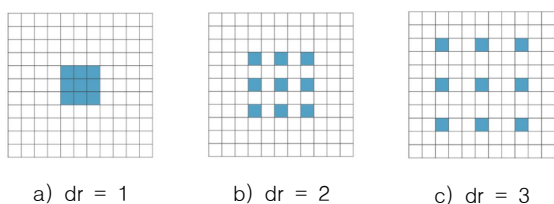


Fig. 5. Dilated Convolution

그림 5에서 dr은 팽창 비율을 나타내는데 공식 (1)에 따라 $dr = n$ 인 팽창 컨볼루션의 수용 영역 크기를 계산할 수 있다.

$$F_n = (2n + 1) \times (2n + 1) \quad \text{식(1)}$$

그림 5(a)는 팽창 비율이 1인 경우로 표준 컨볼루션이다. 그림 5(b)는 팽창 비율이 2인 팽창 컨볼루션으로 수용 영역의 크기는 $F_2=5*5$ 이다. 그림 5(c)는 팽창 비율이 3인 팽창 컨볼루션이며, 수용 영역 크기는 $F_3=7*7$ 이다. 팽창 컨볼루션의 목적은 특징 맵을 축소시키지 않고서도 똑같이 수용 영역을 확장시켜서 컨볼루션된 출력이 더 넓은 범위의 정보를 가질 수 있도록 하는 것이다.

팽창 컨볼루션은 특징 맵 해상도와 수용 영역 사이의 상충 문제를 해결할 수 있지만, 팽창 컨볼루션을 통해 출력된 특징 맵에서는 모든 뉴런이 동일한 수용 영역을 가지며, 의미 마스크 생성 과정이 단일 척도 특징만을 이용한다는 것을 의미한다. 그러나 다중 척도 정보는 경계가 모호한 상황을 해결하고 보다 강건한 추출 결과를 낳는 데 도움이 될 것이다. 특히 열화상 이미지는 가시광선 이미지에 비해 경계가 희미하고 명암대비가 약한 특징이 있기 때문에 본 논문에서 서로 다른 팽창 비율을 가진 여러 개의 팽창 컨볼루션 레이어를 병렬로 연결하여 다중 척도 정보를 포착한다. 또한 인코더 블록의 단계별로 특징 맵의 크기가 작아질 때 팽창 비율도 함께 작아지게 하여 적절한 수용 영역이 유지되도록 한다.

그림 6에 인코더 블록의 구조가 나타나 있다. 하나의 인코더 블록은 여러 개의 컨볼루션 경로가 병렬로 연결된 구조이며 각 경로의 결과 특징 맵은 직렬로 연결되어 출력된다. 각각의 경로는 ResNet 구조를 가지고 있으며 각 경로마다 팽창 비율이 다른 팽창 컨볼루션을 사용하고 있다.

제안된 인코더 블록은 서로 다른 팽창 비율의 팽창 컨볼루션 레이어를 병렬로 결합하여 더 큰 수용 영역에서 더 많은 척도 특징을 발생시킨다. 일련의 특징 연결을 통해 각 중간 특징 맵의 뉴런은 여러 척도로부터 의미 정보를 인코딩하고, 다른 중간 특징 맵은 다른 수용 영역으로부터 다중 척도 정보를 인코딩한다. 일련의 팽창 컨볼루션을 통해 퇴화 문제없이 뒤에 있는 레이어의 뉴런에 비해 점점 더 큰 수용 영역을 얻을 수 있다.

큰 수용 영역의 컨볼루션은 큰 목표의 세부정보를 추출할 수 있으며, 작은 수용 영역의 컨볼루션은 작은 목표에 더욱 효과적이다. 따라서, 본 논문에서 제시한 인코더 블록을 통하여 최종 출력되는 특징 맵은 넓은 범위의 의미

정보를 포함할 뿐만 아니라 작은 범위의 인접 영역에 대한 의미 정보를 포함하고 있다.

제안된 인코더 블록에서는 병렬 팽창 컨볼루션을 사용하여 특징을 직렬 연결하므로 병렬 레이어 수가 많으면 특징 맵의 차원이 너무 크고 연산량이 증가하는 문제가 발생하므로 제안한 네트워크 구조는 계산량을 줄이기 위해 각 인코더 블록 시작 부분에 채널 크기를 절반으로 줄이는 컨볼루션 레이어를 사용한다.

제안된 인코더 블록의 각 경로에서 컨볼루션 커널 크기는 모두 3이며, 첫 번째 컨볼루션은 2 픽셀씩 건너뛰는 컨볼루션으로써 특징 맵의 크기를 줄이고 마지막 컨볼루션은 팽창 컨볼루션이다.

Encoder_block1의 경우 입력 크기는 144x192, 입력 채널은 48이고, 이때 여러 개의 병렬 팽창 컨볼루션의 팽창 비율은 각각 5, 9, 13, 17, 21과 같이 팽창 비율 간격은 4이고, 출력 채널은 64이다. Encoder_block2의 경우 입력 크기를 72x96이고 팽창 컨볼루션의 팽창 비율이 각각 3, 5, 7, 9, 11과 같이 팽창 비율 간격은 2이고, 출력 채널은 80이다. Encoder_block3의 경우 입력 크기가 36x48이고 팽창 컨볼루션의 팽창 비율이 각각 2, 3, 4, 5, 6과 같이 각 팽창 비율 간격은 1이고, 출력 채널은 96이다. 각 인코더 블록의 출력은 다음 블록의 입력으로 연결되며 마지막으로 디코더 블록에 연결된다.

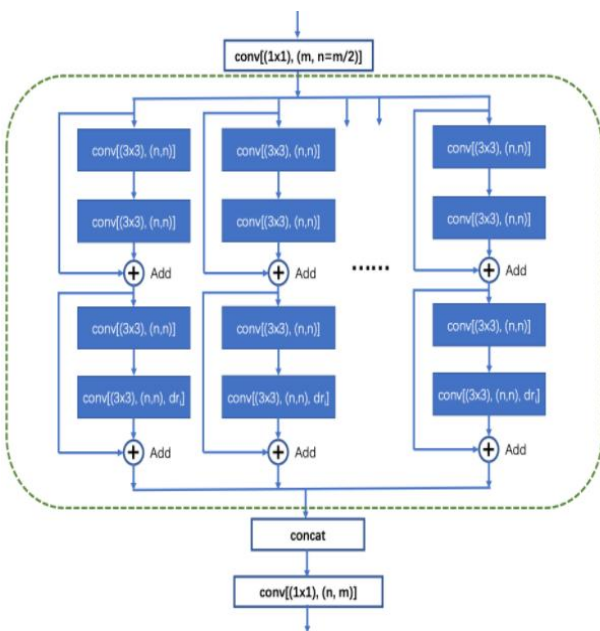


Fig. 6. Encoder-block (j)

3.2 Decoder

본 논문에서 제시한 네트워크의 디코더 부분은 전치된 컨볼루션을 사용하여 특징 맵의 해상도를 36x48에서

576x768로 복원한다. 본 논문에서 사용하는 디코더 블록의 구조는 그림 7과 같다. 먼저, 1x1의 커널을 사용하여 채널 크기를 절반으로 줄인 다음에 3x3 크기의 전치된 컨볼루션을 이용하여 특징 맵의 크기를 키우고 마지막에 다시 1x1 컨볼루션을 실행하여 채널 크기를 키운다. 여기서, m은 입력 특징 맵의 채널 크기를 나타내고, n은 출력 특징 맵의 채널 크기를 나타낸다. 디코더 블록에서 m과 n의 값은 표 2과 같다.

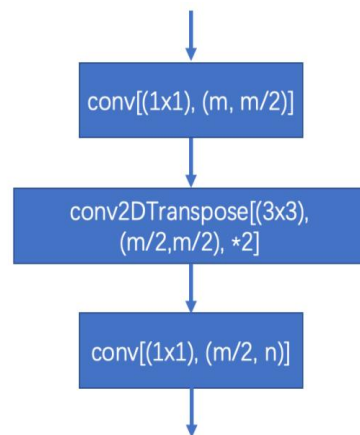


Fig. 7. Decoder_block(j)

Table 2. Channel sizes of Input and output feature maps of Decoder Block

Decoder Blocks	m	n
Decoder Block4	96	80
Decoder Block3	80	64
Decoder Block2	64	48
Decoder Block1	48	32

IV. Experiment

4.1 Infrared Road Black Ice Dataset

논문 [22]에서 만든 블랙 아이스 생성 시뮬레이션 실험을 이용하여 열화 도로 블랙 아이스 데이터 세트를 구축하였다. 시뮬레이션 실험 환경은 그림 8과 같다. 본 논문에서는 새롭게 블랙 아이스 영상을 얻기 위해서 그림 9와 같이 두께 5cm, 가로 세로 1미터 크기의 아스팔트 도로와 시멘트 도로 총 9개를 만든 다음에 냉동 시설에 설치하고 물을 뿌린 다음에 열화상 카메라로 촬영하여 블랙 아이스 영상을 생성하였다. 물을 뿌릴 때는 면적과 위치를 달리하여 서로 다른 9개 케이스를 만들었다. TPV-IAHDR 열화상 카메라를 이용하여 처음부터 블랙 아이스가 생성되는 전 과정을

촬영하였다. 촬영된 동영상에 대해 200ms마다 프레임을 샘플링함으로써 열화 도로 블랙 아이스 이미지를 생성하였다. 논문 [22]에서 구축된 열화 도로 블랙 아이스 영상을 추가해서 최종적으로 10 가지 경우에 대하여 블랙 아이스 도로 이미지를 총 1,156장 생성하였다. 생성된 이미지는 6:2:2의 비율에 따라 훈련 데이터 세트, 검증 데이터와 테스트 데이터 세트로 나뉘었다. 따라서 본 논문에서 구축된 열화 도로 블랙 아이스 데이터 세트는 표 3과 같다.

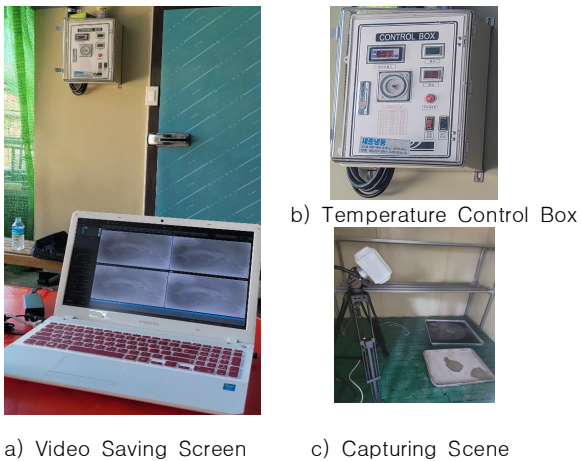


Fig. 8. The Environment of Taking Images of Infrared Road Black Ice Experiment



Fig. 9. The Samples of Asphalt Road (Left) and Cement Road (Right)

Table 3. The Number of Infrared Black Ice Road Image

Type	Black Ice Road
Train	697
Val	229
Test	229
Total	1156

본 논문에서 구축한 열화 도로 블랙 아이스 이미지 데이터 세트의 예는 그림 10과 같다.

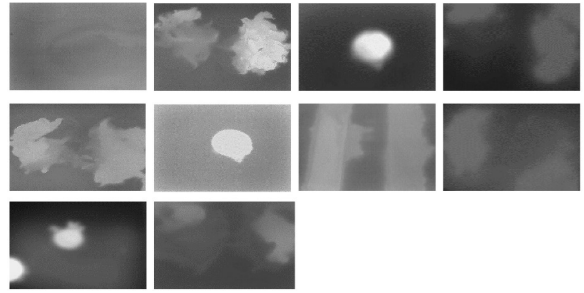


Fig. 10. Examples of Infrared Road Black Ice Image in Different Cases

4.2 Experiments Environment

본 논문의 실험은 다음과 같은 플랫폼에서 진행하였다. 딥러닝 모델 학습에 사용한 운영체제는 Ubuntu18.04 LTS 이고, GPU는 11GB의 메모리가 장착된 NVIDIA GeForce RTX2080ti이고 사용한 딥러닝 프레임 워크는 Keras 및 TensorFlow이다.

4.3 Evaluation metrics

본 논문에서는 분할된 결과와 실제 값의 양자의 교집합과 병합의 비율로 정의되는 mIoU (Mean Intersection Over Union) 지표를 사용하여 블랙 아이스 영역 검출 정확도를 판단하였으며, 그 계산 공식은 다음 식 2와 같다.

$$mIoU = \frac{\sum_{i=1}^N \frac{X_{ii}}{T_i + \sum_{j=1}^N (X_{ji} - X_{ii})}}{N} \quad \text{식(2)}$$

여기서 N 은 이미지 픽셀의 부류 수를 나타내며 T_i 는 부류 i 의 픽셀 총수이고, X_{ii} 은 실제 부류가 i 이고 예측된 부류가 i 인 픽셀 수를 나타내며, X_{ji} 는 실제 부류가 j 이고, 예측된 부류가 i 인 픽셀 수를 나타낸다. mIoU는 간단하고 대표성이 강하여 대부분의 이미지 분할에서 mIoU로 네트워크의 분할 결과를 평가한다.

4.4 Experimental Results and Analysis

실험에서 배치 크기는 8로, epoch는 100으로 설정하였다. 학습 과정에서 손실함수로 교차 엔트로피 손실을 사용하였으며, 최적화기로는 Adam 최적화기를 사용하였다. 네트워크의 학습 과정에서는 매개 변수 초기화와 데이터 선택 등에서 무작위 선택을 사용하기 때문에 학습할 때마다 네트워크 매개 변수가 달라지고 네트워크 성능도 약간씩은 달라진다. 따라서 본 논문에서는 10번 학습과 테스트

과정을 반복하고 결과의 평균값을 사용함으로써 보다 안정적인 결과 값을 산출할 수 있도록 하였다.

표 4에는 본 논문에 제안된 네트워크와 일반적인 이미지 분할 네트워크의 블랙 아이스 영역 분할 결과가 나타나 있다.

Table 4. Experimental Results for Black Ice Dataset

Model	Parameters(KB)	mIoU
U-Net	31,055	70.15%
FCN8	65,810	89.09%
PSPNet101	134,325	85.85%
ENet	371	94.85%
LinkNet	11,555	95.39%
BINet_P2	492	95.93%
BINet_P3	687	96.16%
BINet_P4	882	96.35%
BINet_P5	1,077	96.39%
BINet_P6	1,273	96.40%
BINet_P7	1,468	96.43%
BINet_P8	1,663	96.44%
BINet_P9	1,858	96.46%

표 4에서 본 논문에서 제안한 네트워크 구조는 BINet_Pi 로 표시한다. 여기서, Pi는 인코더 블록에서 컨볼루션들이 병렬로 수행된다는 것을 나타내며, i는 병렬 경로의 수를 나타낸다. Encoder_block1, Encoder_block2, Encoder_block3에서 팽창 비율의 간격은 각각 4, 2, 1이다. 따라서 BINet_P9에서는 병렬 경로가 9개이고 각 경로에서의 팽창 비율은 Encoder_block1에서는 1, 5, 9, 13, 17, 21, 25, 29, 33이고 Encoder_block2에서는 1, 3, 5, 7, 9, 11, 13, 15, 17이고 Encoder_block3에서는 1, 2, 3, 4, 5, 6, 7, 8, 9이다.

Table 4의 블랙 아이스 영역 분할 결과를 보면, 본 논문에서 제안한 여러 가지 네트워크의 블랙 아이스 영역 분할 mIoU 값이 U-Net, FCN, PSPNet, ENet, LinkNet 등 기존의 영상 분할 모델 비하여 모두 향상된 것을 볼 수 있다. 또한 인코더 블록에서 병렬 컨볼루션의 수가 증가하여 수용 영역이 다양화될수록 mIoU 값이 향상되는 것을 볼 수 있으며 9개의 병렬 컨볼루션을 사용한 BINet_P9 네트워크의 mIoU 값이 96.46%로 최고 성능을 보였다. 네트워크의 파라미터 크기를 보면, BINet_P2는 492K에 불과하며, LinkNet의 11,555K보다 23배로 축소할 뿐 아니라 mIoU 값도 향상되는 것을 알 수 있다. 그리고 파라미터 크기가 1,858K인 BINet_P9는 LinkNet 비하여 파라미터 량도 6배로 축소하였고, mIoU 값이 1.07% 향상되었다.

4.5 Black Ice Segmentation on Jetson Nano in Real-time

본 논문에서 채용한 임베디드 장치는 크기가 작고 성능이 좋은 Jetson Nano 4GB이다. 이 시스템의 운영체제는 Ubuntu18.04이며, 사용 소프트웨어는 Tensorflow2.5, Python3.6, CUDA10.2 그리고 TensorRT8.0이다. RTX2080ti GPU를 사용하여 훈련된 모델은 크기가 너무 커서 Jetson Nano와 같은 소규모 임베디드 시스템에서는 그대로 사용하는 것은 불가능하다. 따라서 변환 도구를 이용하여 TensorRT 모델로 변환하여 임베디드 시스템에서 사용될 수 있도록 하였다. 변환된 모델을 Jetson Nano에 설치하여 실시간 열화 도로 블랙 아이스 영역을 분할한 결과는 표 5와 같다. 이 중 FCN8은 네트워크 모델이 너무 커서 TensorRT 모델로 변환할 수 없었다. 표 5에서 FPS는 초당 분할할 수 있는 영상의 총 수를 나타낸다.

Table 5. Experimental Results on Jetson Nano

Model	FPS
U-Net	0.24
PSPNet101	0.14
ENet	1.95
LinkNet	3.72
BINet_P2	5.53
BINet_P3	5.26
BINet_P4	4.98
BINet_P5	4.64
BINet_P6	4.35
BINet_P7	4.10
BINet_P8	3.84
BINet_P9	3.63

표 5에서 볼 수 있듯이 각 네트워크 모델의 분할 속도는 큰 차이를 보인다. 이 중 U-Net, PSPNet는 모델이 크고 계산 량이 많기 때문에 분할 속도가 높지 않고, FPS의 값이 모두 1이하로 실시간으로 블랙 아이스 영역 분할을 만족시킬 수 없다는 것을 알 수 있다. ENet와 LinkNet의 FPS는 각각 1.95, 3.72로 실시간으로 블랙 아이스 영역을 분할할 수 있다. 전체적으로, 본 논문에서 제안한 다중 척도 팽창 컨볼루션을 병렬로 연결해서 특징 융합한 네트워크는 모델 파라미터가 작기 때문에 분할 속도가 빨랐다. BINet_P9의 FPS가 3.63으로 LinkNet보다 약간 낮은 것을 제외하고는, 나머지 모델의 FPS가 LinkNet보다 높다는 것을 보았다. 이 중 BINet_P2는 분할 속도가 5.53 FPS로 가장 빨랐고 나머지 네트워크는 팽창 컨볼루션 병렬 수

가 늘어나면서 계산 량이 커져 분할 속도가 느려지고 FPS 값이 작아졌다.

V. Conclusions

본 논문에서 제안한 다중 척도 팽창 컨볼루션 특징 융합 네트워크는 인코더 블록에 서로 다른 팽창 비율 컨볼루션을 병렬로 추가함으로써 수용 영역을 확장함과 동시에 다층 특징 정보를 융합하여 공간의 세부 정보를 보존하고, 서로 다른 해상도 특징 맵에서 서로 다른 팽창 비율을 설정하는 방식을 사용하여 팽창 컨볼루션의 유효성을 보장하였다.

또한 팽창 컨볼루션 병렬 연결에 따라 네트워크 크기가 커지는 것을 방지하기 위하여 인코더 블록 시작부에 입력 채널 크기를 먼저 2분의1로 축소한 다음에 컨볼루션 작업을 진행하였다. 다운샘플링으로 인한 열화 도로 이미지의 공간 세부 정보 손실을 방지하기 위하여 디코더에서 전치된 컨볼루션 및 인코더의 특징 맵과의 연결을 통하여 열화 도로 블랙 아이스 이미지를 블랙 아이스 영역의 정확한 분할을 가능하게 한다.

본 논문에서 제안한 네트워크 모델과 고전적인 모델은 구축된 열화 도로 블랙 아이스 데이터 세트에서 평가 비교하고, 실험 결과를 통해 본 논문에서 제안한 네트워크는 병렬 팽창 컨볼루션의 수가 증가하여 수용 영역이 다양화될수록 mIoU 값이 향상되는 것을 확인할 수 있으며 9개의 병렬 팽창 컨볼루션을 사용한 BNet_P9 네트워크가 mIoU 값이 96.46%로 가장 높았다. 기존의 영상 분할 모델인 U-Net, FCN, PSPNet, ENet, LinkNet을 열화 도로 블랙 아이스 영상에 대해 학습한 모델에 비하여 본 논문에서 제안된 모델의 크기도 작고 성능도 우수하였으며, Jetson Nano 시스템에서 본 논문에서 제안된 모델은 블랙 아이스 영역을 분할 속도도 더 빠르다는 것을 확인할 수 있었다.

ACKNOWLEDGEMENT

This paper was supported by Wonkwang University in 2021.

REFERENCES

- [1] 2021 Edition (Statistics for 2020) Traffic Accident Statistical Analysis Report, Road Traffic Authority(KoROAD).
- [2] G. Y. Park, S. H. Lee, E. J. Kim, and B. Y. Yun, "A Case Study on Meteorological Analysis of Freezing Rain and Black Ice Formation on the Load at Winter," *Journal of Environmental Science International*, Vol. 26, No. 7, pp. 827-836, Jun. 2017. DOI: 10.5322/JESI.2017.26.7.827
- [3] <https://m.post.naver.com/viewer/postView.nhn?volumeNo=30470274&memberNo=652228>
- [4] O. P. Ghim and T. F. Fwa, "Mechanistic Interpretation of Braking Distance Specifications and Pavement Friction Requirements," *Journal of the Transportation Research Board*, Vol 2155, pp.145-157, Jul. 2010. DOI:10.3141/2155-16.
- [5] Korea Traffic Accident Analysis System [Internet]. Available: <http://taas.koroad.or.kr/>.
- [6] R. Omer, and L. Fu, "An Automatic Image Recognition System For Winter Road Surface Condition Classification," *International IEEE Conference on Intelligent Transportation Systems*, pp.1375-1379, Nov. 2010.
- [7] S. Y. Tan, J. F. Ge, L. Ye and K. Gui, "Road surface icing detecting system based on multiple sensor", *Transducer and Microsystem Technologies*, Vol. 35, No. 1, pp.107-109, Apr. 2016. DOI: 10.13873/J.1000-9787(2016)01-0107-03.
- [8] Q. Li, Y. W. Ji, and Z. P. Wang, "Design of Road Icing Detection System Based on Opencv+Python," *Journal of Shaanxi University of Science & Technology(Natural Science Edition)*, Vol. 35, No. 2, pp. 158-164, Apr. 2017. DOI:10.19481/j.cnki.issn1000-5811.2017.02.030.
- [9] Y. E. Abdalla, M. T. Iqbal and M. Shehata, "Black Ice detection system using Kinect," *2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE)*, pp. 1-4, Aug. 2017. DOI: 10.1109/CCECE.2017.7946722.
- [10] H. Lee, K. Hwang, M. Kang, and J. Song, "Black ice Detection Using CNN for the Prevention of Accidents in Automated Vehicle," *International Conference on Computational Science and Computational Intelligence (CSCI)*, pp. 1189-1192, Jun. 2020. DOI:10.3390/electronics9122178
- [11] Y. C. Du, X. M. Zhang, C. L. Liu and X. Liu, "Visibility Analysis for Freeway Based on Comparison of Ordinary and Infrared Images," *Journal of Transportation Systems Engineering and Information Technology*, Vol.16 No.4, pp. 73-78, Aug. 2016. DOI:1009-6744 (2016) 04-0073-06.
- [12] J. Long, E. Shelhamer and T. Darrell, "Fully convolutional networks for semantic segmentation," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3431-3440, Mar. 2015. DOI: 10.1109/CVPR.2015.7298965.

- [13] M. Everingham, S.M.A. Eslami, L. Van Gool, et al. "The PASCAL Visual Object Classes Challenge: A Retrospective," *International Journal of Computer Vision*, Volume 111, pp. 98-136, Jan. 2015. DOI: 10.1007/s11263-014-0733-5.
- [14] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *International Conference on Medical Image Computing and Computer-Assisted Intervention(MICCAI)*, pp. 234-241, May. 2015. DOI: 10.1007/978-3-319-24574-4_28.
- [15] H. Zhao, J. Shi, X. Qi, X. Wang and J. Jia, "Pyramid Scene Parsing Network," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 6230-6239, Apr. 2017. DOI: 10.1109/CVPR.2017.660.
- [15] L. -C. Chen, G. Papandreou, I. Kokkinos, K. Murphy and A. L. Yuille, "DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 40, No. 4, pp. 834-848, Apr. 2018. DOI: 10.1109/TPAMI.2017.2699184.
- [16] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770-778, Dec. 2015. DOI: 10.1109/CVPR.2015.90.
- [17] L. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation," *European Conference on Computer Vision(ECCV)*, pp. 883-851, Feb. 2018. DOI: 10.1007/978-3-030-01234-2_49.
- [18] F. Chollet, "Xception: Deep Learning with Depthwise Separable Convolutions," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1800-1807. Apr. 2017. arXiv:1610.02357v3
- [19] F. Yu, and V. Koltun, "Multi-Scale Context Aggregation by Dilated Convolutions," *ICLR*, Nov. 2015. arXiv:1511.07122.
- [20] A. Paszke, A. Chaurasia, S. Kim, and E. Culurciello, "ENet: A Deep Neural Network Architecture for Real-Time Semantic Segmentation," Jun. 2016. arXiv:1606.02147
- [21] A. Chaurasia and E. Culurciello, "LinkNet: Exploiting encoder representations for efficient semantic segmentation," *IEEE Visual Communications and Image Processing*, pp. 1-4, Dec. 2017. DOI: 10.1109/VCIP.2017.8305148
- [22] Y. J. Li, and S. K. Kang. "A Black Ice Recognition in Infrared Road Images Using Improved Lightweight Model Based on MobileNetV2," *Journal of the Korea Institute of Information and Communication Engineering*, Vol 25, No. 12, pp. 1835-1845. Dec. 2021. DOI: 10.6109/jkiice.2021.25.12.1835

Authors



Yu-Jie Li received the B.S. degree in Information Engineering from Ludong University, China, in 2008 and received the B.S. and M.S. degrees in Computer Engineering from Wonkwang University,

Korea, in 2008 and 2010, respectively. Yu-Jie Li is currently pursuing the Ph.D. degree in Computer Engineering from Wonkwang University, Korea. She is currently a Lecturer in the School of Computer Science, Weifang University of Science and Technology, China. She is interested in image processing, deep learning and pattern recognition.



Sun-Kyoung Kang received the B.S., M.S. and Ph.D. degrees in Computer Engineering from Wonkwang University, Korea, in 2000, 2004 and 2010, respectively. Dr. Kang is currently a Professor in the Department of

Computer and Software Engineering, Wonkwang University. She is interested in image processing, deep learning and bigdata.



Sung-Tae Jung received the B.S., M.S. and Ph.D. degrees in Computer Engineering from Seoul National University, Korea, in 1987, 1989 and 1994, respectively. Dr. Jung is currently a Professor in the Department of

Computer and Software Engineering, Wonkwang University. He is interested in computer vision, artificial intelligence, image processing and computer graphics.