

딥러닝을 이용한 도용, 원본 이미지의 유사도 측정 및 응용 시스템 개발

조준수¹ · 김승원^{2*}¹전남대학교 인공지능융합학과 석사과정²전남대학교 인공지능학부 조교수

Development of Measure Similarity System between Stolen Image and Original Image Based on Deep Learning

Jun-Su Cho¹ · Seung-Won Kim^{2*}¹Master's Course, Department of AI Convergence, Chonnam National University, Gwangju 61186, Korea²Assistant Professor, Department of Artificial Intelligence, Chonnam National University, Gwangju 61186, Korea

[요약]

본 연구에서는 이커머스 사이트에서 상품 이미지를 도용하는 문제를 해결하고자, 딥러닝을 이용한 이미지 분류 웹 사이트를 소개한다. 최근 이커머스 시장이 확대됨에 따라 상품 이미지를 일부 수정하여 다른 상업 사이트에 도용하는 문제가 빈번히 발생한다. 하지만, 많은 이커머스 사이트와 많은 상품 이미지로 사람이 직접 도용 이미지를 찾아내는 것은 시간과 비용이 많이 든다. 이러한 문제를 해결하기 위해 Siamese Network 기반 딥러닝 모델을 구현하고 도용 이미지 유사도 측정 방법 및 응용 시스템을 개발하였다. 모델의 Top-k 정확도 측정 결과 Top-1, Top-5, Top-10 정확도에서 65.0%, 88.6%, 94.3%의 성능을 보여주었다. 구현된 응용 시스템은 입력된 도용 의심 이미지와 유사도가 높은 10개의 원본 이미지를 나열하였으며 이는 사용자가 94.3%의 확률로 도용 이미지를 확인할 수 있음을 의미한다. 결론적으로 사용자는 유사도가 높은 10개의 원본 이미지를 도용 의심 이미지와 비교 확인함으로써 쉽고 빠르게 도용 이미지임을 확인할 수 있다.

[Abstract]

This study introduces an application that classifies images according to the level of similarity between original and stolen images on e-commerce platforms. Recently, as e-commerce platforms have gained popularity, the theft of product images and their unauthorized use on other platforms on others has become a critical infringement of intellectual property rights. However, it is costly and time-consuming for a worker to seek stolen images from thousands of E-commerce platforms. To address this issue, we propose a website application that measures the similarity between original and stolen images using deep learning technology. The deep learning model is based on a Siamese Network, and the training results of the model show 65.0%, 88.6%, and 94.3% accuracy rates in Top-1, Top-5, and Top-10 measurements, respectively. Our application displays the similarity level with the input image, listing ten original images are listed in a high-order format. This enables users to effectively identify the stolen images by comparing the suspicious image with the ten most similar original images.

색인어 : 딥러닝, 도용 이미지, 원샷 학습법, 데이터 증강, 어플리케이션**Keyword** : Deep Learning, Stolen Image, One-Shot Learning, Data Augmentation, Application<http://dx.doi.org/10.9728/dcs.2024.25.8.2301>

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 16 July 2024; Revised 05 August 2024

Accepted 09 August 2024

***Corresponding Author, Seung-Won Kim**

Tel: +82-62-530-4231

E-mail: seungwon.kim@chonnam.ac.kr

I. 서론

이커머스는 인터넷 네트워크를 기반으로 판매자와 구매자가 상품을 사고 팔 수 있는 공간을 의미한다. 판매자는 구매자의 구매의사 결정을 돕기 위하여 상품의 이미지나 정보 등을 인터넷 상에 공유하고 사용자는 업로드 된 이미지와 정보를 기반으로 상품을 구매할지 결정한다. 하지만 인터넷의 특성상 상품 이미지와 같은 지적 재산권이 포함된 자료를 도용자가 쉽게 다운로드하여 사용할 수 있으며 원본 이미지의 회전 및 추출하는 등의 간단한 수정을 거쳐 사용할 수 있다. 이러한 도용은 원본 이미지 제작자의 지적 재산권을 침해하는 행동이다. 하지만, 이러한 이미지 도용 문제는 인터넷에서 대규모로 일어나며 인력을 통한 해결은 많은 시간과 비용이 드는 문제점이 있다.

딥러닝[1]은 사람의 신경망 구조와 유사한 모델을 생성하여 객체 인식[2],[3], 행동 인식[4],[5], 질병 감지[6],[7] 등과 같은 이미지 인식이 필요한 작업을 효과적으로 수행할 수 있다. 즉, 작업자의 반복적인 이미지 인식 작업을 딥러닝 모델이 대체할 수 있으며 작업자 고용에 비해 적은 비용으로 수행할 수 있다. 딥러닝의 비전 기반 모델로 방대한 양의 도용 이미지를 인식하고 감지하여 시간과 인력을 절약할 수 있다.

본 논문에서는 도용 이미지 추출을 위한 딥러닝 모델 학습 방법과 쉽고 빠르게 이미지 도용을 판단할 수 있는 응용 시스템을 소개한다. 응용 시스템은 사용자가 입력한 도용 의심 이미지와 유사도가 높은 순으로 원본 이미지를 나열한다. Top-k 정확도 측정 결과를 바탕으로 응용 시스템에 표시되는 원본 이미지 수는 10개로 설정되었으며 작업자는 나열된 10개의 원본 이미지를 확인하여 최종적으로 입력된 이미지가 도용된 이미지임을 판단할 수 있다. 이러한 방법은, 도용 의심 이미지와 많은 원본 이미지를 비교하는 시간과 비용을 절약할 수 있게 한다.

본 논문의 구성은 다음과 같다. 2장에서는 기존 이미지 간 유사도 측정에 사용했던 모델과 원샷 학습법에 대한 기본적인 개념을 기술한다. 3장에서는 원본 이미지와 도용 이미지 간 유사도 측정을 위한 데이터셋 구성과 학습 과정 및 실험 결과를 기술한다. 또한 도용 의심 이미지를 이용한 유사도 측정 응용 시스템을 소개한다. 4장에서는 최종 결론과 향후 연구에 대하여 기술한다.

II. 관련 연구 분석

2-1 이미지 간 유사도 측정 모델

딥러닝을 이용한 유사도 측정 모델은 사물 분류[8], 서명 확인[9], 얼굴 인식[10],[11] 등과 같이 여러 분야에 활용되어왔다.

FaceNet[11]은 얼굴 이미지를 이용하여 모델을 학습시키고 다른 배경 및 각도에서 동일 인물임을 구별하는 모델이다. 기준(Anchor), 동일(Positive), 무관련(Negative) 얼굴 이미지 3개를 입력받아 학습에 사용한다. 각 얼굴 이미지는 128 차원의 임베딩 벡터로 변환되며 얼굴의 유사도를 측정하는데 사용된다. 3개의 이미지 임베딩 벡터를 바탕으로 기준과 동일 이미지, 동일과 무관련 이미지, 무관련과 기준 이미지간 상대적 거리를 L2 Distance로 측정한다. 측정된 L2 Distance를 이용하여 유클리드 거리를 계산하고 Loss로 사용한다. 학습 과정에서 기준과 동일 이미지의 유클리드 거리는 가까워지고 기준과 무관련 이미지의 유클리드 거리는 멀어진다. 유클리드 거리가 가까울수록 유사도가 높다고 판단한다.

Siamese Network[8]는 두 이미지의 특징점 차이를 이용하여 물체나 글자의 유사도를 측정한다. 모델은 두 개의 딥러닝 네트워크로 구성되어 있으며 입력 데이터로 두 개의 이미지를 받는다. 각 이미지는 네트워크의 합성곱 계층과 풀링 계층을 통과하여 특징점을 추출한다. 추출된 두 이미지의 특징점은 L1 distance, L2 distance 등을 이용하여 코사인 유사도를 측정하고 Loss로 사용한다. 입력 데이터로 사용되는 두 개의 이미지 중 하나는 기준 이미지, 하나는 비교 이미지로 정의된다. 기준 이미지와 비교 이미지가 동일한 클래스일 경우 True로 분류되고 다른 클래스일 경우 False로 분류된다. 훈련 에포크 중 절반은 True가 되는 비교 이미지 쌍을 사용하고 나머지 절반은 False가 되는 비교 이미지를 이용하여 모델의 유사도 측정 능력을 향상시킨다. 기존 연구는 기준 이미지와 환경이나 각도만 다른 이미지를 같은 클래스로 정의하고 모델을 학습하였다. 하지만 도용 이미지의 경우 기준 이미지의 일부만을 추출하여 사용하는 경우가 많다. 본 연구에서는 도용 이미지 판별에 적합한 모델을 학습시키기 위하여 기준 이미지의 일부만을 추출한 이미지를 같은 클래스로 정의하고 모델을 학습시킨다.

2-2 단일 이미지 클래스를 이용한 학습

이미지 분류 문제를 학습하는 모델은 각 클래스당 기준이 되는 이미지의 개수에 따라 One-Shot 학습법[13],[14]과 Few-Shot 학습법[12]으로 나눌 수 있다. Few-Shot 학습법[12]은 복수의 기준 이미지를 가질 때 정의되는 학습법이다. 동물 분류 문제의 경우 Few-Shot 학습법은 동물의 종류가 클래스로 정의되며 하나의 동물 사진을 입력받아 각 클래스의 기준 동물 사진들을 이용하여 어느 클래스에 속한 동물 사진인지 구별한다. 이때 비교에 사용된 클래스의 총 개수를 way, 기준 이미지의 개수는 shot으로 정의된다. 하지만 Few-Shot 학습법과 다르게 사람은 한 장의 사진으로 동종 동물인지 확인할 수 있다. 이와 같이 사람의 분류 방식을 참고하여 하나의 기준 이미지로 일치, 불일치를 분류하는 방식을 One-Shot 학습법[13],[14]이라고 한다. 이러한 One-Shot 학습법[13],[14]은 복수의 기준 이미지가 주어지지 않을 때 하나의

기준 이미지로도 분류를 할 수 있다는 장점이 있다. 따라서, 본 연구에서는 One-Shot 학습법을 이용하여 학습 모델을 개발하였다.

III. 설계 및 실험결과

3-1 데이터셋 구성

본 연구에서는 AI-Hub에서 제공하는 K-Fasion 이미지[15]를 기반으로 데이터셋을 구축하였으며 그림 1은 데이터 셋 이미지의 예시를 보여준다. 데이터셋은 여성 의류 이미지로 총 2000개 이미지를 포함하며 이를 원본 이미지로 사용하였다. 각 원본 이미지는 클래스를 가지며 같은 상품이라도 다른 각도의 사진일 경우 다른 클래스를 가지게 하였다. 이는 동일 제품을 다른 각도로 찍으면 동일 제품을 판매하는 다른 매매자가 새로운 이미지를 생성한 것으로 판단하고 도용 이미지가 아니라 결정된 것이다.



그림 1. 여성 의류 이미지
Fig. 1. Women's clothing images



그림 2. 도용 이미지 생성 예시
Fig. 2. Example of creating a stolen image

훈련을 위해 2000개의 원본 이미지를 바탕으로 도용 이미지를 생성하였다. 본 연구에서 도용 이미지는 원본 이미지의 전부 및 일부분에서 추출된 이미지로 정의하고 정의에 따라 도용 이미지를 생성하였다. 이 정의는 이커머스를 분석해본 결과 대부분의 도용 이미지가 원본 이미지의 전부 및 일부를 2~3등분으로 잘라 추출한 것으로 확인되어 결정하였으며 그

림 2는 정의된 도용 이미지의 예시를 보여준다. 도용 이미지는 아래와 같은 수식 (1), (2)를 통해 원본 이미지에서 잘려진 도용 이미지를 정의하고 데이터를 증강하였다.

$$min\ height = \int_a^b \frac{(o - size(o, p))x}{(t-1)} dx \quad (1)$$

식 (1)에서 *min height*는 도용 이미지 생성을 위해 원본 이미지가 잘리는 시작 지점을 의미한다. *size(o, p)*는 도용 이미지의 크기를 의미하며 *o*는 원본 이미지의 세로 크기, *p*는 원본 이미지의 세로 크기 비율이다. 본 실험에서는 원본 이미지의 절반 크기를 도용 이미지로 사용하기 위해 *p* = 2를 적용하였다. 또한 *t*는 생성한 도용 이미지 개수로 본 연구에서는 *t* = 3을 적용하였다.

$$max\ height = min\ height + size(o, p) \quad (2)$$

식 (2)에서 *max height*는 도용 이미지 생성을 위해 잘리는 원본 이미지의 끝 지점을 의미한다. 도용 이미지의 가로 크기는 원본 이미지와 동일하며 수식 (1), (2)를 통해 잘려진 이미지는 원본 이미지와 같은 클래스를 가진다.

또한, 데이터 셋에 라벨 값을 추가하였다. 원본과 도용 이미지의 관계는 라벨값 1(True)로 정의하였으며 원본과 도용 이미지가 아닌 관계는 라벨값 0(False)로 정의하여 라벨 값을 추가하였다. 그림 3은 모델 훈련 시 입력 값으로 들어가는 이미지들과 라벨 값을 보여준다.



그림 3. 이미지 구성에 따른 라벨링 예시
Fig. 3. Example labeling based on image composition

3-2 학습 과정

모델 훈련을 위해 2000장의 데이터 중 1700개는 훈련데이터, 300개는 테스트 데이터로 사용하였다.

딥러닝 모델은 Siamese Network[8]를 사용하였다. 입력된 이미지는 Siamese Network에서 1-64-128-256 Conv 레이어를 거쳐 특징점을 추출하며 각 레이어는 Relu 활성화 함수[16]를 사용한다. 두 이미지의 특징점은 fully connected 와 sigmoid를 거쳐 차이값을 통해 L1 distance를 구한다. 결과값은 fully connected와 sigmoid를 한번 더 거치며 최종

L1 distance 값으로 이미지간 유사도 값을 추출하였다. 그림 4는 전체 학습 과정을 보여준다.

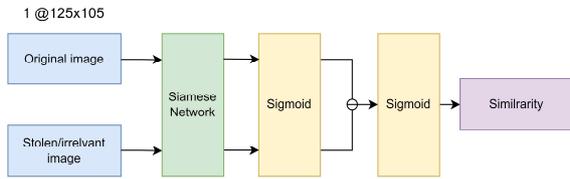


그림 4. 전체 모델 훈련 과정
Fig. 4. Model training process

3-3 모델 성능 측정

데이터셋에 적합한 옵티마이저를 찾기 위하여 기존 Siamese Network[8]에서 사용되었던 옵티마이저 Adam (Adaptive moment estimation)[17]과 SGD(Stochastic Gradient Descent)[18]를 중심으로 모델의 성능을 측정하였다.

성능 측정 테스트는 도용 이미지 1개를 101개의 이미지와 각각 비교하여 측정하였다. 101개의 이미지 중 원본 이미지는 1개 포함 되었다. 모델의 성능 평가는 도용 이미지 판별 성능을 나타내는 정확도(Accuracy), 유사도 측정 능력을 나타내는 원본과 도용 이미지 간 평균 유사도(Similarity)로 나뉜다.

식 (3)에서 정확도(Accuracy)는 테스트 과정에서 모델이 정답을 맞춘 비율을 의미한다. $samples$ 는 테스트셋의 이미지 개수를 의미하며 TN_k 은 도용 이미지를 맞춘 개수를 의미한다.

$$Accuracy = \frac{TN_k}{samples} \tag{3}$$

식 (4)는 Top-k 정확도의 수식을 나타낸다. Top-k 정확도는 입력된 이미지와 유사도가 높은 이미지를 k개 추출하고 추출된 이미지가 입력된 이미지와 동일 클래스일 정확도를 나타낸다. 예를 들어 $k=1$ 의 경우 입력된 도용 이미지와 유사도가 가장 높은 1장의 이미지를 추출하고 그 추출된 이미지가 도용 이미지의 원본 이미지일 확률(정확도)을 나타낸다.

$$TN_k = \sum_{i=0}^{samples-1} \sum_{j=1}^k 1(x_{i,j} = y_i) \tag{4}$$

정확도(Accuracy) 측정 결과는 표 1과 같다. Adam 옵티마이저의 경우 Top-1, Top-5, Top-10, Top-20 순으로 65.0%, 88.6%, 94.3%, 97.3%의 정확도를 보여주었다. SGD 옵티마이저의 경우 Top-1, Top-5, Top-10, Top-20 순으로 63.5%, 85.2%, 91.1%, 95.7%의 성능을 보여주었다.

또한 도용 이미지 판별에 적합한 모델 학습법을 비교하기 위해 기존의 학습법과 도용 이미지를 활용한 학습법의 성능

을 비교하였으며 결과는 표 2에서 볼 수 있다. 테스트 과정은 동일하게 도용 이미지를 입력하였으며 기존 학습법의 경우 훈련 과정에서 동일 이미지는 기준 이미지가 30도 안에서 랜덤으로 회전된 이미지를 사용하였다. 성능 측정 결과 기존 학습법의 경우 Top-1, Top-5, Top-10, Top-20 순으로 34.8%, 56.6%, 75.0%, 79.2%의 성능을 보여주었으며 도용 이미지를 이용한 학습법보다 성능이 낮게 나온 것을 볼 수 있다. 이는 도용 이미지를 이용한 학습법이 기존 학습법보다 도용 이미지 포착에 더 적합하다는 것을 보여준다.

표 1. Top-k accuracy(%) 측정 결과(상: Adam optimizer, 하: SGD optimizer)
Table 1. Results of top-k accuracy(%) (Top: Adam optimizer, Bottom: SGD optimizer)

Model	Top-1	Top-5	Top-10	Top-20
Ours(Adam)	65.0	88.6	94.3	97.3
Ours(SGD)	63.5	85.2	91.1	95.7

표 2. 기존 학습법과 도용 이미지를 활용한 학습법의 Top-k accuracy (%) 측정 결과(상: 기존 학습법, 하: 도용 이미지를 활용한 학습법)

Table 2. Results of base, stolen images training top-k accuracy(%) (Top: Base, Bottom: Stolen images training)

Model	Top-1	Top-5	Top-10	Top-20
Base	34.8	56.6	75.0	79.2
Stolen	65.0	88.6	94.3	97.3

식 (5)에서 $Similarity$ 는 테스트셋의 원본 이미지들과 도용 이미지의 평균 유사도를 의미하며 $sim(x_i, y_i)$ 는 각 원본, 도용 이미지의 유사도를 의미한다.

$$Similarity = \frac{\sum_{i=0}^{samples-1} sim(x_i, y_i)}{samples} \tag{5}$$

표 3은 전체 원본 이미지와 도용 이미지 간 평균 유사도 측정 결과를 보여준다. Adam의 경우 69.4%의 평균 유사도를 보여주며 SGD의 경우 66.3%의 결과를 보여준다.

표 3. 원본, 도용 이미지 평균 유사도(Similarity)(%) 측정 결과(왼쪽: Adam optimizer, 오른쪽: SGD optimizer)
Table 3. Results of original, stolen images similarity(%) (Left: Adam optimizer, Right: SGD optimizer)

Model	Adam	SGD
Ours	69.4	66.3

3-4 세부 구현

본 연구에서는 모델 훈련에 NVIDIA GTX 4090 GPU 한대를 사용하였으며 자세한 실험 환경은 표 4과 같다. 모델의 하이퍼파라미터는 150 훈련 에포크, 옵티마이저 학습률(lr) 0.0003, 훈련 한 번에 1024개의 이미지를 배치 크기로 정의하였다. 옵티마이저는 Adam을 사용하였으며 Loss는 Pytorch에서 제공하는 BCE Loss를 사용하였다. 모든 이미지는 OpenCV 라이브러리[19]를 통해 1차원 grayscale로 변환하였다. 이미지는 105×125 픽셀로 스케일링되며 정규화 과정을 거쳐 평균(mean) 0.8444, 표준편차(standard deviation) 0.5329의 값이 적용된다. 정규화를 마친 이미지는 albumentations 라이브러리[20]의 ToTensorV2를 이용하여 리스트 자료형에서 텐서 자료형으로 변환하였다.

표 4. 모델 학습에 사용된 실험 환경

Table 4. Experimental environments used to train models

Parameter	Value
Ubuntu	22.04 LTS
GPU	NVIDIA GTX 4090
CUDA	Ver 12.1
Python	Ver 3.10.13
Pytorch	Ver 2.2.0

3-5 응용 시스템

응용 시스템의 전체 프레임워크는 그림 5에서 볼 수 있다. 사용자 인터페이스는 Flask 웹 프레임워크를 이용하여 개발하였다. 드래그 앤 드롭 인터랙션을 사용하여 도용 의심 이미지를 입력하고 ‘Upload’ 버튼 클릭 인터랙션으로 백엔드로 전달 및 모델을 통한 유사도 측정을 수행한다. 입력한 도용 의심 이미지와 이미 보유하고 있는 원본 이미지들간 유사도를 측정하고 유사도가 높은 순서대로 10개의 이미지를 나열하였으며 실행 결과는 그림 6에서 볼 수 있다. 10개의 원본 이미지의 개수는 Top-k의 성능 결과를 참고하여 설정하였으며 Top-10의 결과가 94.3%임으로 입력된 이미지가 도용 이미지일 경우 도용 이미지의 원본 이미지가 출력된 10개의 원본 이미지 중 하나일 확률은 94.3%임을 보여준다. 위와 같은 과정은 방대한 양의 원본 이미지와 비교하는 작업을 보다 빠르고 쉽게 수행하여 시간과 비용을 절약하는 효과가 있다.

또한, 사용자는 입력한 이미지와 비교에서 특정 유사도 이상의 원본 이미지 추출 기능을 제공한다. 유사도 Threshold 값을 입력하고 그 이상의 유사도 값을 가진 원본 이미지만을 웹에 표시하는 기능이며 사용 예시는 그림 7에서 볼 수 있다. 예를 들어 Threshold를 40으로 입력할 경우 입력된 이미지와 40% 이상의 유사도를 가진 원본 이미지만을 표시하여 사용자의 도용 이미지 분석을 용이하게 하였다.

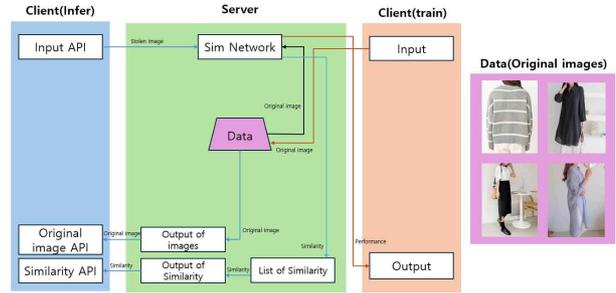


그림 5. 응용 시스템의 전체 프레임워크

Fig. 5. The overall framework of the application system



그림 6. 응용 시스템 실행 결과

Fig. 6. Results of application system

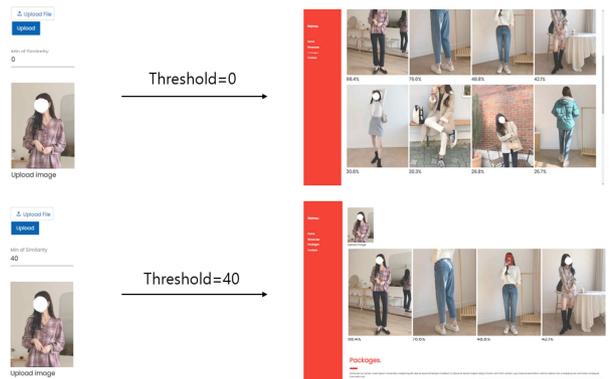


그림 7. 응용 시스템의 유사도 Threshold 기능 사용 예시

Fig. 7. Examples of using the similarity threshold function in application systems

IV. 결론 및 향후 연구

본 연구는 Siamese Network를 활용하여 이커머스에서 도용한 이미지를 쉽고 빠르게 찾는 응용 시스템을 소개하였다. 사용자가 도용 의심 이미지를 입력하면 원본 이미지들과 유사도를 측정하고 유사도가 높은 상위 10개의 원본 이미지만을 출력함으로써 쉽고 빠른 도용 이미지 판별 작업이 가능케 하였다. 또한 Top-k 정확도를 활용하여 유사도가 높은 상위 10개의 원본 이미지와의 비교 작업이 94.3%의 정확도의

도용 이미지 판별을 보장함을 보였다.

본 연구에서는 의류 이미지 데이터만으로 모델을 학습하여 의류 이커머스에서의 이미지 도용 문제를 해결하는 응용 프로그램을 제시하였다. 제시한 응용 시스템은 식품, 사무용품과 같은 다양한 종류의 이미지 데이터로 모델을 학습함으로써 다양한 종류의 이커머스 플랫폼에서 사용 가능한 응용 프로그램으로 발전할 수 있을 거라 기대된다. 또한 색깔 변화나 일부 이미지가 삽입된 도용 이미지의 유사도 측정 연구도 추후 지속할 예정이다.

감사의 글

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 인공지능융합혁신인재양성사업(IITP-2023-RS-2023-00256629), 지역지능화혁신인재양성사업(IITP-2024-00156287)의 연구 결과로 수행되었음.

참고문헌

- [1] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," *Nature*, Vol. 521, No. 7553, pp. 436-444, May 2015. <https://doi.org/10.1038/nature14539>
- [2] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation," in *Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition*, Columbus: OH, pp. 580-587, June 2014. <https://doi.org/10.1109/CVPR.2014.81>
- [3] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in *Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas: NV, pp. 779-788, June 2016. <https://doi.org/10.1109/CVPR.2016.91>
- [4] C. Feichtenhofer, H. Fan, J. Malik, and K. He, "Slowfast Networks for Video Recognition," in *Proceedings of 2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, pp. 6202-6211, November 2019. <https://doi.org/10.1109/ICCV.2019.00630>
- [5] S. Yan, Y. Xiong, and D. Lin, "Spatial Temporal Graph Convolutional Networks for Skeleton-Based Action Recognition," in *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI-18)*, New Orleans: LA, pp. 7444-7452, February 2018. <https://doi.org/10.1609/aaai.v32i1.12328>
- [6] X. Zhao, Y. Wu, G. Song, Z. Li, Y. Zhang, and Y. Fan, "A Deep Learning Model Integrating FCNNs and CRFs for Brain Tumor Segmentation," *Medical Image Analysis*, Vol. 43, pp. 98-111, January 2018. <https://doi.org/10.1016/j.media.2017.10.002>
- [7] L. Kong and J. Cheng, "Classification and Detection of COVID-19 X-Ray Images Based on DenseNet and VGG16 Feature Fusion," *Biomedical Signal Processing and Control*, Vol. 77, 103772, August 2022. <https://doi.org/10.1016/j.bspc.2022.103772>
- [8] G. Koch, R. Zemel, and R. Salakhutdinov, "Siamese Neural Networks for One-Shot Image Recognition," in *Proceedings of the 32nd International Conference on Machine Learning (ICML '15)*, Lille, France, July 2015.
- [9] J. Bromley, I. Guyon, Y. LeCun, E. Säcker, and R. Shah, "Signature Verification Using a "Siamese" Time Delay Neural Network," in *Proceedings of the 6th International Conference on Neural Information Processing Systems (NIPS '93)*, San Francisco: CA, pp. 737-744, November 1993.
- [10] S. Chopra, R. Hadsell, and Y. LeCun, "Learning a Similarity Metric Discriminatively, with Application to Face Verification," in *Proceedings of 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, San Diego: CA, pp. 539-546, June 2005. <https://doi.org/10.1109/CVPR.2005.202>
- [11] F. Schroff, D. Kalenichenko, and J. Philbin, "Facenet: A Unified Embedding for Face Recognition and Clustering," in *Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston: MA, pp. 815-823, June 2015. <https://doi.org/10.1109/CVPR.2015.7298682>
- [12] A. Parnami and M. Lee, "Learning from Few Examples: A Summary of Approaches to Few-Shot Learning," arXiv:2203.04291, March 2022. <https://doi.org/10.48550/arXiv.2203.04291>
- [13] L. Fe-Fei, Fergus, and Perona, "A Bayesian Approach to Unsupervised One-Shot Learning of Object Categories," in *Proceedings of the 9th IEEE International Conference on Computer Vision*, Nice, France, pp. 1134-1141, October 2003. <https://doi.org/10.1109/ICCV.2003.1238476>
- [14] B. Lake, R. Salakhutdinov, J. Gross, and J. Tenenbaum, "One Shot Learning of Simple Visual Concepts," in *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society (CogSci 2011)*, Boston: MA, pp. 2568-2573, July 2011.
- [15] AI-Hub. K-Fasion Image [Internet]. Available: <https://www.aihub.or.kr/aihubdata/data/view.do?currMenu=115&topMenu=100&aihubDataSe=data&dataSetSn=51>.

- [16] A. F. M. Agarap, "Deep Learning Using Rectified Linear Units (ReLU)," arXiv:1803.08375, March 2018. <https://doi.org/10.48550/arXiv.1803.08375>
- [17] D. P. Kingma and J. L. Ba, "Adam: A Method for Stochastic Optimization," arXiv:1412.6980, December 2014. <https://doi.org/10.48550/arXiv.1412.6980>
- [18] S. Ruder, "An Overview of Gradient Descent Optimization Algorithms," arXiv:1609.04747, September 2016. <https://doi.org/10.48550/arXiv.1609.04747>
- [19] I. Culjak, D. Abram, T. Pribanic, H. Dzapo, and M. Cifrek, "A Brief Introduction to OpenCV," in *Proceedings of the 35th International Convention MIPRO*, Opatija, Croatia, pp. 1725-1730, May 2012.
- [20] A. Buslaev, V. I. Iglovikov, E. Khvedchenya, A. Parinov, M. Druzhinin, and A. A. Kalinin, "Albumentations: Fast and Flexible Image Augmentations," *Information*, Vol. 11, No. 2, 125, February 2020. <https://doi.org/10.3390/info11020125>



조준수 (Jun-Su Cho)

2022년 : 전남대학교 학사
2023년 : 전남대학교 대학원 (공학석사 - 인공지능융합)

2023년~현 재: 전남대학교 인공지능융합학과 석사과정
※ 관심분야 : 딥러닝, 증강/가상현실, 인간-컴퓨터 상호작용



김승원 (Seung-Won Kim)

2009년 : University of Tasmania 학사
2011년 : University of Tasmania 석사
2016년 : University of Canterbury 박사

2016년~2020년: University of South Australia 박사 후 연구원
2017년~2020년: Swinburne University of Technology 박사 후 연구원

2020년~2021년: 한국과학기술연구원 박사 후 연구원

2021년~현 재: 전남대학교 인공지능학부 조교수

※ 관심분야 : 딥러닝, 증강/가상현실, 인간-컴퓨터 상호작용