

특집논문 (Special Paper)

방송공학회논문지 제24권 제1호, 2019년 1월 (JBE Vol. 24, No. 1, January 2019)

<https://doi.org/10.5909/JBE.2019.24.1.67>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

시멘틱 세그멘테이션을 활용한 이미지 오브젝트의 효율적인 영역 추론

임 현 영^{a)}, 이 유 림^{b)}, 지 민 규^{c)}, 고 명 현^{a)}, 김 학 동^{a)}, 김 원 일^{d)†}

Efficient Inference of Image Objects using Semantic Segmentation

Heonyeong Lim^{a)}, Yurim Lee^{b)}, Minkyu Jee^{c)}, Myunghyun Go^{a)}, Hakdong Kim^{a)}, and Wonil Kim^{d)†}

요 약

본 연구에서는 다중 라벨링이 되어 있는 이미지 데이터를 대상으로 시멘틱 세그멘테이션을 활용한 효율적인 오브젝트별 영역 분류 기법을 연구한다. 이미지 데이터에 포함된 색상 정보, 윤곽선, 명암, 채도 등 다양한 픽셀 단위 정보와 프로세싱 기법뿐만 아니라 각 오브젝트들이 위치한 세부 영역을 의미 있는 단위로 추출하여 추론 결과에 반영하는 실험을 진행하고 그 결과에 대해 논의한다. 이미지 분류에서 훌륭한 성능을 검증받은 뉴럴 네트워크를 활용하여 비정형성이 심하고 다양한 클래스 오브젝트가 포함된 이미지 데이터를 대상으로 어떤 오브젝트가 어디에 위치하였는지 파악하는 작업을 진행한다. 이러한 연구를 기반으로 향후 다양한 오브젝트가 포함된 복잡한 이미지의 실시간 세부 영역 분류를 진행하는 인공지능 서비스 제공을 목표로 한다.

Abstract

In this paper, we propose an efficient object classification method based on semantic segmentation for multi-labeled image data. In addition to various pixel unit information and processing techniques such as color information, contour, contrast, and saturation included in image data, a detailed region in which each object is located is extracted as a meaningful unit and the experiment is conducted to reflect the result in the inference. We use a neural network that has been proven to perform well in image classification to understand which object is located where image data containing various class objects are located. Based on these researches, we aim to provide artificial intelligence services that can classify real-time detailed areas of complex images containing various objects in the future.

Keyword : Semantic segmentation, Deep learning, Image processing, Object recognition, Feature extraction

a) 세종대학교 디지털콘텐츠학과(Department of Digital Contents, Sejong University)

b) 세종대학교 인공지능언어공학과(Department of Artificial Intelligence and Linguistic Engineering, Sejong University)

c) 세종대학교 소프트웨어융합학과(Department of Software Convergence, Sejong University)

d) 세종대학교 소프트웨어학과(Department of Software, Sejong University)

† Corresponding Author : 김원일(Wonil Kim)

E-mail: wikim@sejong.ac.kr

Tel: +82-3408-3795

ORCID: <https://orcid.org/0000-0002-1489-8427>

※ This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(2015R1D1A1A01060693)

· Manuscript received November 15, 2018; Revised December 31, 2018; Accepted December 31, 2018.

Copyright © 2016 Korean Institute of Broadcast and Media Engineers. All rights reserved.

“This is an Open-Access article distributed under the terms of the Creative Commons BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited and not altered.”

1. 서론

시맨틱 세그멘테이션(Semantic segmentation)^[1]은 이미지 내에서 오브젝트를 검출해낼 때 바운딩 박스(Bounding box)와 같은 정적인 범위로 나타내는 것이 아니라, 오브젝트의 픽셀 단위 예측을 수행하여 의미 있는 단위로 대상의 픽셀 영역을 분리해내는 것이다. 이미지에 어떤 오브젝트가 있는지를 의미적(Semantic) 단위에서 확인하는 것 뿐 아니라 해당 오브젝트의 공간적 특성(Spatial feature)까지 함께 파악하여 세그멘테이션 결과에 반영해야 좋은 성능을 얻을 수 있다. 최근 이러한 작업에 가장 보편적으로 사용되고 있는 방법은 이미지 분류(Classification)에서 높은 성능을 검증받은 CNN(Convolutional neural network)^[2]을 접목하는 것이다. CNN을 기반으로 한 네트워크들은 대부분 마지막 계층에 클래스 분류를 위한 완전 연결 계층(Fully connected layer)이 위치하게 된다. 고정된 크기의 입력만을 받아들이며 공간적 특성을 사라지게 하는 이러한 아키텍처는 시맨틱 세그멘테이션 분야에 적용이 힘든 상황이었으나, 2014년 발표된 FCN(Fully convolutional networks)^[3]에서는 이러한 완전 연결 계층을 1 x 1 컨볼루션 계층(Convolution layer)으로 대체하는 방법을 제시하였다. 이는 이미지 속 각 오브젝트의 의미적 단위와 공간적 특성을 동시에 추출할 수 있게 하였고, 시맨틱 세그멘테이션의 성능을 크게 향상시키는 계기가 되었다.

본 논문에서는 시맨틱 세그멘테이션 아키텍처를 활용하여 연구용으로 공개된 패션 이미지 데이터^[4]에 대한 오브젝트 영역 추론을 수행하고 그 결과를 공유한다. 패션 이미지의 경우 비정형성이 심하기 때문에 다양한 클래스 및 속성 정보를 효율적으로 태깅하는 것이 큰 관건인데, 현재 대다수의 패션 전자 상거래 서비스에서는 이러한 작업을 인력이 동원된 수작업으로 진행하고 있는 상황이다. 이는 그 속도의 효율성 및 결과의 정확도가 상당히 떨어지는 문제점을 야기하며, 이러한 점을 극복하기 위해 시맨틱 세그멘테이션 아키텍처에 밀집 연결성(Dense connectivity) 개념을 접목하여 59개 카테고리의 픽셀 영역을 효율적으로 추론하는 방법을 제시한다.

실험에 사용되는 패션 이미지 데이터의 해상도는 850x550/600x400로 구성되어 있으며, 이미지 인식에 사용하기

에는 다소 높은 해상도라고 판단되어 이미지의 가로 길이를 224로 리스케일링하여 사용하였다. 다양한 오브젝트가 포함된 패션 이미지에 대하여 컨볼루션(Convolution)과 풀링(Pooling)을 반복했을 시 각 오브젝트의 의미적 단위가 쉽게 유실되어 버리는데, 이러한 문제를 해결하기 위하여 각 단계에서 추출한 특징 지도(Feature map)를 재사용하는 Densenet^[5] 기반 아키텍처를 활용함으로써 다음 계층으로 전달되어야 할 오브젝트 정보의 유실을 방지하였다.

논문의 구성은 다음과 같다. II장에서는 시맨틱 세그멘테이션 기술을 간단히 소개함으로써 입력 이미지가 다운샘플링(Downsampling), 업샘플링(Upsampling) 등의 과정을 거쳐 픽셀 단위로 다양한 오브젝트를 태깅한 라벨링 결과를 얻는 과정에 대한 이해를 돕고자 한다. III장에서는 Densenet 기반 아키텍처를 활용하여 패션 이미지 속 다양한 오브젝트의 영역을 추론하는 방법을 제안한다. IV장에서는 실험 과정과 그 결과에 대해 논의한다. V장에서는 결론 및 추후 연구 방향을 제시한다.

II. Semantic Segmentation

1. Semantic Segmentation 개요

세그멘테이션은 이미지 인식 분야의 주요 응용 기술 중 하나로, 이미지 내의 오브젝트들을 의미 있는 단위로 분리하여 구분하는 기술이다. 여기서 이미지의 스레쉬 홀드(Threshold)나 엣지(Edge), 색(Color), 명암(Contrast) 등의 속성에 기반한 단순 구분이 아닌 오브젝트의 의미적 단위와 공간적 특성을 함께 반영하여 오브젝트들을 의미 있는 단위로 구분하는 것을 시맨틱 세그멘테이션이라고 칭한다. 이미지 인식 분야에서 CNN을 활용하여 그 성능이 대폭 개선된 것처럼, 시맨틱 세그멘테이션 분야 또한 CNN을 접목하여 성능을 개선시키는 연구들이 진행되었다.

2. CNN 기반 Semantic Segmentation

CNN은 공간적 정보를 활용하여 입력 데이터에서 인접한 신호들에 대한 상호연관 관계를 컨볼루션(Convolution)

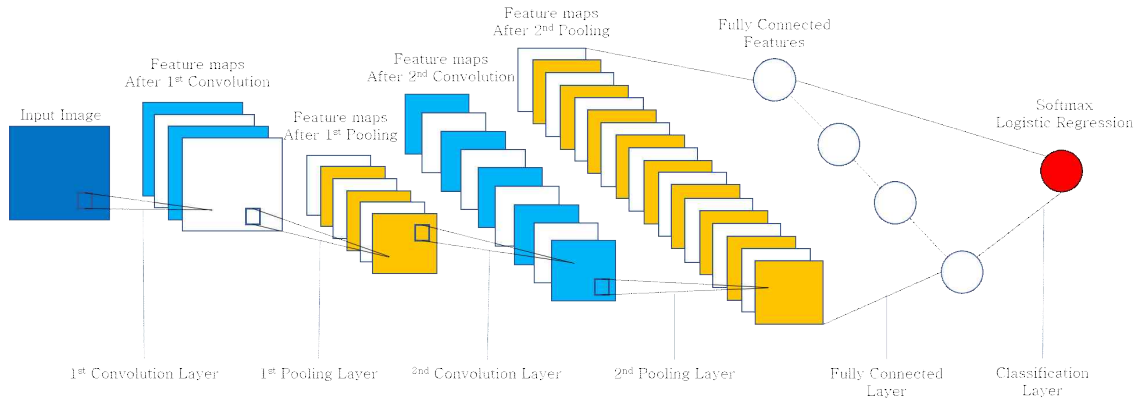


그림 1. 컨볼루션 뉴럴 네트워크의 아키텍처 예시 (이미지 분류)
Fig. 1. Example of convolutional neural networks (Image classification)

필터를 적용하여 추출해 내는 아키텍처이다. 이러한 필터를 여러 번 적용하면 다양한 공간적 특성을 데이터에서 추출하여 특징 지도를 만들어 낼 수 있다. 그 후 맥스 풀링(Max pooling)과 같은 여러 번의 다운샘플링(Downsampling) 과정을 거쳐 영상의 크기를 줄이고 이동이나 변형 등에 무관한 강한 전역적 특징(Global feature)을 얻어냄으로써 이미지 인식 분야에 있어서 뛰어난 성능을 보이고 있다. 이미지를 입력으로 받아 정해진 클래스를 출력하는 분류

아키텍처와 달리 시멘틱 세그멘테이션 아키텍처는 이미지를 입력으로 받아 동일한 사이즈를 유지하면서 각 오브젝트를 정해진 픽셀 단위로 구분해놓은 세그멘테이션 결과를 출력하게 된다. 2014년 발표된 FCN의 개발자들은 여러 번의 컨볼루션 필터와 다운샘플링 계층을 거친 데이터의 전역적 특징들이 시멘틱 세그멘테이션에 필요한 중요한 정보들을 포함하고 있음을 보여주었으며, 정해진 크기의 입력만을 받아들이는 완전 연결 계층을 1 x 1 컨볼루션 계층으

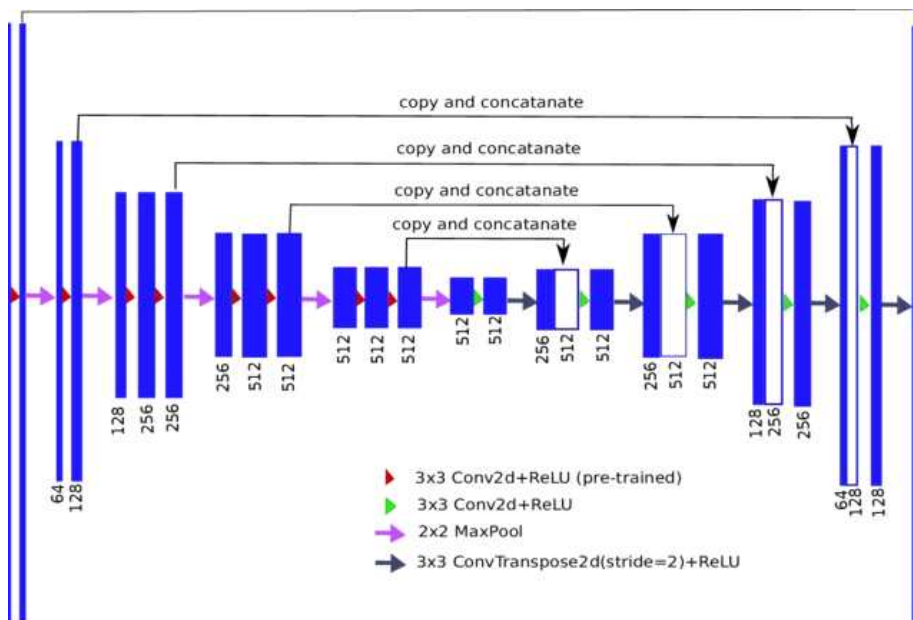


그림 2. VGG11 아키텍처를 기반으로 한 U-net 아키텍처 (시멘틱 세그멘테이션)^[8]
Fig. 2. U-net architecture based on VGG11 model (Semantic segmentation)^[8]

로 대체함으로써 전체 아키텍처를 하나의 CNN으로 보는 방식을 제안하였다. 그러나 이미지 분류와 달리 시멘틱 세그멘테이션은 픽셀 단위의 디테일한 예측이 필요하다는 차이점이 존재한다. 이미지 분류를 목적으로 구성된 CNN 아키텍처는 여러 번의 컨볼루션, 다운샘플링을 거치면서 강하고 전역적인 특징만을 전달하게 하는데, 이렇게 전달받은 특징을 단순히 업스케일링(Upscaling)하는 식으로 시멘틱 세그멘테이션 아키텍처를 구성하게 되면 오브젝트의 디테일한 특징을 얻는 것이 어렵게 되는 것이다.

이러한 문제의 해결을 위해 FCN을 기반으로 한 아키텍처로 U-net^[6]이 2015년에 제시되었다. U-net의 저자들은 이미지의 시멘틱 세그멘테이션을 위한 다운샘플링-업샘플링을 진행할 때, 아래 계층에서 얻어낸 오브젝트의 특징을 상위 계층에서 얻어낸 특징에 연쇄(Concatenate)시켜줌으로써 좀 더 정확한 오브젝트의 지역화(Localization)와 디테일한 예측을 얻어내는 아키텍처를 제시하였다. U-net은 시멘

틱 세그멘테이션 분야에서 폭넓게 사용되고 있으며, 이미지 입력 - 다운샘플링 - 특징 추출 - 업샘플링 - 디테일 확보 - 그라운드 트루쓰(Ground truth) 출력이라는 시멘틱 세그멘테이션의 기본 과정들을 상세히 보여주고 있다. 본 논문에서는 이러한 U-net을 기반으로 하되, 이미지에서 추출한 특징들을 재사용하는 방식으로 오브젝트의 디테일을 확보한 Fully convolutional densenet^[7]을 활용하여 시멘틱 세그멘테이션을 진행한다.

III. 제안 방법

본 논문에서는 멀티 라벨 이미지의 효과적인 오브젝트 영역 추론을 위하여 밀집 연관성 개념을 추가한 DenseNet 기반 시멘틱 세그멘테이션 아키텍처를 사용한다. FCN과 U-net 아키텍처는 여러 번의 컨볼루션과 풀링 과정을 거치



그림 3. 다운샘플링 과정에서 패션 아이템의 특징을 잃은 사례 (FCN 기반 아키텍처)

Fig. 3. Case that lost feature of fashion item in downsampling process (FCN based model)

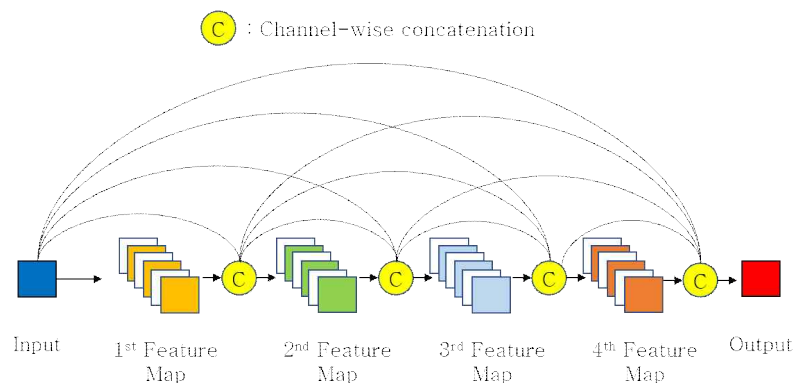


그림 4. 밀집 블록(Dense block)에서의 밀집 연결성 개념

Fig. 4. Dense connectivity concept of dense block

며 굉장히 작아진 이미지 특징 맵의 디테일 보강을 위하여 앞선 단계에서의 컨볼루션 및 풀링 결과를 병합해주는 스킵 계층(Skip layer) 기법을 활용하여 업샘플링을 진행한다. 그러나 이미지 인식 레벨에서의 패션 이미지 데이터는 의류 패턴의 모호한 인식(Ambiguous patterns), 옷감 색깔간의 유사성 판단(Similar appearance), 배경과의 혼선(Disordered background), 높은 그림자 가중치(Heavy shadows) 등의 특성을 지니고 있기 때문에 컨볼루션 및 풀링을 거치며 다운샘플링되었을 때 각 패션 오브젝트의 디테일을 상당 부분 잃게 된다.

시맨틱 세그멘테이션 분야에서는 성능의 지표로 IoU (Inter-section over union)를 사용하게 되는데, 다운샘플링 단계에서 이미 패션 아이템의 특징 정보가 사라지게 되면 업샘플링 단계에서 스킵 계층, 특징 지도 연쇄 메소드 등을 통하여 오브젝트의 디테일을 보강해도 IoU를 충분히 확보하지 못한다는 문제가 발생하게 된다. 이러한 문제의 해결을 위하여 다운샘플링 및 업샘플링 과정에 밀집 연결성(Dense connectivity)이라는 전체 아키텍처의 계층 간 완전한 연결

개념을 추가함으로써 오브젝트의 디테일을 보강하였다.

DenseNet의 저자들은 이전 계층에서 얻어낸 정보를 네트워크가 깊어짐에 따라 계속해서 쌓아나가는 밀집 블록(Dense block)을 아키텍처에 포함함으로써 이전에 지니고 있었던 특징이 사라지거나 희미해지는 문제점을 해결할 수 있다고 여쭈었다. 실험에 사용한 패션 데이터의 경우 전체 클래스가 다양하며 같은 클래스에 속할 지라도 비정형성이 심하여 사이즈나 텍스처, 모양, 패턴 등이 달라 인식에 있어서 모호함이 생기는데, 이러한 밀집 연결성 개념을 사용하면 전체적인 네트워크 내부의 특징 및 연관성이 열은 특징까지 효율적으로 시맨틱 세그멘테이션 결과에 반영할 수 있게 된다. 다른 뉴럴 네트워크 아키텍처와 비교하였을 때, DenseNet 기반 아키텍처는 배치 정규화(Batch normalization)^[9], ReLU^[10] 활성화 함수, 3 x 3 컨볼루션, 0.2 드롭아웃(Dropout)^[11]으로 구성된 하나의 Dense layer에서 정해진 성장률(Growth rate)만큼의 특징 지도를 생성하여 다음 Dense layer로 연쇄시켜 주는데, 이러한 특성상 굉장히 좁은 계층(Very narrow layers)을 가지게 된다. 이러한 특징

표 1. Fully convolutional densenet layers
Table 1. Fully convolutional densenet layers

Dense layer	Architecture of FCDenseNet67(growth rate = 16)	
Batch Normalization	Input	Input, ch = 3(RGB image)
ReLU	First convolution	3 x 3 convolution, ch = 48
3 x 3 Convolution	Downsampling	DB (5 Dense layers) + TD, ch = 128
Dropout = 0.2		DB (5 Dense layers) + TD, ch = 208
		DB (5 Dense layers) + TD, ch = 288
		DB (5 Dense layers) + TD, ch = 368
		DB (5 Dense layers) + TD, ch = 448
	Bottleneck	DB (5 Dense layers), ch = 528
	Upsampling	TU + DB (5 Dense layers), ch = 448
		TU + DB (5 Dense layers), ch = 368
		TU + DB (5 Dense layers), ch = 288
		TU + DB (5 Dense layers), ch = 208
		TU + DB (5 Dense layers), ch = 288
	Fully convolution	1 x 1 convolution, ch = 59 (classes)
	Classification	Softmax

Transition down(TD)
Batch Normalization
ReLU
1 x 1 Convolution
Dropout = 0.2
2 x 2 Max pooling

Transition up(TU)
3 x 3 Transposed Convolution stride = 2

재사용성 덕에 학습 파라미터 개수 및 계산 복잡도를 큰 폭으로 줄이게끔 하고, 적은 양의 데이터 세트에도 쉽게 과적합(Overfitting)^[12]되지 않는 장점을 가진다.

IV. 실험 결과 및 논의

1. 실험 데이터

실험에서는 고해상도 스트리트 패션 이미지에 대하여 59개 클래스로 시멘틱 세그멘테이션 되어 있는 그라운드 트루쓰가 쌍(Pair)으로 포함되어 있는 데이터를 사용한다. 이 데이터는 Clothing Co-Parsing by Joint Image Segmenta-

표 2. 데이터 세트 명세
Table 2. Data Set Description

Data	Description
Quantity	1689 samples, 1007 images with train and 341 images with validation, 341 images with test.
Number of classes	58 (+1 background)
Image size	Resize images with a width of 224, height dependent on original size.
Remarks	We expanded 57 classes of fashionista dataset to 59 categories of ccp dataset.



그림 5. 데이터 세트 예시
Fig. 5. Sample of data set

tion and Label^[4] 연구의 결과물으로써 공개되어 있으며, 논문에서 보여준 CCP, Fashionista 두 가지 데이터 세트를 병합하여 사용하였다. 학습 단계에서는 GPU 메모리 사용량 및 학습 파라미터 수 감소를 위하여 가로 길이 224 픽셀을 기준으로 이미지를 다운스케일링(Downscaling)하여 진행하였으며, 데이터 학습의 효율성을 높이기 위하여 랜덤하게 좌우 대칭(Horizontal flip) 연산을 적용하였다.

2. 실험 결과

제안 방법의 성능을 평가하기 위한 손실 함수(Loss function)로는 범주형 크로스 엔트로피 (Categorical cross-entropy) 역할을 하는 음의 로그 우도 함수(Negative log-likelihood function)를 사용한다. 이는 사용 아키텍처의 마지막단에 있는 Softmax unit을 거쳐 출력된 데이터의 확률 분포와 그라운드 트루쓰의 확률 분포의 차이, 즉 출력 데이터와 그라운드 트루쓰 간의 클래스 손실 값을 나타낸다. 출력 데이터의 픽셀 정확도를 나타내는 평가 지표로는 IoU 정확도 (Accuracy)를 사용한다. 이 값은 그라운드 트루쓰 이미지의 각 픽셀에 라벨링된 클래스 인덱스와 출력 데이터의 동일 차원 픽셀에 라벨링된 클래스 인덱스 간의 교집합 수치를 나타낸다.



그림 6. 시멘틱 세그멘테이션 결과 예시
Fig. 6. Examples of semantic segmentation



그림 7. FCN 기반 아키텍처와의 결과 비교
Fig. 7. Comparison of results with FCN-based architecture



그림 8. 배경과 인물의 영역 분류에는 성공하였으나, 오브젝트의 디테일을 추론하지 못한 사례
Fig. 8. Successful classification of the background and person, but the case of not inferring the detail of the object

표 3. 데이터 세트별 학습 결과

Table 3. Learning results by data set

Data set	NLL loss	IoU accuracy(%)
Train Set	0.490	86.46
Validation Set	0.892	81.93
Test Set	0.868	82.04

표 4. 학습 관련 파라미터

Table 4. Learning parameters

Hyperparameter	Value
Optimizer	RMSprop
Loss function	NLLoss(Negative log likelihood loss) for categorical classification
Learning rate	1e-3
Learning rate decay	0.995 (every 1 epoch)
Mini batch size	1
Epoch	900
Weight initialization	Initialize to zero

$$-\sum_{i=1}^{n=59} o_i \log(g_i) \quad (1)$$

전체 클래스 $n=59$ 에 대하여 Softmax를 거쳐서 나온 출력(Output) 데이터의 1번째 확률분포값과 같은 공간에 대한 그라운드 트루쓰(Ground truth)의 1번째 확률분포값에 대한 NLL loss.

$$IoU(c) = \frac{\sum_i (o_i \equiv c \wedge g_i \equiv c)}{\sum_i (o_i \equiv c \vee g_i \equiv c)} \quad (2)$$

주어진 클래스 c 에서 출력(Output) 데이터의 1번째 픽셀과 그라운드 트루쓰(Ground truth)의 1번째 픽셀에 대한 IoU.

테스트 데이터에 대한 IoU 정확도는 82.04%이며 동일 아키텍처에 대한 CamVid^[13] 데이터 세트(11개 클래스 +1 배경)의 결과 값(86.80%)보다는 수치가 낮지만, 패션 이미지 특유의 비정형성과 59 클래스에 대한 영역 추론의 복잡성을 고려한다면 이는 충분히 의미 있는 결과로 볼 수 있다. 또한 음의 로그 우도 손실 함수의 값이 충분히 최적화되지 않았는데, 이는 밀집 연결성, 스킵 계층 등 특징 재사용만으로는 해결이 힘든 패션 이미지 데이터의 강한 비정형성으로 인한 것으로 생각된다.

V. 결 론

본 논문에서는 밀집 연결성 개념을 포함한 DenseNet 기반 아키텍처로 다양한 오브젝트 클래스가 포함된 패션 이미지 데이터에 대한 실험을 수행하고 그 결과에 대하여 논의하였다. 관련 선행 연구에서는 CNN 기반의 효율적인 모델을 제시한 FCN과 시멘틱 세그멘테이션의 기본적인 프로세스를 정립한 U-net을 언급하고 각 방식의 차이점을 알아보았다. 실험에는 [4] 연구에서 공개한 픽셀 레벨 라벨링 데이터 1,689건을 활용하였으며, 밀집 연결성 개념을 통한 이전 연구와의 차이점 및 개선점을 확인할 수 있었다. 평가 지표로 활용한 음의 로그 우도 손실 값은 적절한 최적화 단계까지 도달하지는 못하였으나, 비정형성이 심하고 클래스가 다양한 패션 이미지 데이터의 효율적 영역 추론에 대한 가능성을 제시하였다. 추후 CRF (Conditional random fields)^[14]와 같은 후처리 기법과 더 많은 학습 데이터를 확보하여 최적화를 실시함으로써 성능 향상이 가능할 것으로 보인다. 본 연구 및 여러 기업에서 진행 중인 동일 주제의 연구를 기반으로 현대 패션 상거래 서비스의 패션 오브젝트 클래스 태깅에 대한 개선점을 찾을 수 있을 것으로 예상된다.

참 고 문 헌 (References)

- [1] M. Thoma, "A survey of semantic segmentation," CoRR, vol. abs/1602.06541, 2016, [Online]. Available: <http://arxiv.org/abs/1602.06541>
- [2] Krizhevsky, A., Sutskever, I., and Hinton, G. E. "ImageNet classification with deep convolutional neural net-works." In NIPS, 2012, pp. 1106 - 1114
- [3] J. Long, E. Shelhamer, and T. Darrell. "Fully convolutional networks for semantic segmentation." In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 3431-3440
- [4] X. Liang, L. Lin, W. Yang, P. Luo, J. Huang, and S. Yan. "Clothes co-parsing via joint image segmentation and labeling with application to clothing retrieval." In IEEE Transactions on Multimedia, 2016. pp. 1175-1186
- [5] G. Huang, Z. Liu, K. Q. Weinberger, and L. van der Maaten. "Densely connected convolutional networks." In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.
- [6] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," In Navab N., Hornegger J., Wells W., Frangi A. (eds) Medical Image Computing and Computer-

- Assisted Intervention - MICCAI 2015. Lecture Notes in Computer Science, vol 9351. Springer, Cham
- [7] Simon Jégou, M. Drozdal, D. Vazquez, et al. "The one hundred layers tiramisu: Fully convolutional densenets for semantic segmentation." In 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW).
- [8] V. Iglovikov and A. Shvets. "Ternausnet: U-net with vgg11 encoder pre-trained on imagenet for image segmentation." arXiv preprint arXiv:1801.05746, 2018.
- [9] S. Ioffe and C. Szegedy. "Batch normalization: Accelerating deep network training by reducing internal covariate shift." In ICML, 2015, pp. 446-456
- [10] Nair, V. and Hinton, G. E. (2010). "Rectified linear units improve restricted boltzmann machines." In Proc. 27th International Conference on Machine Learning, 2010
- [11] N. Srivastava, G. E. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov. "Dropout: a simple way to prevent neural networks from overfitting." In JMLR, 2014. pp. 1929-1958
- [12] R. Caruana, S. Lawrence, and L. Giles, "Overfitting in neural nets: Backpropagation, gradient, and early stopping," In Advances in Neural Information Processing Systems, 2001, vol. 13, pp. 402 - 408.
- [13] J. Brostow, J. Fauqueur, and R. Cipolla. "Semantic object classes in video: A high-definition ground truth database." In Pattern Recognition Letters, 2009, vol. 30, Issue 2, pp. 88 - 97
- [14] J. Lafferty, A. McCallum, and F. Pereira. "Conditional random fields: Probabilistic models for segmenting and labeling sequence data." In Proceedings of the 18th International Conference on Machine Learning 2001 (ICML 2001), pp. 282-289

저 자 소 개

임 현 영



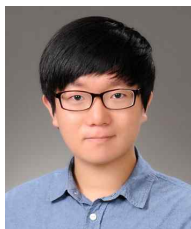
- 2017년 : 세종대학교 디지털콘텐츠학과 학사
- 2017년 ~ 현재 : 세종대학교 디지털콘텐츠학과 석사과정
- ORCID : <https://orcid.org/0000-0002-8547-6248>
- 주관심분야 : 컴퓨터 비전, 기계학습, 딥러닝

이 유 림



- 2018년 : 세종대학교 디지털콘텐츠학과 학사
- 2018년 ~ 현재 : 세종대학교 인공지능언어공학과 석사과정
- ORCID : <https://orcid.org/0000-0001-8309-090X>
- 주관심분야 : 텍스트 마이닝, 자연어 처리, 딥러닝

지 민 규



- 2018년 : 세종대학교 천문우주학과 학사
- 2018년 ~ 현재 : 세종대학교 소프트웨어융합학과 석사과정
- ORCID : <https://orcid.org/0000-0002-3089-1452>
- 주관심분야 : 텍스트 마이닝, 기계학습, 딥러닝

저 자 소 개



고 명 현

- 2016년 : 세종대학교 디지털콘텐츠학과 학사
- 2016년 ~ 현재 : 세종대학교 디지털콘텐츠학과 석사과정
- ORCID : <https://orcid.org/0000-0002-6036-4717>
- 주관심분야 : 텍스트 마이닝, 기계학습, 딥러닝



김 학 동

- 2016년 : 경상대학교 컴퓨터공학과 학사
- 2017년 ~ 현재 : 세종대학교 디지털콘텐츠학과 석, 박사통합과정
- ORCID : <https://orcid.org/0000-0003-3816-1224>
- 주관심분야 : 머신러닝, 딥러닝, 자연어처리



김 원 일

- 1981년 12월 ~ 1985년 7월 : ㈜대한항공 전산실 재무 시스템 개발원
- 1982년 : 한양대학교 공과대학 금속공학 학사
- 1987년 : 미국 일리노이주 서던일리노이대학교 컴퓨터 과학 학사
- 1990년 : 미국 일리노이주 서던일리노이대학교 컴퓨터 과학 석사
- 1994년 : 미국 인디애나주 인디애나 대학교 대학원 컴퓨터 과학 전공
- 2000년 : 미국 뉴욕주 시러큐스 대학교 대학원 컴퓨터 & 정보과학 공학 박사
- 2000년 1월 ~ 2001년 3월 : 미국 펜실베이니아주 외인시 소재 Bhasha, INC Technical Staff (연구원)
- 2002년 3월 ~ 2003년 8월 : 아주대학교 정보통신전문대학원 BK 교수
- 2003년 9월 ~ 2017년 2월 : 세종대학교 전자정보공학대학 교수
- 2017년 3월 ~ 현재 : 세종대학교 소프트웨어융합대학 교수
- ORCID : <https://orcid.org/0000-0002-1489-8427>
- 주관심분야 : 인공지능, 지능형 시스템, 딥러닝 등