

회색조 영상의 시맨틱 분할을 위한 사전 학습된 딥러닝 모델 확장 기법

정진원*, 신요안^o

An Extension of Pre-Trained Deep Learning Model for Sematic Segmentation of Gray-Scale Images

Jin-won Jung*, Yoan Shin^o

요약

데이터 양이 적은 데이터 세트로 딥러닝 모델을 학습할 때, 전이학습을 통해 업데이트와 재학습을 활용하면 학습 시간과 연산 리소스를 줄이고 알고리즘의 성능을 크게 향상할 수 있다. 하지만 영상처리에서 대부분의 사전 학습된 모델은 컬러 영상을 사용하여 학습되어 있어 3개 색상 채널이 입력 영상 값으로 사용된다. 반면, 회색조 영상은 컬러 영상에 비해 크기가 작고 1개의 채널을 가져 여러 개의 채널을 학습한 모델의 입력으로 사용할 수 없다. 따라서 기존 기법들은 데이터를 입력에 맞게 변형시키거나 사전 학습된 모델의 레이어를 수정해야 한다는 한계가 있다. 본 논문에서는 시맨틱 분할 영상처리를 위한 딥러닝 모델 앞부분에 컨볼루션 레이어를 추가해 회색조 영상에서 사전 학습된 모델을 사용하는 새로운 방법을 제안한다. 모의실험 결과, 제안 기법의 정확도가 높았으며 기존 기법보다 우수한 성능을 얻으면서도 좀 더 효과적으로 사전 학습된 모델을 사용할 수 있음을 확인하였다.

Key Words : image processing, semantic segmentation, RGB 3-channel color

image, gray-scale image, transfer learning, pre-trained model, InceptionResNetV2, convolution layer

ABSTRACT

When learning deep learning models on datasets with small amounts of data, utilizing updates and re-learning through transfer learning can reduce learning time and computational resources, and significantly improve the performance of the algorithm. However, most pre-trained models in image processing applications are trained using color images, so three color channels are used as input image values. On the other hand, the gray-scale image is smaller in size than the color image and has one channel, so it cannot be used as an input of a model that has learned several channels. Consequently, existing techniques have limitations in that they need to transform the data to fit the input or modify the layer of the pre-trained model. In this paper, we propose a novel method to use a pre-trained model for semantic segmentation of gray-scale images by adding convolutional layers to the front of a deep learning model. Simulation results show that the proposed scheme has higher accuracy than the existing scheme, and that pre-trained models can be used effectively with good performance.

1. 서론

전이학습 (Transfer Learning)은 대용량 데이터 세트에서 학습된 모델을 이용해 데이터 양이 적은 데이터 세트의 일반화 성능을 개선하고 딥러닝 모델 (또는 인공지능망)의 성능을 향상하는 기계학습 기술이다^[1]. 전이학습에서는 구축한 사전 학습된 딥러닝 모델이 핵심이라 할 수 있는데, 영상처리에서 이러한 모델은 일반적으로 굉장히 많은 수 (예를 들어, 수백만 개)의 샘플 학습 영상을 통해 학습된다. 하지만 사전 학습된

* 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터지원사업의 연구결과로 수행되었습니다. (IITP-2022-2018-0-01424)

• First Author : (ORCID:0000-0003-3026-0594) School of Electronic Engineering, Soongsil University, jinwonj@soongsil.ac.kr, 학생 (석사), 학생회원

o Corresponding Author : (ORCID:0000-0002-4722-6387) School of Electronic Engineering, Soongsil University, yashin@ssu.ac.kr, 정교수, 종신회원

논문번호 : 202211-285-A-LU, Received November 28, 2022; Revised December 9, 2022; Accepted December 9, 2022

딥러닝 모델은 학습에 사용된 원본 영상의 모양과 구조에 맞게 사전 처리해야 한다. 따라서 RGB 3-채널 구조의 컬러 영상이 학습된 모델에 회색조 (Gray-scale) 영상을 적용하기 위해서는, 회색조 영상을 3-채널 유사 컬러 영상으로 쌓아 만들어야 하나 이렇게 쌓인 3-채널 회색조 영상에는 색상 정보가 없음에도 불구하고 데이터가 불필요하게 많아지게 된다.

본 논문에서는 컬러 영상으로 사전 학습된 딥러닝 모델 앞부분에 컨볼루션 레이어 (Convolution Layer)를 추가하여, 전처리 과정과 데이터 증가 없이 회색조 영상의 전이학습에 사용할 수 있는 새로운 방법을 제안한다.

II. 컨볼루션 레이어를 활용한 모델 적용 기법

그림 1은 컨볼루션 레이어를 추가하여 확장된 제안 딥러닝 모델의 구성이다. 제안된 모델은 전형적인 영상처리용 딥러닝 모델 (본 연구에서는 InceptionResNetV2^[2])에, 3개 채널로 변환하는 기능을 수행하는 역할을 가진 새로운 컨볼루션 레이어가 추가되어 확장이 일어난다. 이 추가되는 새로운 레이어는 크기가 3x3인 3개 커널을 가진 두 개 컨볼루션2D 레이어가 쌓여 이루어져 있다. 컨볼루션2D 연산은 커널을 입력 벡터 상에서 움직여 선형 모델과 합성함수가 적용되는 구조로 아래 식과 같이 정의된다^[3].

$$y[i,j] = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} h[m,n] \cdot x[i-m,j-n], (1)$$

여기서 x , y 는 각각 입력 이미지와 출력 이미지를 나타내며, 인덱스 i , j 와 m , n 은 각각 해당 이미지의 행렬과 커널의 행렬을 나타낸다. 해당 층에 회색조

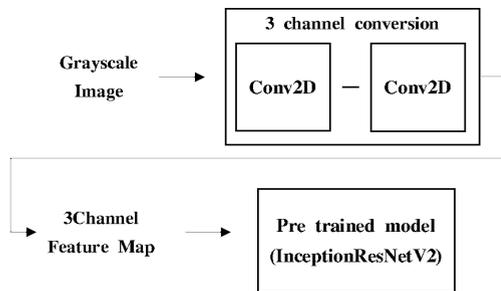


그림 1. 컨볼루션 레이어를 추가하여 확장된 제안 딥러닝 모델.
Fig. 1. Proposed deep learning model extended by adding convolutional layers.

영상이 입력되면, 컨볼루션 레이어가 영상의 채널을 맞춰주는 기능을 수행하게 된다.

회색조 영상은 그림 1과 같이 컨볼루션 레이어를 거치면서 가로 및 세로 방향으로 인접한 픽셀들에 대한 필터를 학습하게 되며, 3개 채널로 변환된다. 이후 나온 3개 채널의 Feature Map을 사전 훈련된 딥러닝 모델 (InceptionResNetV2)에 입력값으로 사용하여 학습한다. 사전 훈련된 InceptionResNetV2 모델은 ImageNet의 100만개가 넘는 영상에 대해 훈련된 인공신경망으로 164개의 계층으로 이루어져 있는 구조이다^[2].

III. 시맨틱 분할 모의실험 결과

본 연구에서는 제안 기법을 영상 처리의 주요 연구 분야 가운데 하나인 시맨틱 분할 (Semantic Segmentation)에 적용한다. 시맨틱 분할은 픽셀 수준 분류라고도 하며, 동일한 객체 클래스에 속하는 영상의 일부를 함께 클러스터링 하는 작업이다. 시맨틱 분할은 도로 표지판 검출, 토지 표지 분류, 종양 검출 등 다양한 응용 분야에 적용되고 있다^[4].

모의실험을 위한 데이터 세트는 MBRSC 위성에서 얻은 두바이의 컬러 위성 영상으로 구성되어 있으며, 256x256의 크기로 총 1,136개의 영상을 전처리하여 데이터 세트로 사용하였다^[5]. 본 논문에서 제안하는 모델 적용 기법은 Python 3.8.13과 TensorFlow를 사용하여 구현되었다. 모델 훈련 및 알고리즘 실행은 NVIDIA GeForce RTX 3070과 Intel i7-10700 2.90 GHz CPU, 8GB RAM의 실험 환경에서 수행하였다.

기존 기법의 전처리 과정에서 회색조 영상을 3개 채널로 변형시킨 것과는 달리, 제안 기법은 모델을 확장시켜 모델 안에서 회색조 영상을 3개 채널로 변형시킨다. 때문에 기존 기법은 입력으로 전처리된 3개 채널의 회색조 영상이 사용되는 반면, 제안 기법은 1개 채널의 회색조 영상이 사용된다. 모의실험에서는 제안 기법과 기존 기법의 성능을 비교하였으며, 모델 내의 파라미터는 전부 동일하게 적용하였다.

그림 2는 훈련 중 Epoch에 따른 성능 평가 지표의 변화를 나타낸 결과이다. 기존 기법은 학습 중 변화 폭이 크지 않았지만, 제안 기법은 동일 Epoch에서의 변화가 크고 민감하게 나타났다. 여기서, 성능 평가 지표는 정확도 (Accuracy), 손실 (Loss) 및 아래 식으로 정의되는 Dice Coefficient를 사용하였다^[6].

$$Dice = 2 \cdot |X \cap Y| / (|X| + |Y|), (1)$$

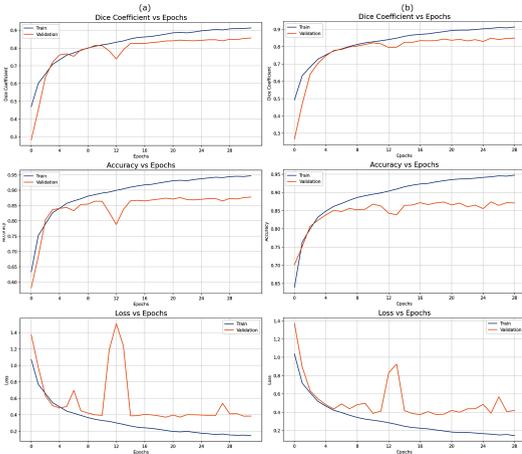


그림 2. 훈련 중 Epoch에 따른 성능 평가 지표 변화. (a) 제안 기법, (b) 기존 기법.
 Fig. 2. Performance evaluation metrics according to training epochs. (a) Proposed scheme, (b) conventional scheme.

위 식에서 X 는 정답 픽셀, Y 는 예측 픽셀로서, 분모에는 정답 픽셀과 예측 픽셀의 합, 분자에는 정답 픽셀과 예측 픽셀의 교집합 부분을 2배한 값을 나타내어, 두 영역이 정확히 같다면 1이 되며 그렇지 않을 경우에는 0이 된다.

그림 3은 시맨틱 분할 모의실험의 실제 결과 영상들을 도시한다. 여기서, 제안 기법이 기존 기법보다 작은 물체를 좀 더 잘 식별하였으며, 시각적으로 정답 마스크에 훨씬 더 유사한 결과를 낸 것을 알 수 있다.

한편, 표 1은 다양한 성능 평가 지표에 대한 제안 기법과 기존 기법의 구체적인 결과를 비교하여 정리한다. 여기서, 제안 기법의 정확도와 Dice Coefficient

표 1. 성능 평가 지표 비교 결과.

Table 1. Comparison of performance evaluation metric results.

	Accuracy	Dice Coeff.	Loss
제안 기법	0.875	0.841	0.368
기존 기법	0.868	0.832	0.388

가 기존 기법에 비해 더욱 높게 나온 것을 알 수 있으며, 손실 역시 제안 기법이 더 좋은 걸 알 수 있다. 이것은 전처리 과정에서 회색조 영상을 3개 채널로 변형시킨 기존 기법에 비해, 제안 기법이 사전 학습된 모델을 효과적으로 사용할 수 있음을 나타낸다.

IV. 결 론

본 논문에서는 시맨틱 분할 영상처리에서 회색조 영상을 사전 학습된 모델에 적용하기 위해, 컨볼루션 레이어를 추가해 딥러닝 모델을 확장하는 기법을 제안하였다. 전처리 과정에서 이미지를 3개 채널로 변형시킨 기존 기법과 모델을 확장시켜 모델 안에서 회색조 영상을 3개 채널로 변형시킨 제안 기법을 비교해, 모의실험 결과 본 제안 기법이 시각적으로 정답 마스크에 더 유사하였으며, 기존 기법과 비교해 전체적으로 여러 성능 평가 지표에서 우수한 결과를 얻을 수 있었다. 따라서, 제안 기법이 기존 기법보다 우수한 성능을 얻으면서도 동시에 사전 학습된 모델을 더욱 효과적으로 사용할 수 있음을 확인하였다.

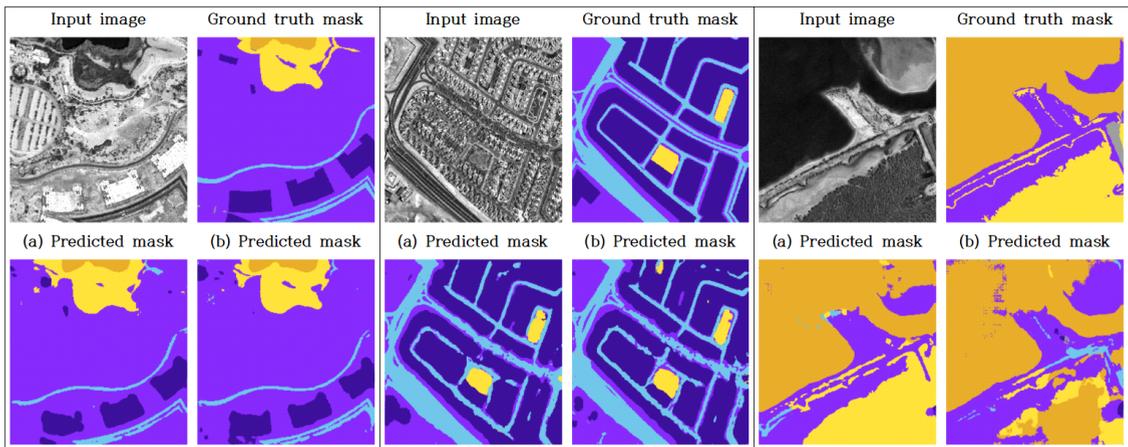


그림 3. 회색조 영상의 시맨틱 분할 결과. (a) 제안 기법, (b) 기존 기법.
 Fig. 3. Semantic segmentation results of gray-scale images. (a) Proposed scheme, (b) conventional scheme.

References

- [1] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, and Q. He, "A comprehensive survey on transfer learning," in *Proc. IEEE*, vol. 109, no. 1, pp. 43-76, Jan. 2021.
(<https://doi.org/10.1109/JPROC.2020.3004555>)
- [2] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, "Inception-v4, Inception-ResNet and the impact of residual connections on learning," in *Proc. AAAI 2017*, pp. 4278-4284, San Francisco, USA, Feb. 2017.
(<https://doi.org/10.1609/aaai.v31i1.11231>)
- [3] J.-H. Choi, "Parallel-addition convolution algorithm in grayscale image," *J. Korea Inst. Inf. Electr. & Commun. Technol.*, vol. 10, no. 4, pp. 288-294, Oct. 2017.
(<https://doi.org/10.17661/jkiiect.2017.10.4.288>)
- [4] X. Liu, Z. Deng, and Y. Yang, "Recent progress in semantic image segmentation," *Artificial Intell. Rev.*, vol. 52, pp. 1089-1106, Aug. 2019.
(<https://doi.org/10.1007/s10462-018-9641-3>)
- [5] <https://humansintheloop.org/resources/datasets/semantic-segmentation-dataset-2/>
- [6] L. R. Dice, "Measures of the amount of ecologic association between species," *J. Ecology*, vol. 26, no. 3, pp. 297-302, Jul. 1945.
(<https://doi.org/10.00002307/1932409>)