

GLOW 기반 에지 탐지 기법

전 형 주¹ · 정 순 철¹ · 최 윤 석¹ · 김 재 우¹ · 김 진 서^{2*}
^{1,2*}한국전자통신연구원

GLOW-based edge detection method

Hyeong-Ju Jeon¹ · Soonchul Jung¹ · Yoon-Seok Choi¹ · Jae Woo Kim¹ · Jin Seo Kim^{2*}

^{1,2*}Electronics and Telecommunications Research Institute

[요 약]

이미지 내에서 객체는 다양한 크기로 존재하므로, 딥러닝 기반의 에지 탐지에서는 계층적 표현자를 학습하는 것이 중요하다. 또한, 에지 탐지는 픽셀 단위의 세밀한 연산을 요구하므로, 에지 탐지를 위해 이미지 정보 손실이 최소화되어야 한다. 이를 위해서 기존 에지 탐지 기법에서는 인코더와 디코더가 밀접 연결된 구조로 구성된다. 그러나, 인코딩 단계의 컨볼루션 및 풀링 연산은 여전히 이미지 정보를 유실시키는 문제점이 존재한다. 이를 해결하기 위해 본 논문에서는 GLOW를 기반으로 한 에지 탐지 기법을 제안한다. GLOW는 역변환 가능한 함수들로 구성되어, 이미지 정보 손실을 줄여준다. 제안 기법을 평가하기 위해 BSDS500 데이터셋과 BIPED 데이터셋을 각각 학습하고 비교하여, 유의미한 에지를 탐지함을 확인하였다. 더 나아가, 제안 기법을 사진 이미지가 아닌 다른 도메인인 회화 작품 이미지에 적용했을 때 제안 기법이 기존 기법에 비해 에지를 세밀하게 탐지함을 확인하였다.

[Abstract]

Since objects in images exist in various sizes, it is important to learn hierarchical representation for edge detection. In addition, since edge detection requires detailed operation at the pixel level, content leak in an image should be minimized. To this end, the traditional edge detection methods consist of a structure in which an encoder and a decoder are densely connected. However, content leak in an image occurs through pooling and convolution of encoding. In this paper, we propose an edge detection method based on GLOW to solve the problem. As GLOW consists of reversible functions, reducing content leak in an image. In order to evaluate the proposed method, BSDS500 dataset and BIPED dataset were respectively trained and comparatively analyzed. The experimental results showed that the proposed method detected reasonable edges in images. Furthermore, it was confirmed that when the proposed method was applied to a painting work image, which is a domain other than a photographic image, the proposed method detected edges more precisely than the existing method.

색인어 : 딥러닝, 인코더, 디코더, GLOW, 에지 탐지, 계층적 표현자

Keyword : Deep learning, Encoder, Decoder, GLOW, Edge detection, Hierarchical representation

<http://dx.doi.org/10.9728/dcs.2021.22.12.2089>



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.

Received 12 November 2021; **Revised** 29 November 2021

Accepted 17 December 2021

***Corresponding Author, Jin Seo Kim**

Tel: +82-42-860-5552

E-mail: kjseo@etri.re.kr

I. 서론

에지 탐지는 그림 1처럼 이미지 내에서 중요한 정보를 포함하는 경계 부분을 추출하는 것을 말한다. 이미지 내 불필요한 정보를 최소화함으로써 영역 분할, 원격탐사, 의료 영상 등 다양한 컴퓨터 비전 분야에 성능을 향상시키는 보조적인 용도로 활용된다. 따라서, 최근에도 에지 탐지 관련 연구가 활발히 진행 중이다.

기존 에지 탐지 기법들은 Sobel, Laplacian 등과 같이 연산자를 이용하여 밝기값, 질감 등에서의 불연속성을 통해 에지를 추출한다[1], [2]. 더 나아가, Pb와 같이 수동 설계된 특징들을 이용하여 이미지 내 에지를 탐지한다[3]. 그러나 기존 에지 탐지 기법들은 낮은 수준 특징들을 이용하므로, 이미지에서 의미적으로 중요한 경계 부분만을 탐지하는 것은 어렵다.

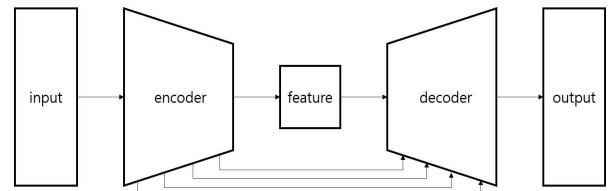
이에 높은 수준 표현자를 학습할 수 있는, RCF (richer convolutional features), HED (holistically-nested edge detection), DeepEdge, ContourGAN[4]-[7] 등과 같이 딥러닝을 활용한 여러 에지 탐지 연구들이 활발히 진행되고 있으며[15], 에지 탐지 성능 또한 급격히 향상되었다. 기존 딥러닝 기반 에지 탐지 기법들은 주로 인코더와 디코더를 결합한 형태의 구조로 구성된다. 이미지 내 객체들은 다양한 크기로 존재하므로, 에지 탐지를 위해서는 객체에 대한 다양한 크기에 대한 표현자를 학습하는 것이 중요하다. 이를 위해 기존 딥러닝 기반 에지 탐지 기법들은 그림 2(a)와 같이 인코더와 디코더를 밀접 연결한 구조를 통해 계층적 표현자를 학습하여, 다양한 크기의 객체의 에지를 추출한다. 그러나, 인코더와 디코더 결합 구조인 경우 인코더 과정에서 이미지 정보 손실이 발생하는 문제점이 존재한다[8].

본 논문에서는 이를 보완하기 위해 GLOW (generative flow) 기반 에지 탐지 기법을 제안한다. 제안 기법은 딥러닝 기반 생성 모델 중 하나인 NF (neural flow)를 기반으로 구성된다[9], [10]. 그림 2(b)는 제안 기법의 개념도를 보여준다. 이 기법은 역변환이 가능하여 순방향과 역방향 추론을 하나의 네트워크에서 수행한다. 이를 통해 이미지를 재구성할 때 정보 손실을 최소화한다[8]. 순방향을 통해서 이미지에서 특징을 추출하고, 역방향을 통해 추출된 특징을 이미지로 재구성한다. 이 기법의 네트워크는 기본적으로 additive coupling, invertible 1×1 convolution, actnorm으로 구성된다. 본 연구에서는 BSDS500 (berkeley segmentation dataset 500) 데이터셋과 BIPED (barcelona images for perceptual edge detection) 데이터셋을 이용하여 제안 기법을 학습 및 평가하였다.

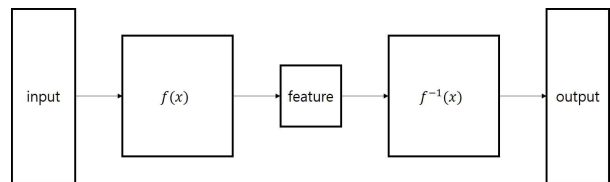
2절에서는 에지 탐지 관련 연구에 대해 서술하며, 3절에서는 제안 기법에 대해 상세히 설명하고, 4절에서는 제안 기법을 통한 실험 결과에 대한 분석을 논의하고자 한다. 마지막 절에서는 결과와 향후 연구 방향에 대해 기술한다.



그림 1. 이미지 내 에지 예시
Fig. 1. Examples of edge in images



(a) 기존 에지 탐지 기법 개념도
(a) The conceptual diagram of the traditional edge detection method



(b) 제안 에지 탐지 기법 개념도
(b) The conceptual diagram of the proposed edge detection method

그림 2. 기존 에지 탐지 개념도와 제안 에지 탐지 기법 개념도
Fig. 2. The Comparison of conceptual diagrams between the traditional and the proposed edge detections

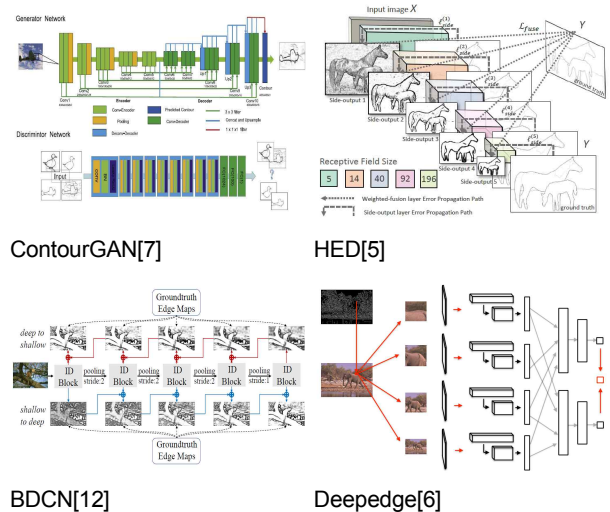


그림 3. 기존 딥러닝 기반 에지 탐지 구조
Fig. 3. The structure of the traditional deep learning-based edge detections

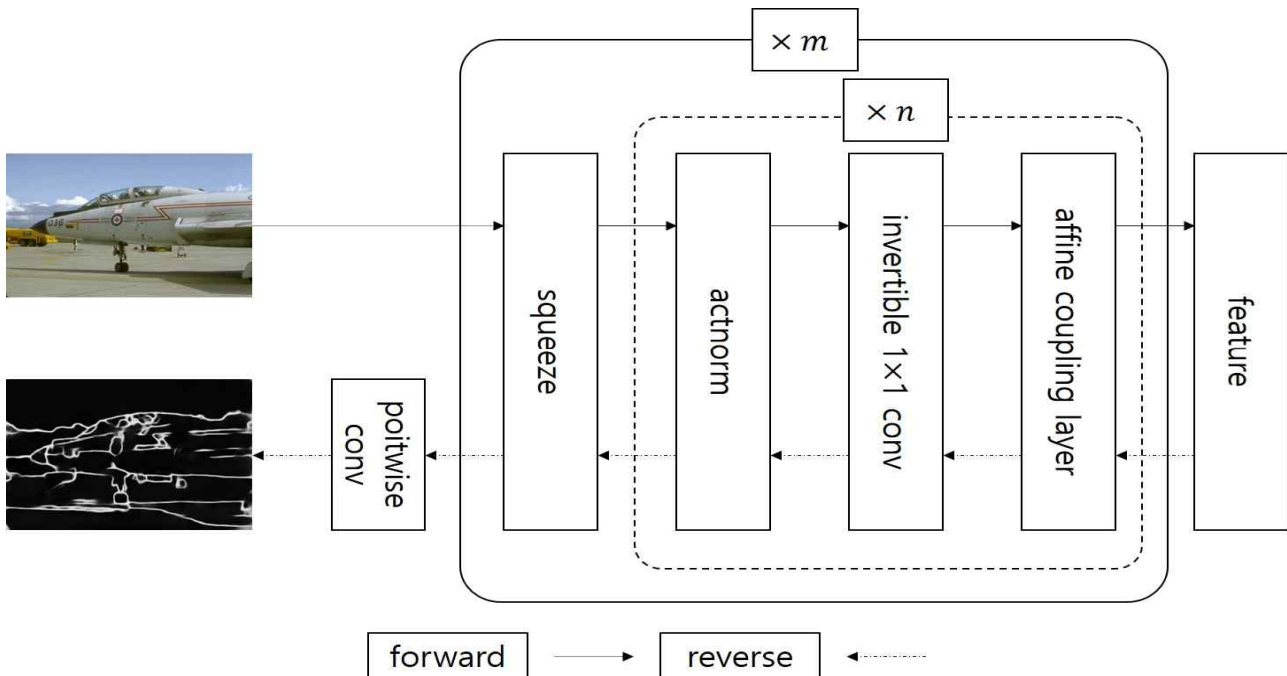


그림 4. GLOW 기반 에지 탐지 기법 구조

Fig. 4. The structure of the GLOW-based edge detection method

II. 관련 연구

에지 탐지는 크게 전통적 기법, 규칙 기반 기법, 딥러닝 기반 기법으로 나뉜다.

전통적 기법은 연산자를 이용하여 이미지에서 밝기값, 색, 텍스처 등의 불연속성을 탐지하여 에지를 추출한다. 대표적인 기법으로는 Sobel, Laplacian 연산자가 존재한다[1], [2]. 규칙 기반 기법은 수동으로 설계한 특징을 이용하여 에지를 추출하거나, 에지를 추출하기 위한 논리적 단계를 통해 에지를 추출한다. 대표적인 기법으로는 Canny 에지 탐지 기법과 Pb 기법 등이 존재한다[3], [11]. 전통적 기법과 규칙 기반 기법들은 낮은 수준 표현자를 이용하여 에지를 탐지하므로, 이미지에서 의미적으로 중요한 경계 부분만을 추출하는 데 한계가 존재한다.

딥러닝 성능이 향상되면서 높은 수준 표현자를 추출하는 것이 가능하다. 이에 따라, 딥러닝을 기반으로 한 에지 탐지 연구들이 활발히 진행되고 있다. 에지 탐지는 픽셀 단위의 에지 여부를 판단하여야 하고, 이미지에서 유의미한 에지를 추출하는 것이 중요하다. 또한, 이미지 내에서 객체가 다양한 크기로 존재하기 때문에 다중 계층 구조에 대한 표현자를 학습하는 것이 에지 탐지에 유용하다. 따라서, 기존 딥러닝 기반 기법들은 그림 3과 같이 계층적 표현자를 학습하기 위해 인코더와 디코더를 밀접 연결하거나 각 계층에 손실을 역전파하는 기법으로 학습한다[4]-[7], [12]-[14], [16].

III. 방법론

딥러닝 기반 에지 탐지 기법들은 주로 인코더와 디코더를 결합한 구조로 구성된다. 인코더의 풀링 연산 과정에서 공간 정보 손실로 인해 이미지 정보가 손실되어, 세밀한 에지를 추출하는데 문제가 발생한다[8].

Kingma et al.[10]은 딥러닝 생성 모델 중 하나인 GLOW를 제안하여 현실적인 외양의 얼굴 이미지를 합성하는데 성공하였다. An et al.[8]은 GLOW를 이용한 이미지 스타일 변환 기법을 제안하였다.

본 논문에서는 GLOW를 활용하여 이미지에서 의미 있는 에지를 세밀하게 탐지하는 기법을 제안한다. 그림 4는 하나의 네트워크를 통해 순방향과 역방향 추론을 수행하는 제안 기법의 구조를 보여준다. 순방향 추론을 통해 이미지의 특징을 추출하고, 역방향 추론을 통해 추출된 특징을 이미지로 재구성한다. 제안 기법은 크게 m 개의 block으로 구성되며, 각 block 당 n 개의 NF를 가진다. NF는 additive coupling, invertible 1x1 convolution, actnorm 순으로 이루어진 가역변환 함수들의 그룹이다.

Additive coupling은 어파인 결합 계층 (affine coupling layers)으로, 수식은 다음과 같다.

$$\begin{aligned} x_a, x_b &= \text{split}(x) \\ y_b &= NN(x_a) + x_b \\ y &= \text{concat}(x_a, y_b) \end{aligned} \quad (1)$$

여기서, x 는 입력값을 의미하고, $split()$ 은 x 의 차원을 절반으로 나눠주는 역할을 한다. $NN()$ 은 입력과 출력 크기를 동일하게 출력하는 어파인 변환을 수행하는 네트워크이다. $concat()$ 은 두 개의 입력값을 연결시켜주는 역할을 한다. 식 (1)의 역계산은 다음과 같다.

$$\begin{aligned} x_a, y_b &= split(y) \\ x_b &= y_b - NN(x_a) \\ x &= concat(x_a, x_b) \end{aligned} \quad (2)$$

Invertible 1×1 convolution은 입력값의 채널 순서를 조합시켜준다.

$$y_{i,j} = Wx_{i,j} \quad (3)$$

여기서, i, j 는 공간 위치를 의미하며, W 는 $c \times c$ 크기의 가중 행렬을 나타낸다. c 는 x 와 y 의 채널 차원 수를 의미한다. 역함수는 $x_{i,j} = W^{-1}y_{i,j}$ 이다.

Actnorm은 배치 정규화를 대체하기 위한 활성화 정규화(activation normalization) 층이다[10]. 다음과 같은 식 (4)를 통해 각 채널을 어파인 변환을 한다.

$$y_{i,j} = w \odot x_{i,j} + b \quad (4)$$

여기서, w 와 b 는 어파인 변환의 크기와 편향 변수를 의미한다. 역함수는 $x_{i,j} = (y_{i,j} - b)/w$ 이다.

Squeeze 연산자는 특징맵의 공간 크기를 줄이기 위한 역할을 수행한다. 구체적으로는 특징맵을 작은 패치로 나눈 후 채널 차원으로 연결한다. Pointwise 컨볼루션은 에지 탐지에서 색상 영향을 최소화시키기 위해 출력 채널을 하나의 채널로 통합한다. 또한, 본 논문에서는 이진 크로스 엔트로피(binary cross entropy) 손실함수를 사용한다.

IV. 실험 결과

4-1 데이터셋

본 논문에서 제안 기법을 학습 및 평가하기 위해 BSDS500 데이터셋과 BIPED 데이터셋을 이용하였다.

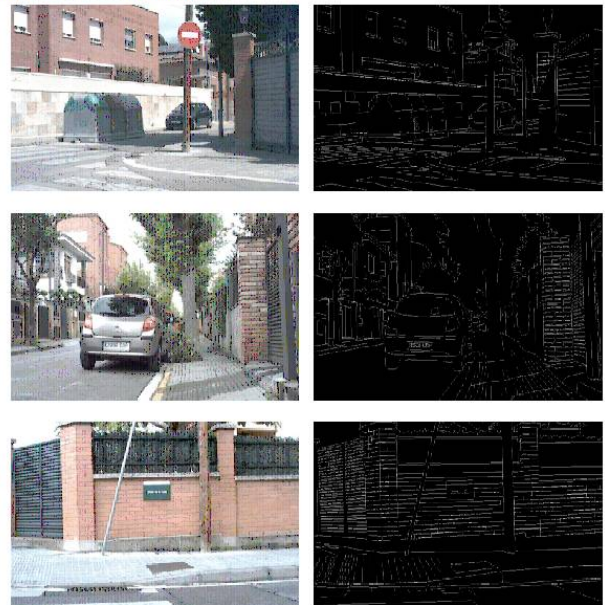
BSDS500 데이터셋은 500개의 자연 이미지 데이터셋으로 구성되며, 다수 사용자에 의해 라벨링 된 데이터셋이다. 그림 5(a)는 BSDS500 데이터셋의 일부를 보여주는데, 데이터셋의 일관성이 다소 떨어짐을 확인할 수 있다. 구체적으로, 그림 5(a) 두 번째와 네 번째 열은 사람 윤곽만을 예지로 라벨링하

였으나, 그림 5(a)의 첫 번째와 세 번째 열은 사람이 입은 옷의 경계도 라벨링한 것을 확인할 수 있다.

이에 비해 BIPED 데이터셋은 250개 실외 이미지 데이터셋으로 구성되며, 컴퓨터 비전 전문가들에 의해 라벨링 되어 BSDS500 데이터셋에 비해 일관성이 높다. 그림 5(b)는 BIPED 데이터셋 일부를 보여주며, 차, 도로, 그리고 벽 등이 일관성 있게 라벨링 되어있다. 또한 BSDS500보다 좀 더 촘촘히 에지가 나와 있다.



(a) BSDS500 데이터셋
(a) BSDS500 dataset



(b) BIPED 데이터셋
(b) BIPED dataset

그림 5. BSDS500 데이터셋과 BIPED 데이터셋 예시
Fig. 5. Examples of BSDS500 and BIPED dataset

4-2 결과 분석

본 논문에서 제안 기법에 대해 분석하기 위해 block과 NF 수에 따른 기법의 결과물을 비교하였다. 또한, BSDS500 훈련 데이터셋과 BIPED 훈련 데이터셋으로 각각 학습을 수행한 후, Poma et al. [13]가 제안한 Dexined 기법의 결과물과 비교하였다. 구체적인 내용은 다음과 같다.

1) Block 수에 따른 분석

제안 기법을 구성하는 네트워크는 크게 m 개 block과 각 block 당 n 개의 NF로 구성된다. 본 실험에서는 block에 수에 따른 결과물을 분석하고, 적절한 m 을 설정하는 것을 목표로 한다. Block에 따른 결과물 분석을 위해 NF 수를 5에서 20까지 5씩 변경한 결과물을 비교하였으며, 그림 6과 7은 NF를 15로 설정한 결과물이다.

그림 6(a)와 7(a)는 제안 기법을 BSDS500 훈련 데이터셋으로 학습한 기법의 결과물을 보여준다. 그림 6(a)의 빨간 영역을 보면, block-1과 block-3에서는 차 외의 불필요한 부분이 에지로 탐지되었으나, block-2에서는 차 객체만의 의미론적인 에지가 탐지되었다. 그러나, 그림 7(a)에서는 block 수와 관계없이 제안 기법이 나무의 형상 또는 차선의 경계들을 에지로 탐지하지 못하였음을 확인할 수 있다. 이를 통해 BSDS500 훈련 데이터셋으로 학습한 기법을 BSDS500 평가 데이터셋에 적용하였을 경우 전반적으로 에지들을 유의미하게 탐지하였으나, BIPED 평가 데이터셋에 적용할 경우 에지를 일관되게 탐지하지 못하였음을 확인할 수 있다.

그림 6(b)와 7(b)는 제안 기법을 BIPED 훈련 데이터셋으로 학습한 기법의 결과물을 보여준다. 그림 6(b) 빨간 영역을 보면, block-1과 block-3에서는 비행기 외 배경이 에지로 탐지되었으나, block-2에서는 배경 부분이 상대적으로 적게 에지로 탐지되었다. 이를 통해 block-2일 때 block-1과 block-3에 비해 불필요한 에지를 탐지하지 않음을 확인할 수 있다. 그림 7(b) 보면, 전반적으로 block 수가 증가할수록 불필요한 에지가 줄어들었으며, block-2와 block-3에서 나무의 형상이 유의미하게 추출됨을 확인할 수 있다. 이를 통해 BIPED 훈련 데이터셋으로 학습한 기법의 경우 BSDS500 평가 데이터셋과 BIPED 평가 데이터셋에서 모두 유의미하게 에지를 탐지함을 알 수 있다.

본 실험을 통해 BIPED 훈련 데이터셋으로 학습한 경우 BSDS500 훈련 데이터셋에 비해 에지들이 일관되게 탐지됨을 확인하였다. 또한, block-2에서 block-1과 block-3에 비해 불필요한 에지가 줄어들고, 유의미한 에지들이 탐지되었음을 확인하였다. 이에 제안 기법의 네트워크를 구성하는 block 수를 2로 설정하였다.

2) NF 수에 따른 분석

제안 기법에서 NF 수에 따른 영향을 분석하기 위해, block의 수를 2로 고정하고 NF 수를 10부터 30까지 10씩 변경한 결과물을 비교하였다.

그림 8(a)와 9(a)는 BSDS500 훈련 데이터셋으로 학습한 기법을 각각 BSDS500 평가 데이터셋과 BIPED 평가 데이터셋에 적용한 결과를 보여준다. 그림 8(a)를 보면, NF 수가 증가할수록 사슴의 형상이 세밀하게 추출되었다. 그러나, 사슴 형상 외의 불필요한 배경들도 에지로 검출되는 문제점을 확인할 수 있다. 그림 9(a)에서는 BIPED 데이터셋에서 전반적으로 나무의 형상이나 도로의 경계들이 에지로 추출되지 않았다. 이는 앞서 말했듯이 BSDS500 데이터셋 구성이 일관되지 않은 요인으로 분석된다.

그림 8(b)와 9(b)는 BIPED 훈련 데이터셋으로 학습한 기법을 BSDS500 평가 데이터셋과 BIPED 평가 데이터셋에 적용한 결과를 보여준다. 그림 8(b)를 보면, 전반적으로 NF 수가 증가할수록 사슴의 형상이 에지로 유의미하게 추출되었다. 또한 그림 9(b)의 빨간 영역과 같이 나무의 형상 외에 불필요한 부분이 NF 수가 증가할수록 줄어들음을 확인할 수 있다.

본 실험을 통해 전반적으로 BSDS500 데이터셋으로 학습하였을 때 에지들이 일관되게 탐지되지 않았으며, NF 수가 증가할수록 객체의 형상인 유의미한 부분이 에지로 탐지되었다. 따라서, 제안 기법의 네트워크를 구성하는 NF 수를 30으로 설정하였다.

3) 기존 기법과 정성적 비교

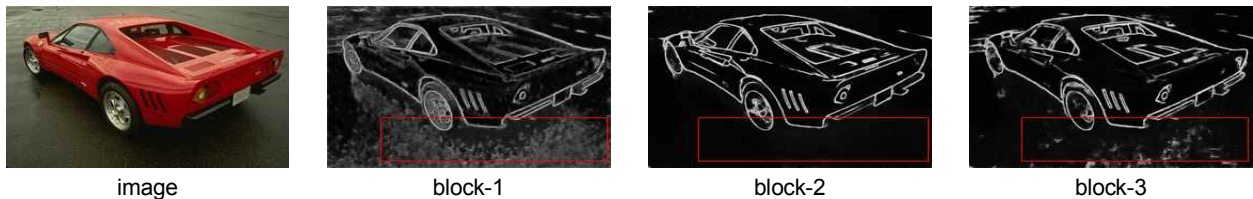
본 절에서는 block-2와 NF-30으로 설정된 제안 기법을 Dexined avg/fused 기법 [13]과 비교하였다. Dexined 기법은 BIPED 데이터셋으로 훈련되었으며, 인코더와 디코더로 구성된 기법이다. 이에, 제안 기법도 BIPED 데이터셋으로 학습한 기법을 비교하였다. 그림 10은 제안 기법과 Dexined 기법을 BSDS500 평가 데이터셋과 BIPED 평가 데이터셋에 적용한 결과물을 보여준다. 그림 10의 첫 번째와 두 번째 행을 보면, Dexined 기법의 경우 비행기와 새 외의 배경이 에지로 다수 탐지되었으나, 제안 기법의 경우 불필요한 에지 탐지가 줄어들었음을 확인할 수 있다. 그림 10의 세 번째와 네 번째 행에서는 Dexined 기법과 제안 기법을 통해 나무의 형상, 도로 경계와 차의 형상 등이 모두 우수하게 에지로 탐지되었다. 그러나, 전반적으로 Dexined 기법이 제안 기법에 비해 에지를 두껍게 탐지하였다. 이는 Dexined 기법의 경우 세밀한 에지를 추출하기 위해 후처리 과정이 필요하다는 것을 의미한다 [14]. 결론적으로, 제안 기법이 Dexined 기법에 비해 객체 외의 배경과 같은 불필요한 에지들을 줄이고, 유의미한 에지를 세밀하게 탐지하였다.

제안 기법과 Dexined 기법은 일상 이미지로 구성된 훈련 데이터셋을 이용하여 학습된 기법이다. 따라서, 일상 이미지로 평가를 하였을 경우, 이미지 정보 손실 여부를 판단하는데 어려움이 존재한다. 따라서, 이 두 기법을 다른 도메인인 예술 작품 중의 회화 이미지에 적용해 보았다. 회화는 작가의 기법, 캔버스 재질, 물감 등에 따라서 다양한 에지가 존재한다. 이러한 에지에는 작품 내에 작가가 의도하는 에지와 재료, 작가 드로잉 기법 등에 의해 의도하지 않는 에지가 존재하게

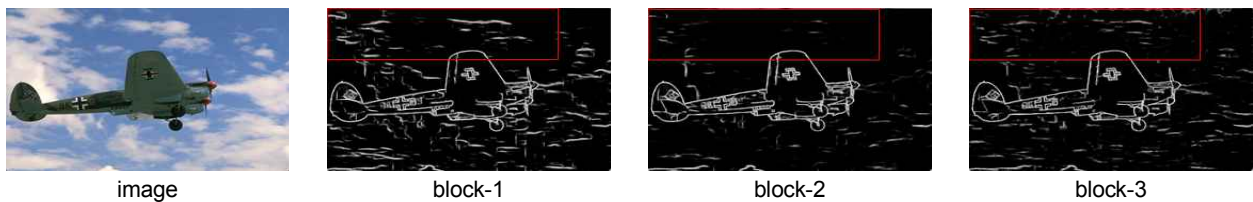
되는데 우리는 작가가 의도한 의미 있는 에지를 추출하고자 한다. 그림 11은 두 기법을 회화 이미지에 적용한 결과를 보여준다. 첫 번째 행에서는 작은 객체들이 Dexined 기법의 경우 에지로 탐지되지 않았으나, 제안 기법에서는 작은 객체들도 에지로 유의미하게 탐지되었다. 두 번째 행에서는 Dexined 기법의 경우 에지가 뭉개져 탐지되었으나, 제안 기법에 경우 세밀하게 에지를 탐지함을 확인할 수 있다. 이를 통해, 제안 기법이 이미지 정보 손실을 최소화하며 에지를 유의미하게 탐지함을 보여준다.

본 논문에서는 대상 이미지에서 보다 유의미한 에지를 탐지하기 위한 GLOW 기반 에지 탐지 기법을 제안하였다. 제안 기법은 역변환 가능한 네트워크로 구성되며, 하나의 네트워크를 통해 이미지의 특징을 추출하고 추출된 특징들을 이미지로 재구성한다. 제안 기법을 학습 및 평가하기 위해 BSDS500 데이터셋과 BIPED 데이터셋을 활용하였다. 이를 통해 제안 기법이 일상 이미지에서 에지를 유의미하게 탐지함을 확인하였다. 또한, 제안 기법과 기존 기법을 회화 작품에 적용하여, 제안 기법이 기존 기법에 비해 에지를 세밀하게 탐지함을 확인하였다.

V. 결 론



(a) BSDS500 훈련 데이터셋으로 학습된 제안 기법의 결과
 (a) Results of the proposed method trained on BSDS500 training dataset



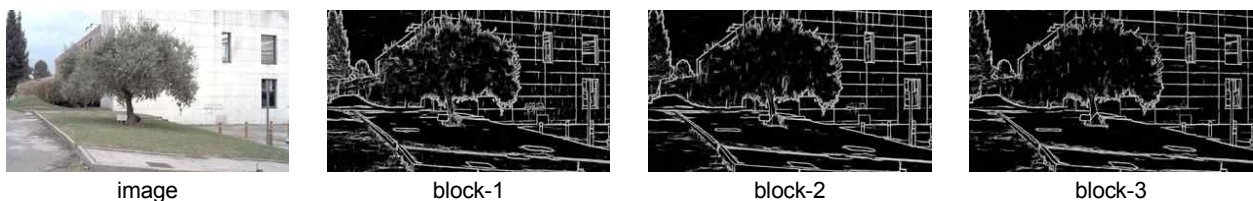
(b) BIPED 훈련 데이터셋으로 학습된 제안 기법의 결과
 (b) Results of the proposed method trained on BIPED training dataset

그림 6. block 수에 따른 제안 기법의 BSDS500 평가 데이터셋 적용 결과

Fig. 6. Results of applying the proposed method to BSDS500 test dataset according to the number of blocks



(a) BSDS500 훈련 데이터셋으로 학습된 제안 기법의 결과
 (a) Results of the proposed method trained on BSDS500 training dataset



(b) BIPED 훈련 데이터셋으로 학습된 제안 기법의 결과
 (b) Results of the proposed method trained on BIPED training dataset

그림 7. block 수에 따른 제안 기법의 BIPED 평가 데이터셋 적용 결과

Fig. 7. Results of applying the proposed method to BIPED test dataset according to the number of blocks



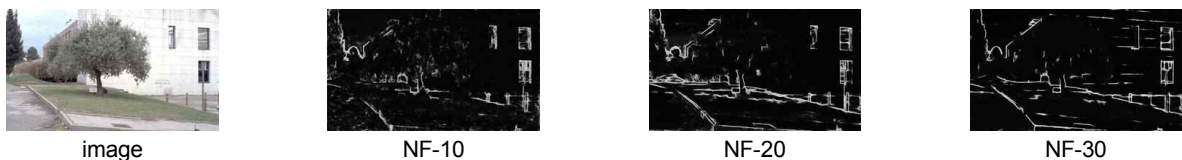
(a) BSDS500 훈련 데이터셋으로 학습된 제안 기법의 결과
 (a) Results of the proposed method trained on BSDS500 training dataset



(b) BIPED 훈련 데이터셋으로 학습된 제안 기법의 결과
 (b) Results of the proposed method trained on BIPED training dataset

그림 8. NF 수에 따른 제안 기법의 BSDS500 평가 데이터셋 적용 결과

Fig. 8. Results of applying the proposed method to BSDS500 test dataset according to the number of NF



(a) BSDS500 훈련 데이터셋으로 학습된 제안 기법의 결과
 (a) Results of the proposed method trained on BSDS500 training dataset



(b) BIPED 훈련 데이터셋으로 학습된 제안 기법의 결과
 (b) Results of the proposed method trained on BIPED training dataset

그림 9. NF 수에 따른 제안 기법의 BIPED 평가 데이터셋 적용 결과

Fig. 9. Results of applying the proposed method to BIPED test dataset according to the number of NF

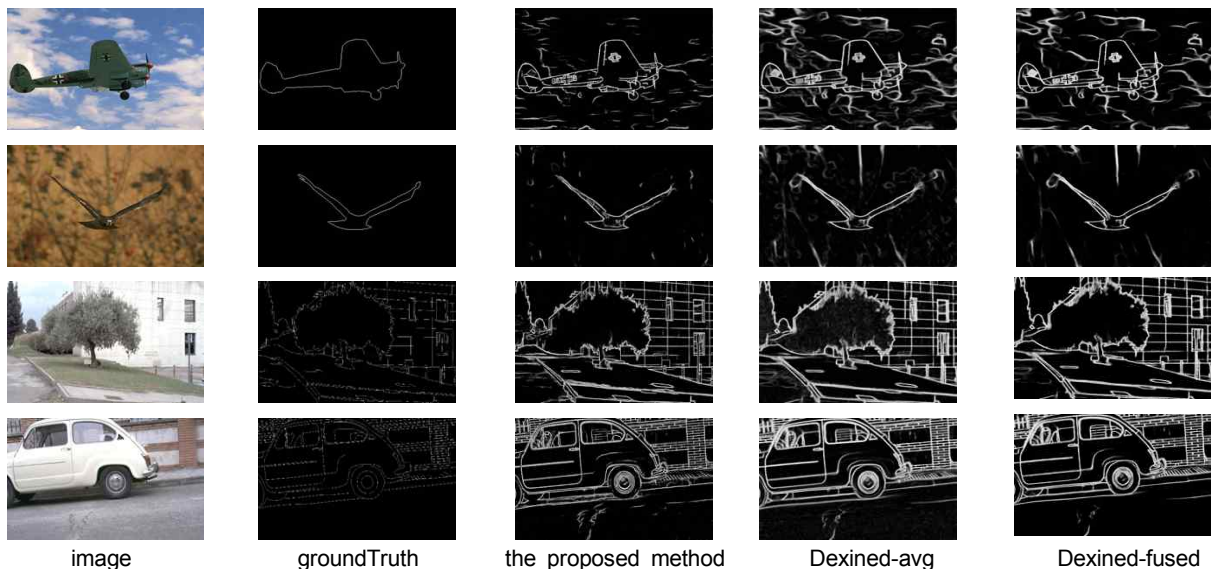


그림 10. 제안 기법과 Dexined 기법의 BSDS500 데이터셋과 BIPED 데이터셋 적용 결과

Fig. 10. Results of applying the proposed method and Dexined method to BSDS500 dataset and BIPED dataset

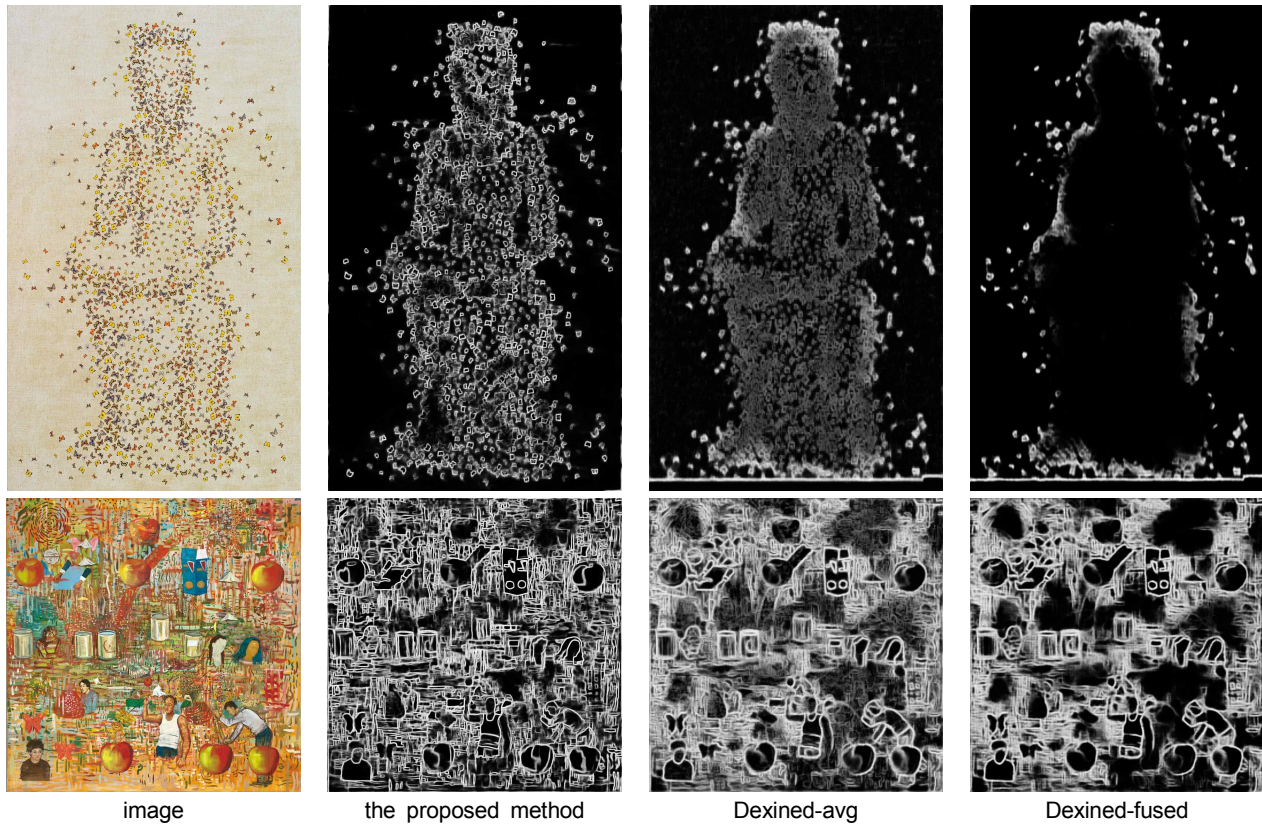


그림 11. 제안 기법과 Dexined 기법의 회화 작품 적용 결과

Fig. 11. Results of applying the proposed method and Dexined method to painting work images

추후 정량적인 분석을 통해 제안 기법의 네트워크를 최적화할 예정이며, 회화 이미지와 같이 작품에서 작가가 의도한 유의미한 에지를 추출함으로써 회화 작품의 밀그림을 추출하는 방법 등 다양한 도메인에 적용할 수 있도록 발전시킬 예정이다.

감사의 글

본 연구는 문화체육관광부 및 한국콘텐츠진흥원의 2020년도 문화기술연구개발지원사업의 연구결과로 수행되었음(과제번호: R2020060004)

참고문헌

[1] N. Kanopoulos, N. Vasanthavada, and R. L. Baker, "Design of an image edge detection filter using the Sobel operator," *IEEE Journal of solid-state circuits*, Vol. 23, No. 2, pp. 358-367, April, 1988. <https://doi.org/10.1109/4.996>

[2] A. Anand, S. S. Tripathy, and R. S. Kumar, "An improved edge detection using morphological Laplacian of Gaussian operator," *In 2015 2nd International conference on signal processing and integrated networks (SPIN)*, Noida, pp.

532-536, February, 2015.

<https://doi.org/10.1109/SPIN.2015.7095391>

[3] D. R. Martin, C. C. Fowlkes, and J. Malik, "Learning to detect natural image boundaries using local brightness, color, and texture cues," *IEEE transactions on pattern analysis and machine intelligence*, Vol. 26, No. 5, pp. 530-549, March, 2004. <https://doi.org/10.1109/TPAMI.2004.1273918>

[4] Y. Liu, M. M. Cheng, X. Hu, K. Wang, and X. Bai, "Richer convolutional features for edge detection," *In Proceedings of the IEEE conference on computer vision and pattern recognition*, Honolulu, pp. 3000-3009, October, 2017. <https://doi.org/10.1109/TPAMI.2018.2878849>

[5] S. Xie, and Z. Tu, "Holistically-nested edge detection," *In Proceedings of the IEEE international conference on computer vision*, Boston, pp. 1395-1403, February, 2015. <https://doi.org/10.1109/ICCV.2015.164>

[6] G. Bertasius, J. Shi, and L. Torresani, "Deepedge: A multi-scale bifurcated deep network for top-down contour detection," *In Proceedings of the IEEE conference on computer vision and pattern recognition*, Boston, pp. 4380-4389, October, 2015. <https://doi.org/10.1109/CVPR.2015.7299067>

- [7] H. Yang, Y. Li, X. Yan, and F. Cao, "ContourGAN: Image contour detection with generative adversarial network," *Knowledge-Based Systems*, Vol. 164, pp. 21-28, January, 2019. <https://doi.org/10.1016/j.knosys.2018.09.033>
- [8] J. An, S. Huang, Y. Song, D. Dou, W. Liu, and J. Luo, "ArtFlow: Unbiased Image Style Transfer via Reversible Neural Flows," *In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 862-871, June, 2021. <https://doi.org/10.1109/CVPR46437.2021.00092>
- [9] L. Dinh, D. Krueger, and Y. Bengio, "Nice: Non-linear independent components estimation," arXiv preprint arXiv:1410.8516, October, 2014.
- [10] D. P. Kingma, and P. Dhariwal, "Glow: Generative flow with invertible 1x1 convolutions," *32nd Conference on Neural Information Processing Systems*, Montréal, December, 2018.
- [11] L. Ding, and A. Goshtasby, "On the Canny edge detector," *Pattern Recognition*, Vol. 34, No. 3, pp. 721-725, March, 2001. [https://doi.org/10.1016/S0031-3203\(00\)00023-6](https://doi.org/10.1016/S0031-3203(00)00023-6)
- [12] J. He, S. Zhang, M. Yang, Y. Shan, and T. Huang, "Bi-directional cascade network for perceptual edge detection," *In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, pp. 3828-3837, June, 2019. <https://doi.org/10.1109/TPAMI.2020.3007074>
- [13] X. S. Poma, E. Riba, and A. Sappa, "Dense extreme inception network: Towards a robust cnn model for edge detection," *In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, Colorado, pp. 1923-1932, May, 2020. <https://doi.org/10.1109/WACV45572.2020.9093290>
- [14] R. Deng, C. Shen, S. Liu, H. Wang, and X. Liu, "Learning to predict crisp boundaries," *In Proceedings of the European Conference on Computer Vision (ECCV)*, Munich, pp. 562-578, September, 2018. https://doi.org/10.1007/978-3-030-01231-1_35
- [15] M. Li, Z. Lin, R. Mech, E. Yumer, and D. Ramanan, "Photo-sketching: Inferring contour drawings from images," *In 2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*, Hawaii, pp. 1403-1412, January, 2019. <https://doi.org/10.1109/WACV.2019.00154>
- [16] Z. Yu, C. Feng, M. Y. Liu, and S. Ramalingam, "Casenet: Deep category-aware semantic edge detection," *In Proceedings of the IEEE conference on computer vision and pattern recognition*, Hawaii, pp. 5964-5973, November, 2017. <https://doi.org/10.1109/CVPR.2017.191>



전형주(Hyeong-Ju Jeon)

2015년 : 서울대학교 학사
2018년 : 서울대학교 대학원 (공학석사)

2018년~현 재: 한국전자통신연구원

※관심분야 : 딥러닝 (Deep Learning), 기계 학습 (Machine Learning), 컴퓨터비전 (Computer Vision) 등



정순철(Soonchul Jung)

1998년 : 카이스트 학사
2000년 : 서울대학교 대학원 석사 (공학석사)
2006년 : 서울대학교 대학원 박사 (공학박사 - 컴퓨터공학과)

2006년~2012년: KT 연구원

2013년~현 재: 한국전자통신연구원

※관심분야 : 기계 학습 (Machine Learning), 진화 연산 (Evolutionary Computation), 컴퓨터비전 (Computer Vision), 가상현실 (VR) 등



최운석(Yoon-Seok Choi)

1996년: 중앙대학교 학사
1998년: 중앙대학교 대학원 석사 (공학석사)
2006년: 서울대학교 대학원 박사 (공학박사 - 컴퓨터공학과)

2006년~현 재: 한국전자통신연구원

※관심분야 : 비사실적렌더링 (NPR), 가상현실 (VR), 증강현실 (AR), 기계 학습 (Machine Learning), 진화 연산 (Evolutionary Computation) 등



김재우(Jae Woo Kim)

1991년 : 한국외국어대학교 학사
1993년1: 한국외국어 대학원 석사 (이삭석사)
2009년 : 조지워싱턴대학원 박사 (이학박사 - 컴퓨터공학과)

1993년~1999년: 한국전자통신연구원

2003년~2010년: VR Sonic Inc. Senior Research Scientist

2011년: Hampton University Physics Dept. 연구조교수

2012년~현 재: 한국전자통신연구원

※관심분야 : 컴퓨터 비전 (Computer Vision), 컴퓨터 그래픽스 (Computer Graphics), 인공지능 (AI)



김진서(Jin Seo Kim)

1991년 : 인천대학교 학사
1993년 : 뉴욕 폴리텍 대학원 석사 (공학석사)
2009년 : 리즈대학교 대학원 박사 (공학박사 - 칼라 과학)

1993년~현 재: 한국전자통신연구원

※관심분야 : 컬러과학 (Color Science), 장치 간 컬러 재현 (Cross Media Color Reproduction), 이미지 처리 (Image Processing), 이미지 품질 개선 (Image Quality Enhancement) 등