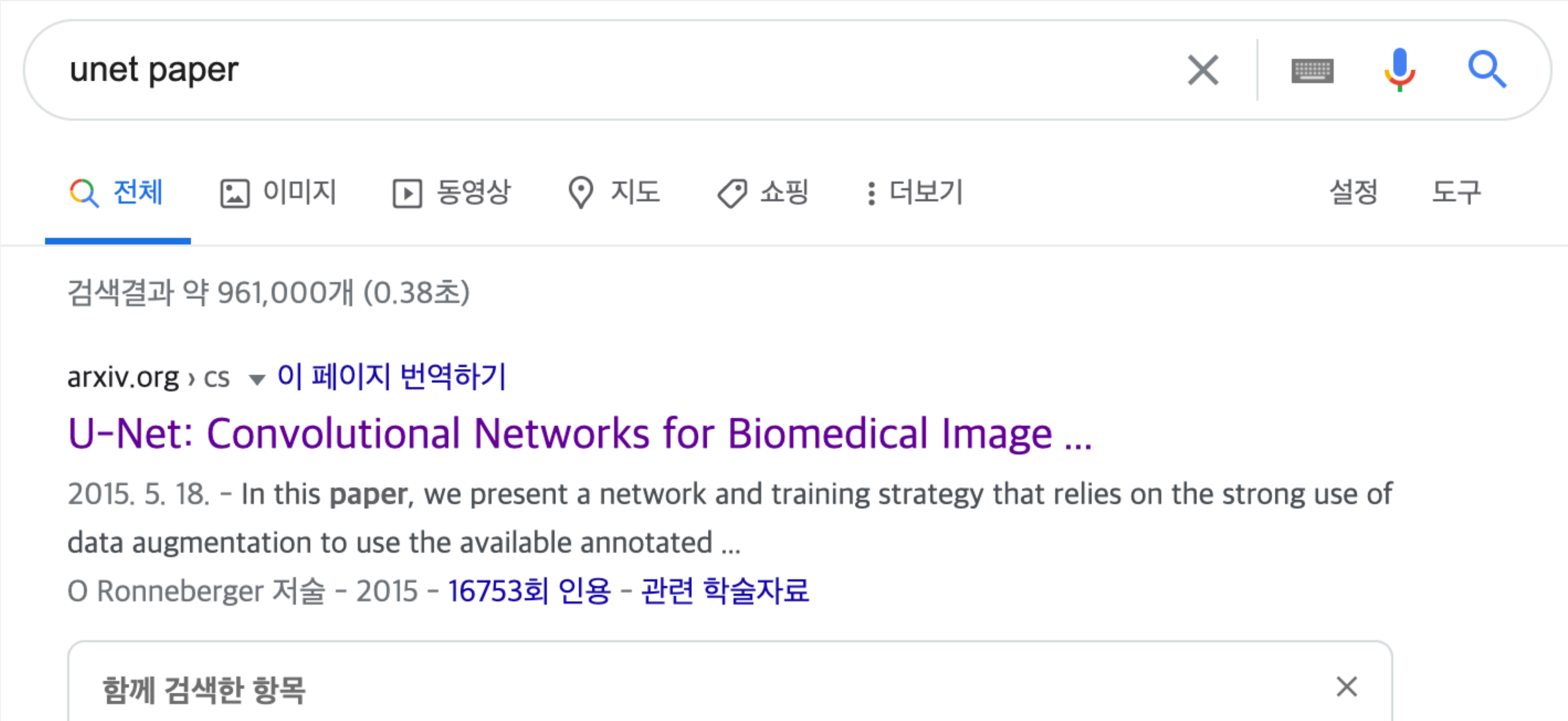
7~8차시. Mobile Unet

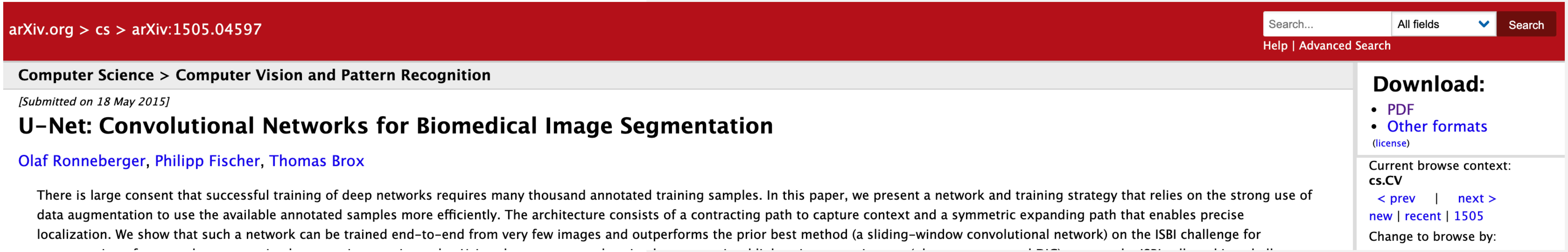
MobileUnet

논문보기

구글 검색

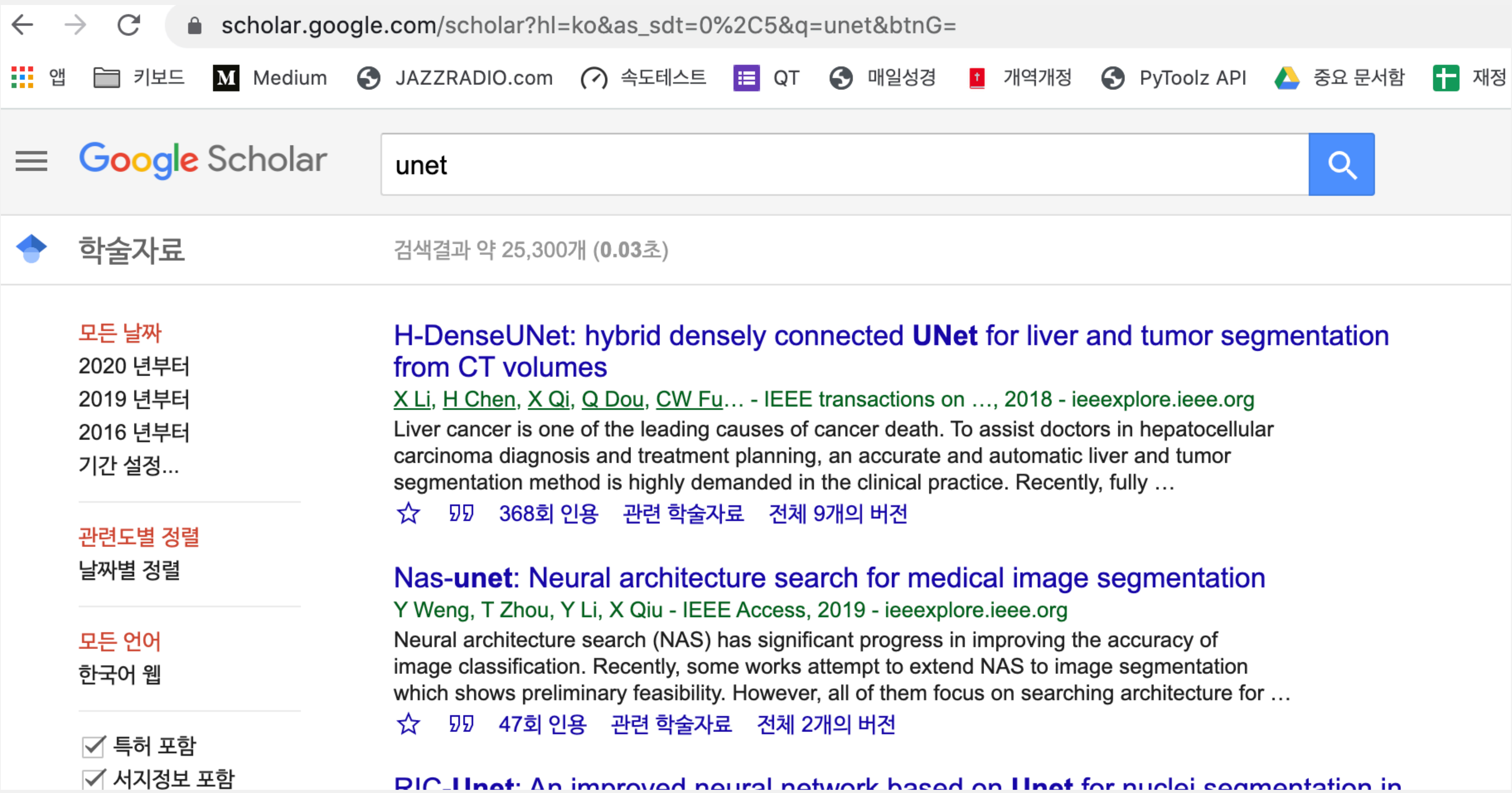
- arXiv.org





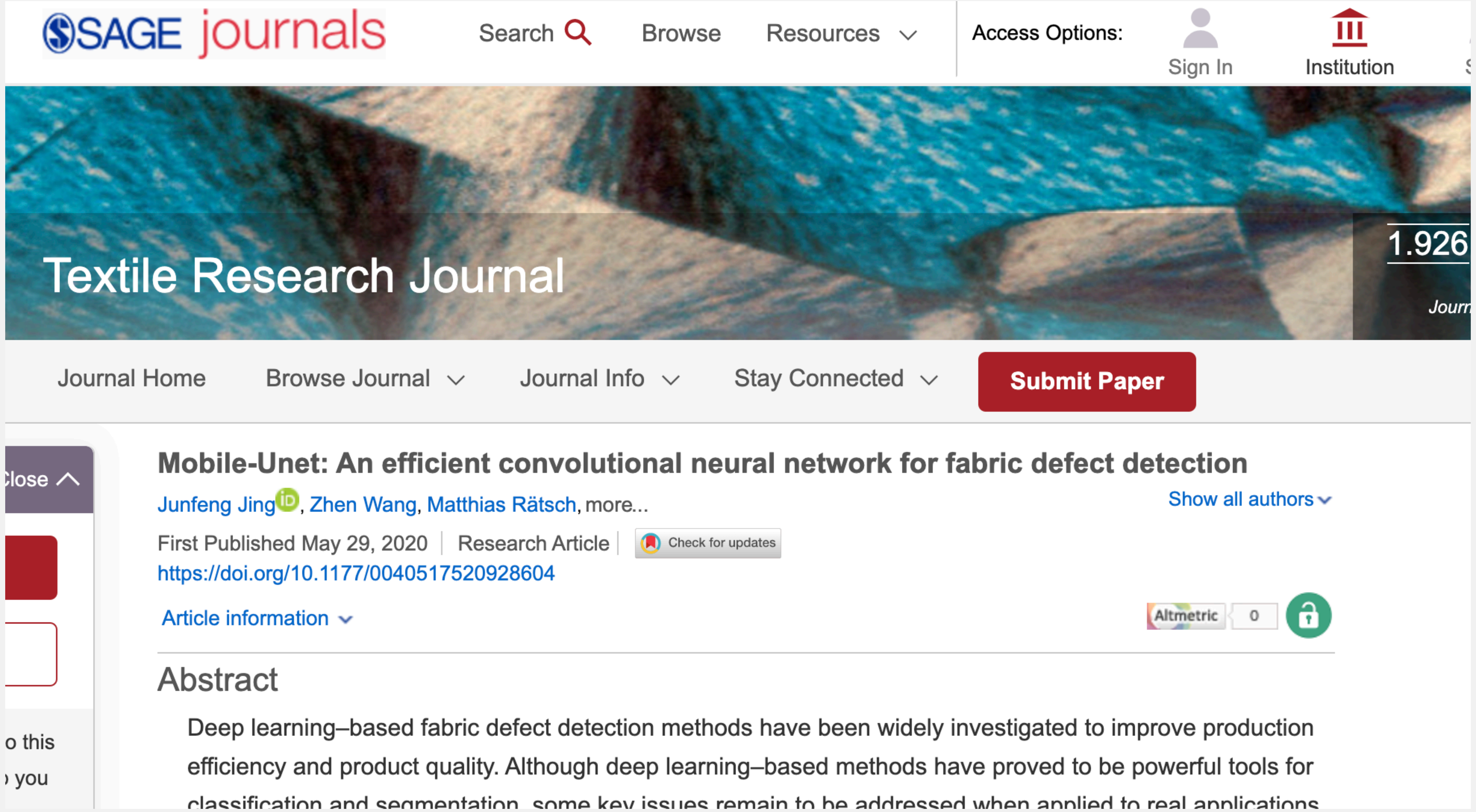
구글 스칼라 검색

- scholar.google.com
- 이 두 가지 방법으로 대부분 무료로 볼 수 있다.



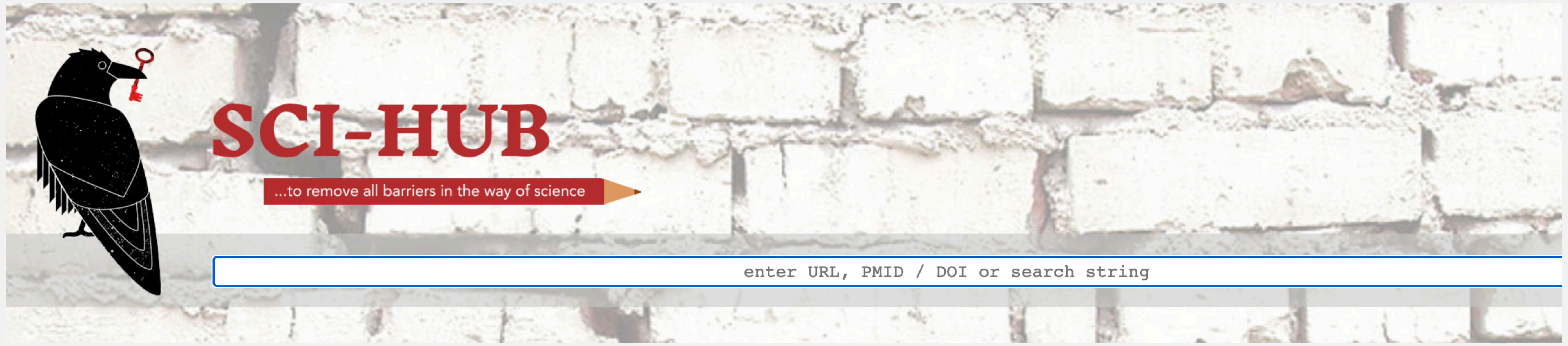
Mobile-Unet

- 두 가지 방법으로 다 무료로 볼 수 없다. (올라와 있지 않음)
- 대학생이라면, 학교 도서관과의 제휴로, 학교 IP에 한해, 논문을 다운 받을 수는 있다.
- 일반인이라면, 이 방법은 불가능



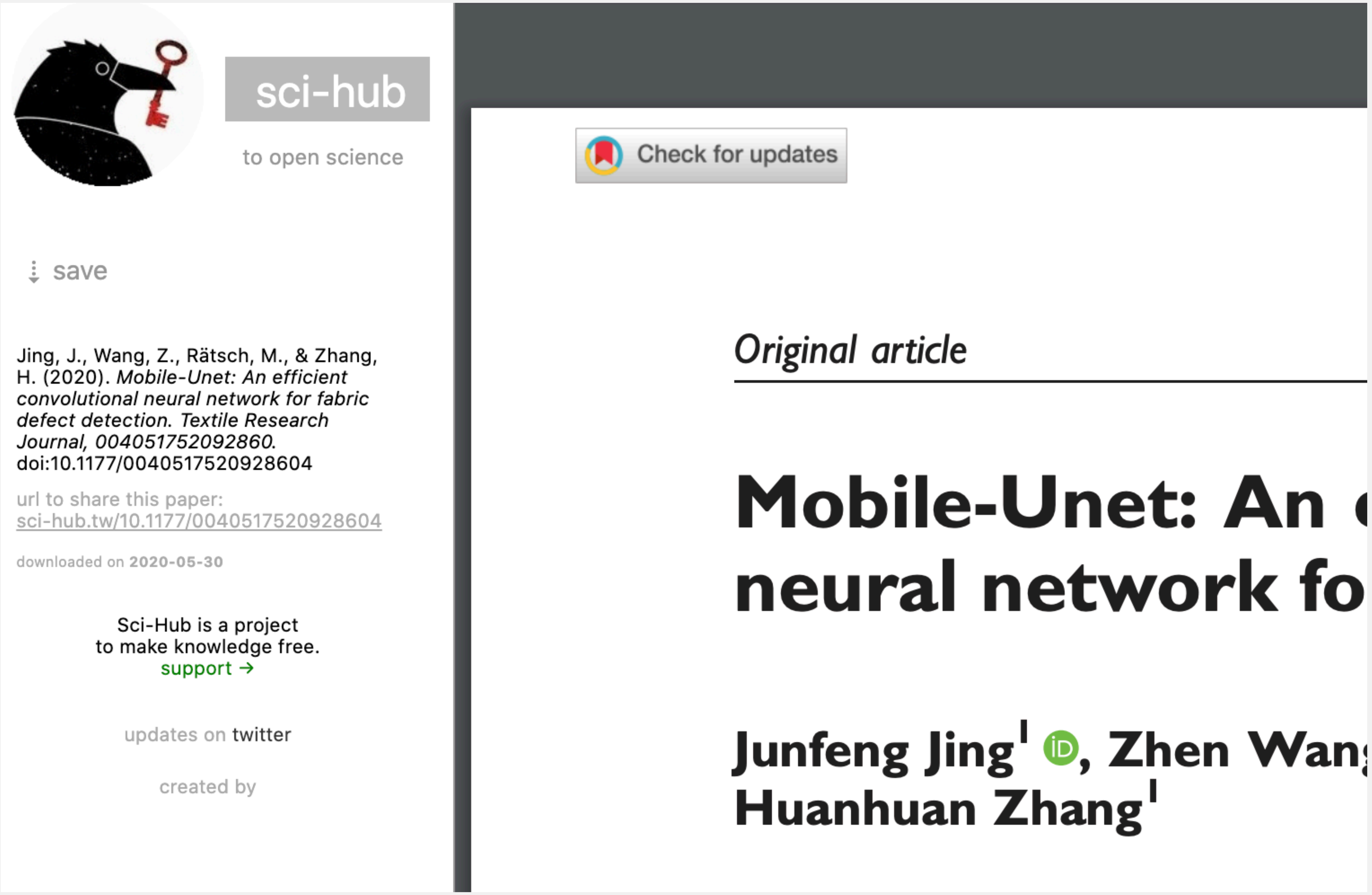
SCI-HUB

- https://sci-hub.tw/
- 10.1177/0040517520928604



SCI-HUB

- DOI 번호로 검색 가능
- https://ko.wikipedia.org/wiki/%EC%82%AC%EC%9D%B4%ED%97%88%EB%B8%8C
  - 관심있다면, 위키를 읽어보길..



Mobile-Unet

02

Abstract

- Deep learning-based fabric defect detection methods
- 해결하고자 하는 문제
    - 더 높은 실시간 성능이 필요
    - 비정상 샘플의 개수가 정상 샘플 개수에 비해 아주 적어서, 트레이닝 할 데이터 자체가 적다.

Abstract

Deep learning–based fabric defect detection methods have been widely investigated to improve production efficiency and product quality. Although deep learning–based methods have proved to be powerful tools for classification and segmentation, some key issues remain to be addressed when applied to real applications. Firstly, the actual fabric production conditions of factories necessitate higher real-time performance of methods. Moreover, fabric defects as abnormal samples are very rare compared with normal samples, which results in data imbalance. It makes model training based on deep learning challenging. To solve these problems, an extremely efficient convolutional neural network, Mobile-Unet, is proposed to achieve the end-to-end defect segmentation. The median frequency balancing loss function is used to overcome the challenge of sample imbalance. Additionally, Mobile-Unet introduces depth-wise separable convolution, which dramatically reduces the complexity cost and model size of the network. It comprises two parts: encoder and decoder. The MobileNetV2 feature extractor is used as the encoder, and then five deconvolution layers are added as the decoder. Finally, the softmax layer is used to generate the segmentation mask. The performance of the proposed model has been evaluated by public fabric datasets and self-built fabric datasets. In comparison with other methods, the experimental results demonstrate that segmentation accuracy and detection speed in the proposed method achieve state-of-the-art performance.

Abstract

- 볼링공을 생산하는 업체
  - 볼링공을 생산할 때, 검수 과정을 딥러닝 기반으로 자동화하고자 할 때, 데이터가 필요
  - 현재 불량률이 1% 정도
  - 10000개의 이미지를 얻었을 때, 그 중 100개 만이 불량인 이미지

Abstract

Deep learning–based fabric defect detection methods have been widely investigated to improve production efficiency and product quality. Although deep learning–based methods have proved to be powerful tools for classification and segmentation, some key issues remain to be addressed when applied to real applications. Firstly, the actual fabric production conditions of factories necessitate higher real-time performance of methods. Moreover, fabric defects as abnormal samples are very rare compared with normal samples, which results in data imbalance. It makes model training based on deep learning challenging. To solve these problems, an extremely efficient convolutional neural network, Mobile-Unet, is proposed to achieve the end-to-end defect segmentation. The median frequency balancing loss function is used to overcome the challenge of sample imbalance. Additionally, Mobile-Unet introduces depth-wise separable convolution, which dramatically reduces the complexity cost and model size of the network. It comprises two parts: encoder and decoder. The MobileNetV2 feature extractor is used as the encoder, and then five deconvolution layers are added as the decoder. Finally, the softmax layer is used to generate the segmentation mask. The performance of the proposed model has been evaluated by public fabric datasets and self-built fabric datasets. In comparison with other methods, the experimental results demonstrate that segmentation accuracy and detection speed in the proposed method achieve state-of-the-art performance.

Abstract

- Mobile-Unet은 end-to-end defect segmentation을 달성하기 위한 극도로 효율적인 CNN
  - 1의 문제를 해결하기 위해, depth-wise separable convolution 도입
  - 2의 문제를 해결하기 위해, median frequency balancing loss 도입
  - 둘 다 기존에 있는 것들

Abstract

Deep learning–based fabric defect detection methods have been widely investigated to improve production efficiency and product quality. Although deep learning–based methods have proved to be powerful tools for classification and segmentation, some key issues remain to be addressed when applied to real applications. Firstly, the actual fabric production conditions of factories necessitate higher real-time performance of methods. Moreover, fabric defects as abnormal samples are very rare compared with normal samples, which results in data imbalance. It makes model training based on deep learning challenging. To solve these problems, an extremely efficient convolutional neural network, Mobile-Unet, is proposed to achieve the end-to-end defect segmentation. The median frequency balancing loss function is used to overcome the challenge of sample imbalance. Additionally, Mobile-Unet introduces depth-wise separable convolution, which dramatically reduces the complexity cost and model size of the network. It comprises two parts: encoder and decoder. The MobileNetV2 feature extractor is used as the encoder, and then five deconvolution layers are added as the decoder. Finally, the softmax layer is used to generate the segmentation mask. The performance of the proposed model has been evaluated by public fabric datasets and self-built fabric datasets. In comparison with other methods, the experimental results demonstrate that segmentation accuracy and detection speed in the proposed method achieve state-of-the-art performance.

Abstract

- Encoder, Decoder 구조의 네트워크
    - 인코더에 MobileNetV2 feature extractor 도입
        - 이게 전이학습을 사용하는 요소.
    - 디코더에 다섯개의 컨볼루션 네트워크 도입
    - 맨 마지막에 softmax 레이어 도입

Abstract

Deep learning–based fabric defect detection methods have been widely investigated to improve production efficiency and product quality. Although deep learning–based methods have proved to be powerful tools for classification and segmentation, some key issues remain to be addressed when applied to real applications. Firstly, the actual fabric production conditions of factories necessitate higher real-time performance of methods. Moreover, fabric defects as abnormal samples are very rare compared with normal samples, which results in data imbalance. It makes model training based on deep learning challenging. To solve these problems, an extremely efficient convolutional neural network, Mobile-Unet, is proposed to achieve the end-to-end defect segmentation. The median frequency balancing loss function is used to overcome the challenge of sample imbalance. Additionally, Mobile-Unet introduces depth-wise separable convolution, which dramatically reduces the complexity cost and model size of the network. It comprises two parts: encoder and decoder. The MobileNetV2 feature extractor is used as the encoder, and then five deconvolution layers are added as the decoder. Finally, the softmax layer is used to generate the segmentation mask. The performance of the proposed model has been evaluated by public fabric datasets and self-built fabric datasets. In comparison with other methods, the experimental results demonstrate that segmentation accuracy and detection speed in the proposed method achieve state-of-the-art performance.

Introduction

- 기존에 Fabric 결함 감지를 찾아냈던 방법들
- 딥러닝 기반 결함 감지 방법
  - CNN 기반 방법 여러가지의 소개
- U-Net의 초반부를 MobileNetV2로 교체하겠다.
  - real-time application 및 better feature extraction 간의 균형 맞추기

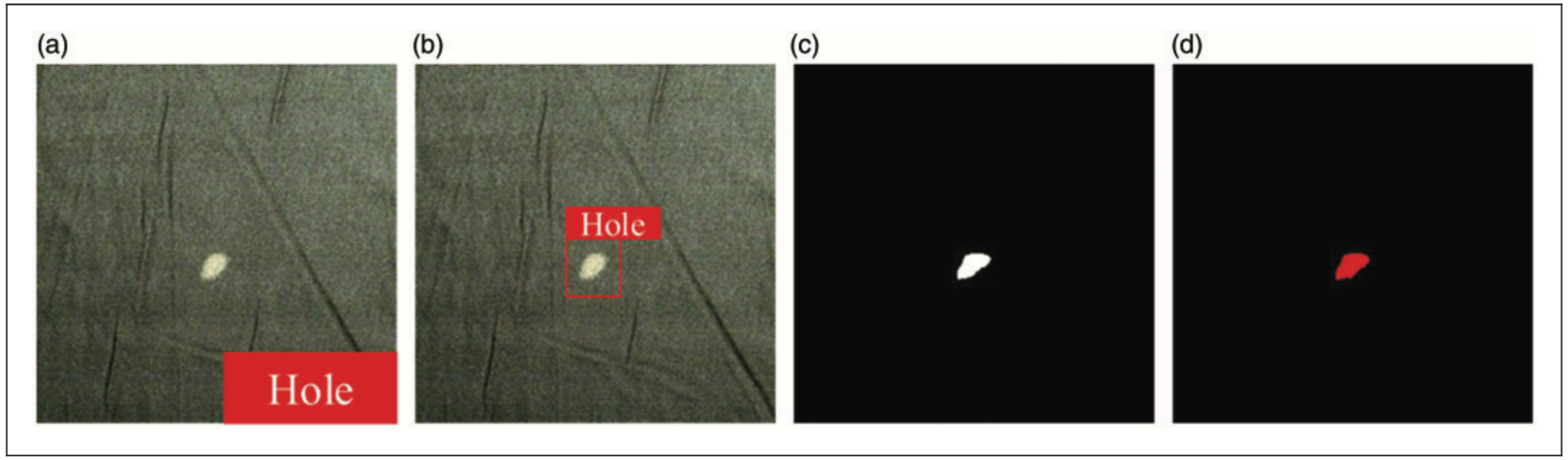


Figure 1. Four levels of defect detection: (a) defect classification; (b) defect location; (c) defect segmentation; (d) defect semantic segmentation.

Introduction

- Contributions
  - **end-to-end** 네트워크
  - 더 적은 파라미터. 온라인. (속도 및 실시간)
  - 샘플 간의 불균형 해소방안. (불균형 해소 방안)

1. Different from traditional algorithms for defect classification or defect location, Mobile-Unet is a novel end-to-end defect segmentation network that achieves pixel-level fabric defect classification.
2. Compared with extant methods, Mobile-Unet has fewer model parameters and can significantly shorten detection time. It is also more suitable for online automated detection.
3. In order to solve the imbalance between defective and non-defective samples, a cross-entropy loss function weighted with median frequency balancing is used to improve the convergence speed and detection accuracy of the model.

Proposed Model

- 사전 트레이닝 단계
    - MoibleNetV2
        - 결함 분류기
- 시맨틱 세그먼테이션 단계
    - 첫 다섯 컨볼루션을 인코더로, 디코더로서, 다섯 개의 디콘볼루션 도입

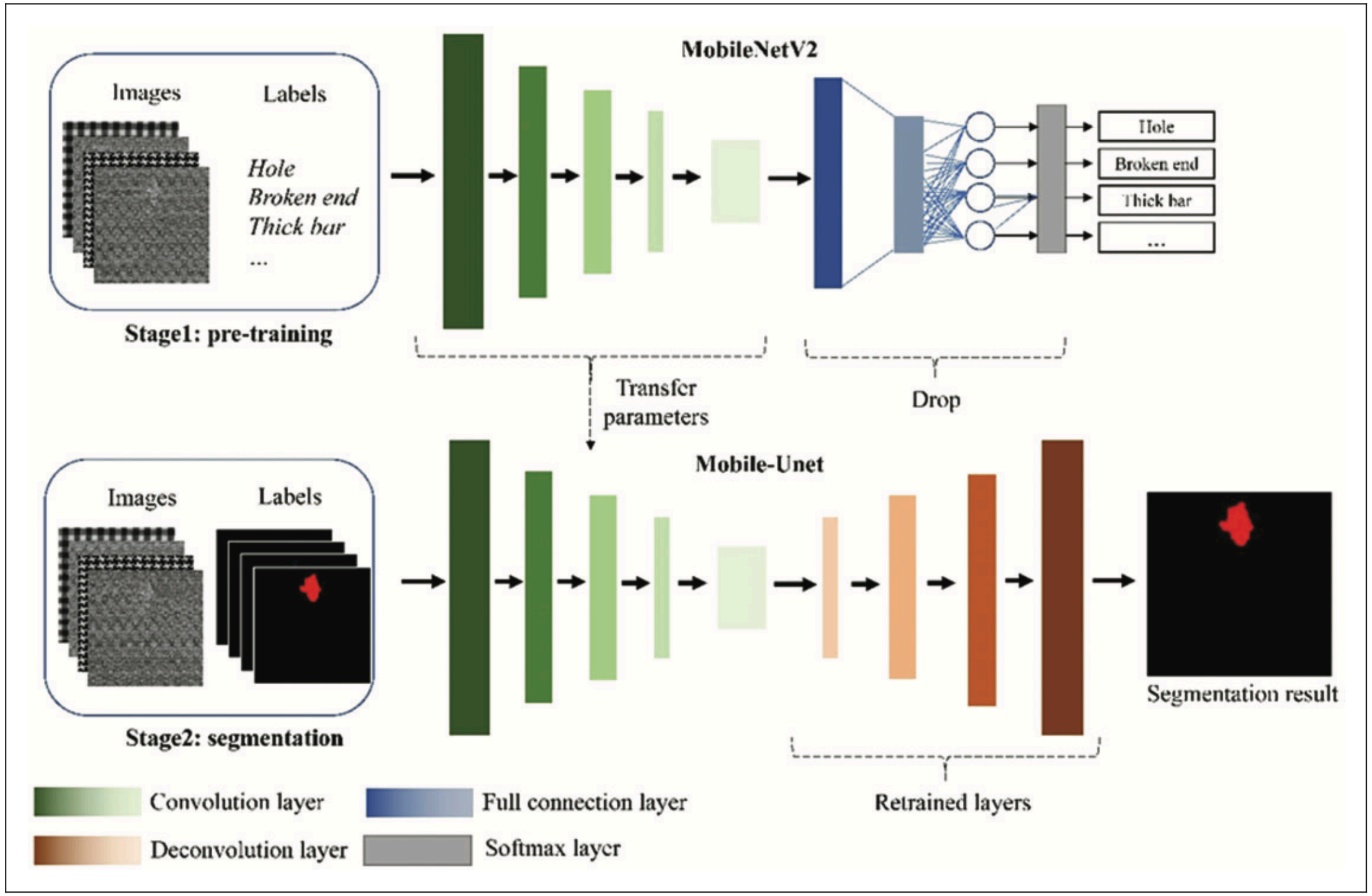


Figure 2. Process of our method.

relationship between the two stages is illustrated in Figure 2. The pre-training stage (stage 1 in Figure 2) comprises training MobileNetV2 for defect classification, and the input of the network is the defect image

Proposed Model
사전 트레이닝 단계

- MobileNetV2
  - 경량 CNN을 위한 depthwise. separable convolutions 사용
  - depthwise separable convolutions.
  - real-time tasks에서 널리 도입됨
    - 이유
      - 더 적은 파라미터
      - 계산량이 적다.

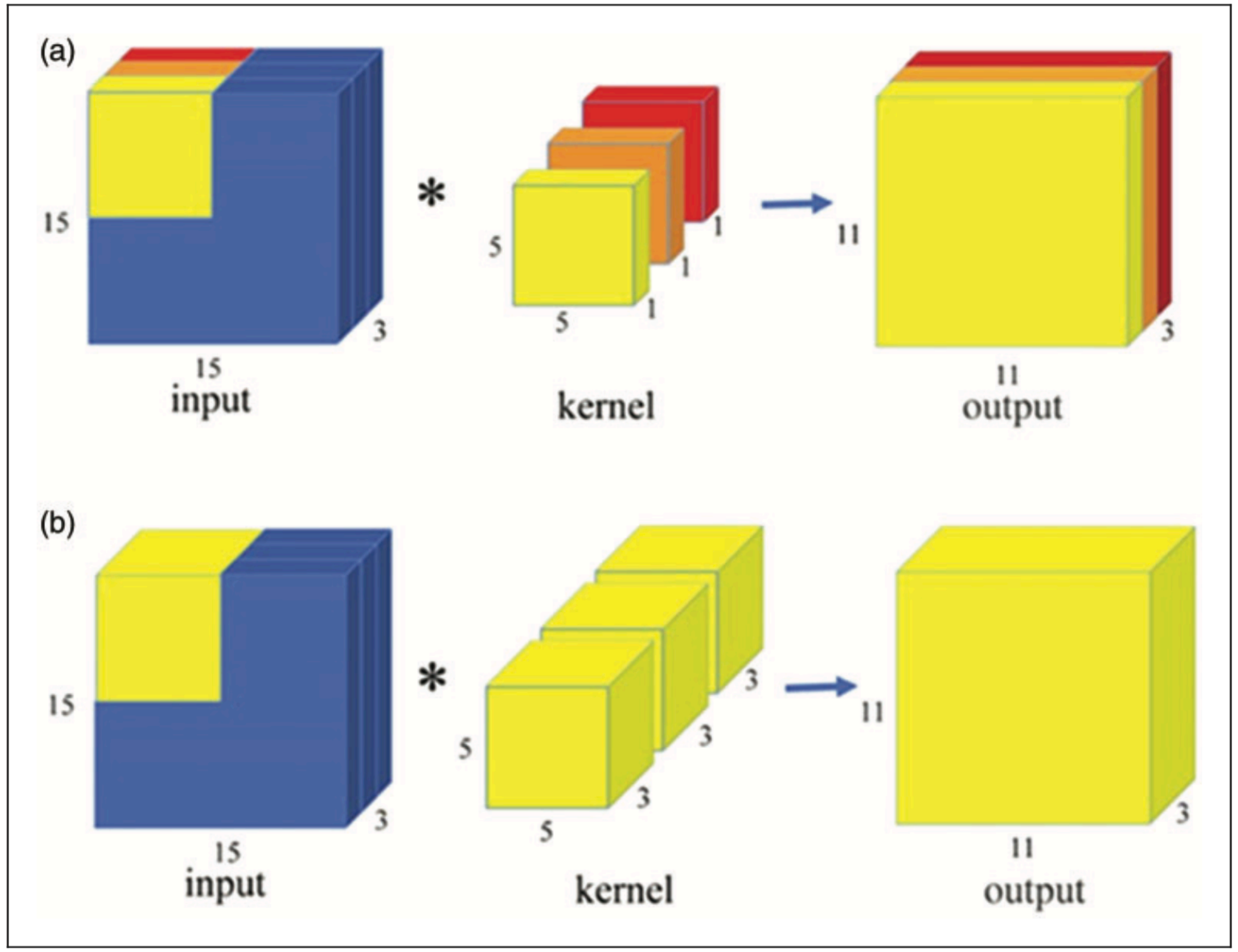


Figure 3. Two types of convolution using three kernels to transform a 15 × 15 × 3 image into an 11 × 11 × 3 image: (a) depth-wise convolution; (b) standard convolution.

Proposed Model
사전 트레이닝 단계

- inverted residual blocks.
    - 하나는 stride가 1인 residual 블록.
    - 다른 하나는 stride가 2인 다운사이징을 위한 블록.
    - 두 타입의 블록 모두는 세 레이어로 구성
        - 1x1 컨볼루션 레이어. ReLU 6.
        - 3x3 깊이별 컨볼루션 레이어
        - 1x1 컨볼루션 레이어. 활성화 함수 없음

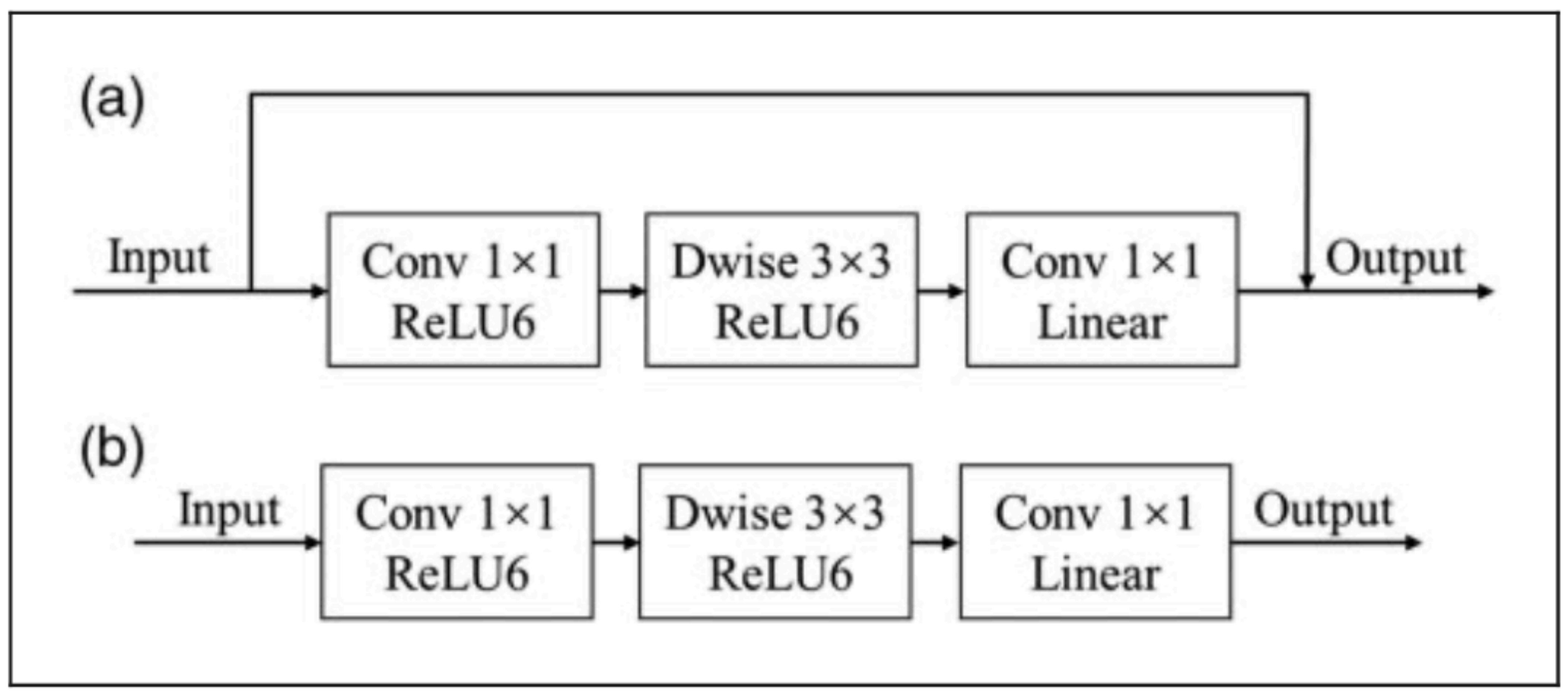
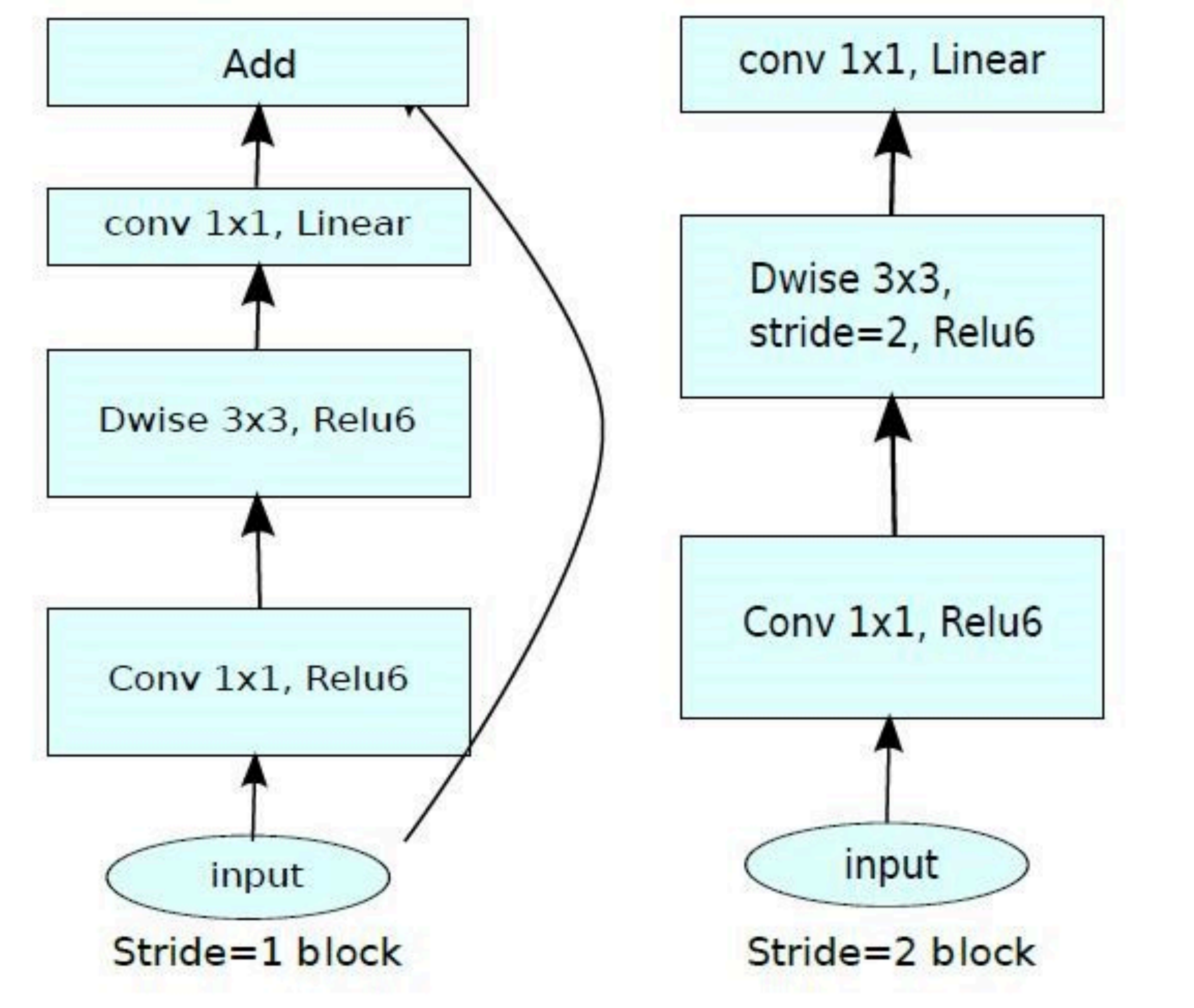


Figure 4. Inverted residual blocks: (a) stride = 2 block; (b) stride = 1 block.



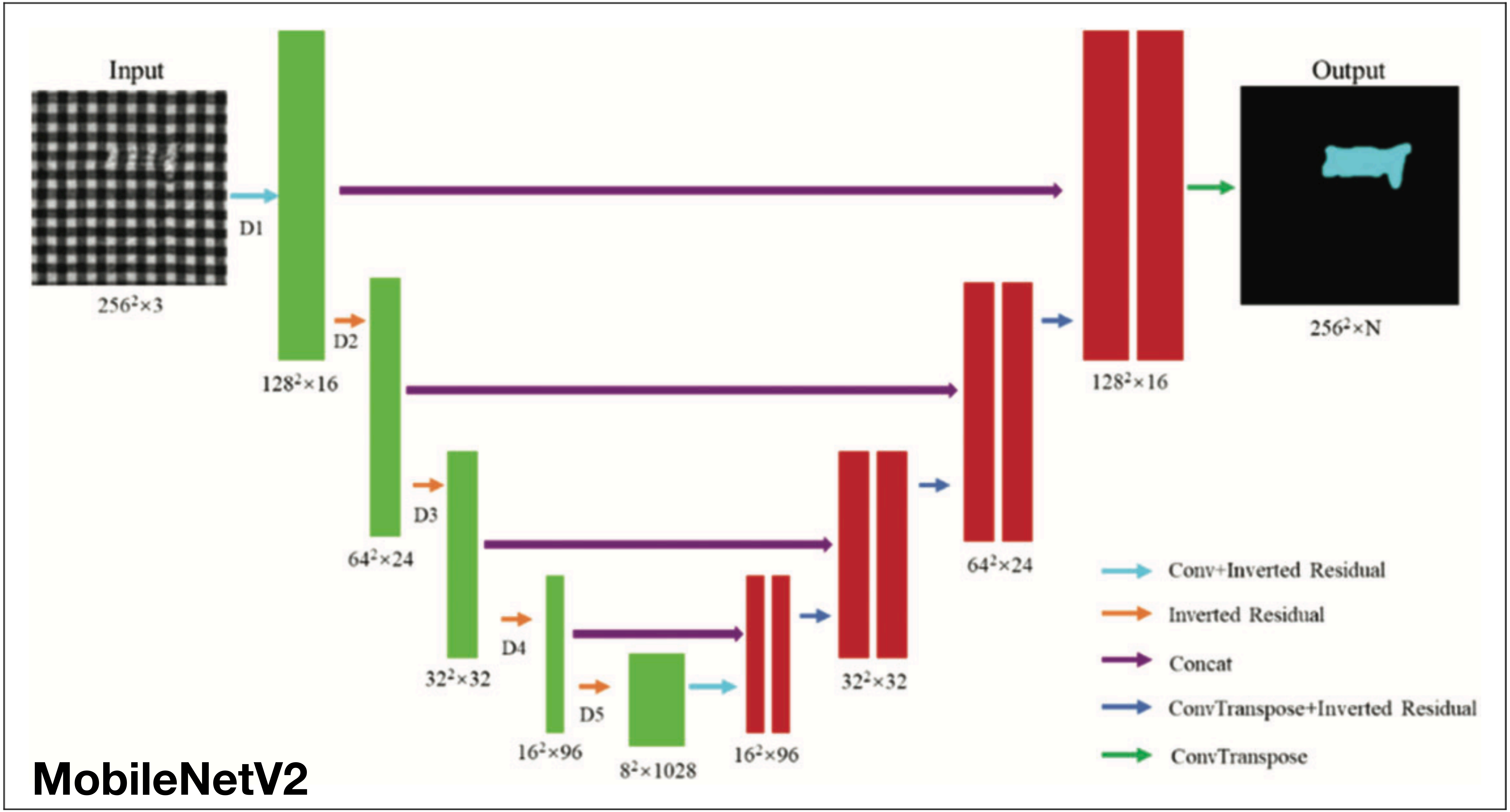


Figure 5. Architecture of Mobile-Unet.

Table 1. The structural configuration of MobileNetV2

Layers	Input size	Operation	t	k	n	s	Output size
Input image	–	–	–	–	–	–	(3,256,256)
D1	(3,256,256)	Conv + BN + ReLU6	–	3	1	–	(16,128,128)
		Inverted residual	1	–	1	1	
D2	(16,128,128)	Inverted residual	6	–	2	2	(24,64,64)
D3	(24,64,64)	Inverted residual	6	–	3	2	(32,32,32)
D4	(32,32,32)	Inverted residual	6	–	4	2	(96,16,16)
		Inverted residual	6	–	3	1	
D5	(96,16,16)	Inverted residual	6	–	3	2	(128,8,8)
		Inverted residual	6	–	1	1	
		Inverted residual	6	3	1	–	
C1	1280	Linear	–	–	–	–	N

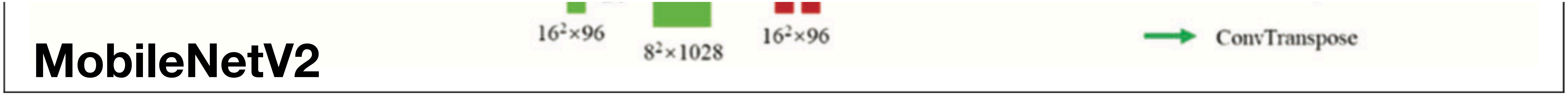


Figure 5. Architecture of Mobile-Unet.

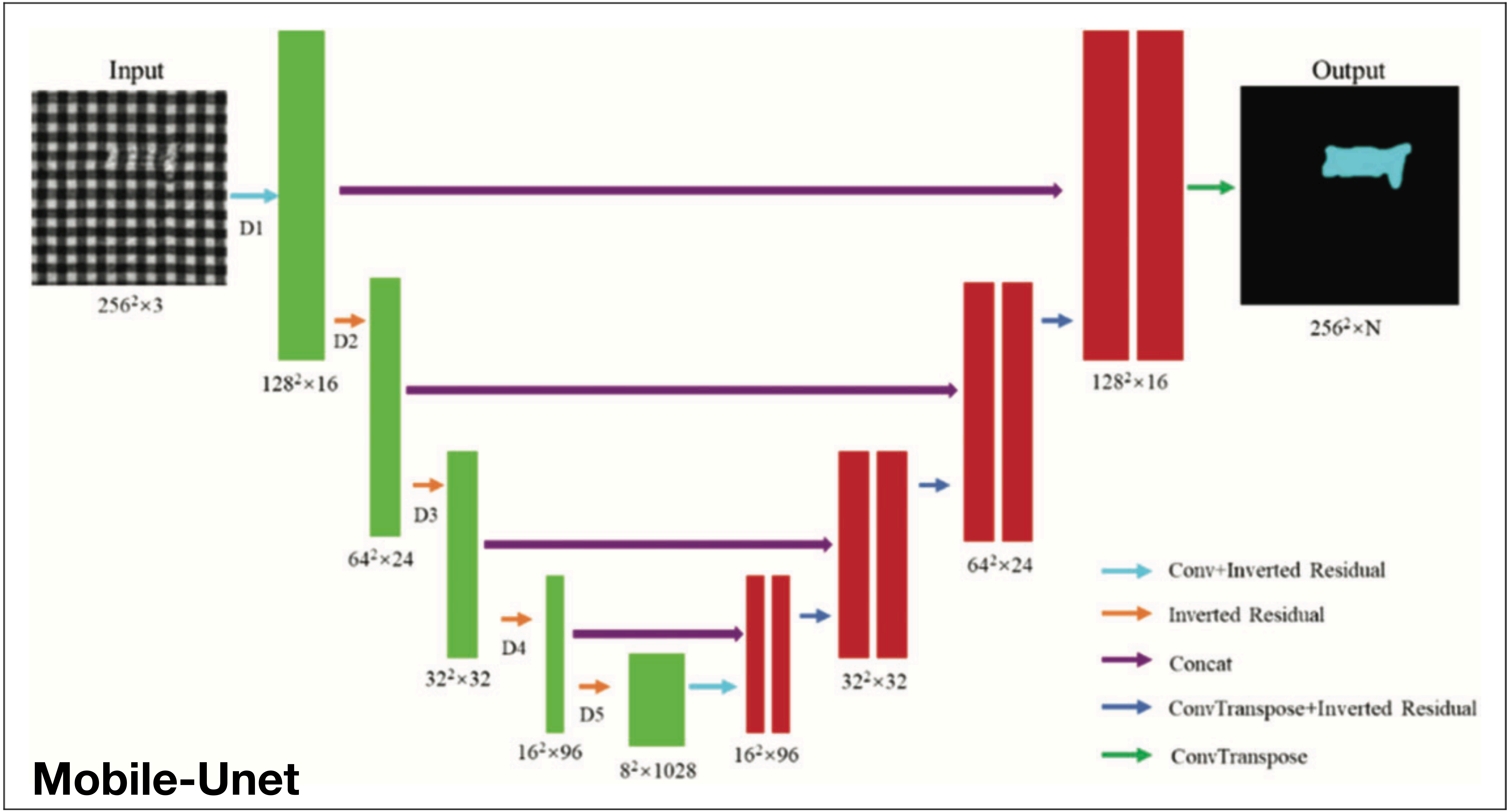


Figure 5. Architecture of Mobile-Unet.

Table 2. Definition and operation of each layer of Mobile-Unet

Layer	Input	Output	Output size	k	s	t
Input image	–	–	(3,256,256)	–	–	–
D1	Input image	X1	(16,128,128)	–	–	–
D2	X1	X2	(24,64,64)	–	–	–
D3	X2	X3	(32,32,32)	–	–	–
D4	X3	X4	(96,16,16)	–	–	–
D5	X4	X5	(1280,8,8)	–	–	–
ConvTranspose1	X5	L1	(96,16,16)	4	–	2
Inverted Residual1	L1 + X4	L2	(96,16,16)	–	6	1
ConvTranspose2	L2	L3	(32,32,32)	4	–	2
Inverted Residual2	L3 + X3	L4	(32,32,32)	–	6	1
ConvTranspose3	L4	L5	(24,64,64)	4	–	2
Inverted Residual3	L5 + X2	L6	(24,64,64)	–	6	1
ConvTranspose4	L6	L7	(16,128,128)	4	–	2
Inverted Residual4	L7 + X1	L8	(16,128,128)	–	6	1
ConvTranspose5 + SoftMax	L8	Output	(N,256,256)	4	–	2

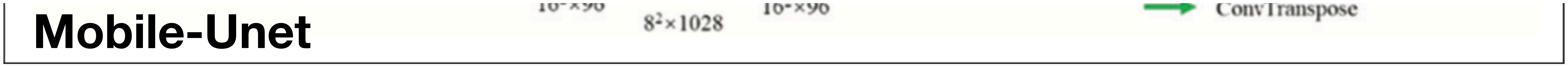


Figure 5. Architecture of Mobile-Unet.

Proposed Model

개선된 손실 함수

- 결함 비율이 아주 낮은 문제를 해결하자.
- $Loss = -\frac{1}{N}\sum_{n=1}^{N}\sum_{c=1}^{C} w_c l_c^{(n)} \log(p_c^{(n)})$
- $w_c = \frac{median(\{f_c \mid c \in C\})}{f_c}$
  - N : 트레이닝 샘플의 픽셀 수
  - w_c : 클래스 c의 가중치
  - f_c : 클래스 c에 있는 픽셀의 frequency
  - $p_c^{(n)}$: 클래스 c에 있는 픽셀 n의 softmax 확률
  - $l_c^{(n)}$: 클래스 c에 대한 샘플 n의 해당되는 라벨 (one-hot 인코딩)
  - C : 모든 클래스의 세트

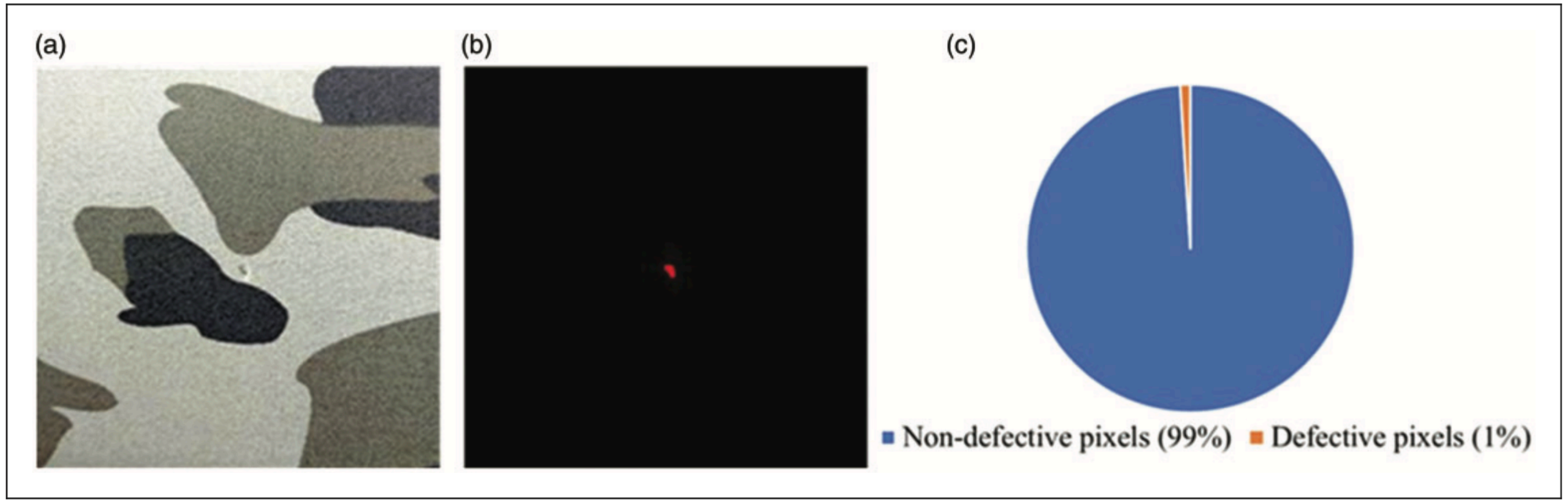


Figure 6. Imbalanced number of defects and background pixels: (a) original images; (b) ground truth; (c) comparison of the number of defective and non-defective pixels.

ground pixels. An ordinary loss function may cause the network to fall into local optimum. In order to solve the sample imbalance problem, an improved cross-entropy loss with median frequency balancing weights (CE-MFB)[34] is used. The class loss is weighted by the ratio of the median class frequency in the training set and the actual class frequency. The improved cross-

Proposed Model

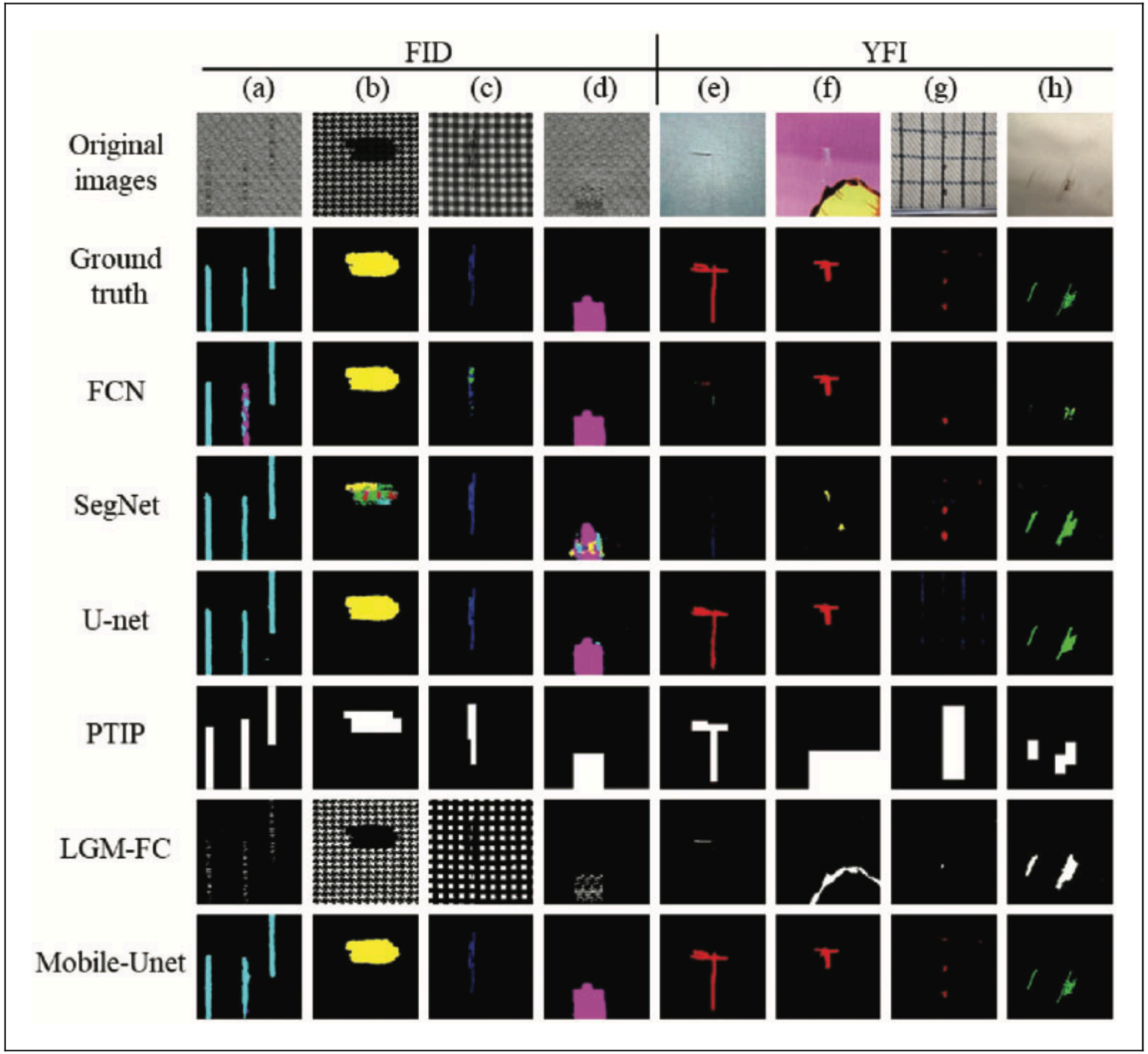
개선된 손실 함수

- 결함 비율이 아주 낮은 문제를 해결하자.
- $Loss = -\frac{1}{N}\sum_{n=1}^{N}\sum_{c=1}^{C} w_c l_c^{(n)} \log(p_c^{(n)})$
- $w_c = \frac{median(\{f_c \mid c \in C\})}{f_c}$
    - N : 트레이닝 샘플의 픽셀 수
    - w_c : 클래스 c의 가중치
    - f_c : 클래스 c에 있는 픽셀의 frequency
    - $p_c^{(n)}$: 클래스 c에 있는 픽셀 n의 softmax 확률
    - $l_c^{(n)}$: 클래스 c에 대한 샘플 n의 해당되는 라벨 (one-hot 인코딩)
    - C : 모든 클래스의 세트

```
def _compute_cross_entropy_mean(class_weights, labels, softmax):
    cross_entropy = -tf.reduce_sum(
        tf.multiply(labels * tf.log(softmax), class_weights),
        reduction_indices=[1])

    cross_entropy_mean = tf.reduce_mean(
        cross_entropy, name='xentropy_mean')
    return cross_entropy_mean
```

shows the imbalance in the number of defects and back-ground pixels. An ordinary loss function may cause the network to fall into local optimum. In order to solve the sample imbalance problem, an improved cross-entropy loss with median frequency balancing weights (CE-MFB)[34] is used. The class loss is weighted by the ratio of the median class frequency in the training set and the actual class frequency. The improved cross-



FID
YFI
(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
Original images
Ground truth
FCN
SegNet
U-net
PTIP
LGM-FC
Mobile-Unet

결과

실습 Session

4. Mobile-Unet.ipynb

Mobile-Unet

딥러닝을 도와주는 보조도구

03

https://deepcognition.ai/

딥러닝 GUI

https://github.com/lutzroeder/netron

딥러닝 GUI 분석기