

Optimizing CNN Structure to Improve Accuracy of Artwork Artist Classification

Ji-Seon Park*, So-Yeon Kim*, Yeo-Chan Yoon**, Soo Kyun Kim*

*Student, Dept. of Computer Engineering, Jeju National University, Jeju, Korea

*Student, Dept. of Computer Engineering, Jeju National University, Jeju, Korea

**Professor, Dept. of Artificial Intelligence, Jeju National University, Jeju, Korea

*Professor, Dept. of Computer Engineering, Jeju National University, Jeju, Korea

[Abstract]

Metaverse is a modern new technology that is advancing quickly. The goal of this study is to investigate this technique from the perspective of computer vision as well as general perspective. A thorough analysis of computer vision related Metaverse topics has been done in this study. Its history, method, architecture, benefits, and drawbacks are all covered. The Metaverse's future and the steps that must be taken to adapt to this technology are described. The concepts of Mixed Reality (MR), Augmented Reality (AR), Extended Reality (XR) and Virtual Reality (VR) are briefly discussed. The role of computer vision and its application, advantages and disadvantages and the future research areas are discussed.

▶ **Key words:** Computer Vision, CNN, Artwork Classification, Fine-tuning, ResNet50

[요 약]

컴퓨터 비전 분류 연구에서 합성곱 신경망 (Convolutional Neural Network)은 탁월한 이미지 분류 성능을 보여준다. 이에 영감을 받아 예술 관련 이미지 분류 작업에 대한 적용 가능성을 분석해 본다. 본 논문에서는 예술 작품 아티스트 분류의 정확도를 향상시키기 위해 최적화된 합성곱 신경망 구조를 제안한다. 미세 조정 범위 시나리오와 완전연결층 조정 시나리오를 세운 뒤 그에 따른 예술 작품 아티스트 분류의 정확도를 측정했다. 즉, 학습 컨볼루션 레이어(Convolution layer) 수와 완전연결층 수 등 ResNet50 모델의 구조를 변경하며 예술 작품 아티스트 분류의 정확도가 향상되도록 최적화했다. 본 논문에서 제안하는 합성곱 신경망 구조는 기존 예술 작품 아티스트 분류에서 쓰이던 AlexNet 모델을 1-GPU 버전으로 수정한 CaffeNet 모델보다 더 높은 정확도를 실험 결과에서 증명한다.

▶ **주제어:** 컴퓨터 비전, 합성곱 신경망, 예술 작품 아티스트 분류, 미세 조정, ResNet50

-
- First Author: Ji-Seon Park, Corresponding Author: Soo Kyun Kim
 - *Ji-Seon Park (wls1020528@gmail.com), Dept. of Computer Engineering, Jeju National University
 - *So-Yeon Kim (hbna567111@gmail.com), Dept. of Computer Engineering, Jeju National University
 - **Yeo-Chan Yoon (ycyoon@jejunu.ac.kr), Dept. of Artificial Intelligence, Jeju National University
 - *Soo Kyun Kim (kimsk@jejunu.ac.kr), Dept. of Computer Engineering, Jeju National University
 - Received: 2023. 06. 29, Revised: 2023. 08. 16, Accepted: 2023. 08. 16.

I. Introduction

컴퓨터 비전 분류 연구[1]에서 합성곱 신경망(Convolutional Neural Network)의 성공에 대한 핵심 근거 중 하나는 대규모 라벨링 데이터셋이 있다는 것이다. 대표적으로 1,500만 개 이상의 고해상도 이미지가 약 22,000개의 카테고리 라벨링 된 이미지넷(ImageNet)[2]이 있다. 만약 사용 가능한 온라인 상의 모든 예술 작품 컬렉션을 모두 모으더라도 이는 이미지넷 데이터셋의 이미지 수보다 훨씬 적기 때문에 과적합 없이 합성곱 신경망을 처음부터 훈련시키기에 충분하지 않다. 그러나 제한된 크기의 데이터셋을 다루는 다양한 이미지 분류 작업[3]에서 이미지넷 데이터셋으로 사전 훈련된 합성곱 신경망을 새로운 목표 데이터셋과 과제에 맞게 미세 조정하여 분류 성능을 높인 바 있다. 본 연구에서는 이미지넷에서 사전 훈련된 합성곱 신경망 중 Resnet50[4]을 이용하여 아티스트 분류와 같은 예술 작품에 특화된 작업을 위해 어떻게 미세 조정할 수 있는지에 대해 검증해 보기로 한다.

II. Preliminaries

1. Related works

예술 작품 분류에서의 합성곱 신경망은 특징 추출기로서 처음 도입되었다. 이미지넷에서 학습된 합성곱 신경망의 레이어 활동을 예술적 스타일 인식을 위한 특징으로 사용하는 접근 방식은 Karayev 등(2014)[5]에 의해 소개되었다. Karayev 등은 비예술적 이미지의 객체 인식을 위해 학습된 합성곱 신경망의 레이어에서 파생된 특징을 활용한 예술 작품 분류 작업에서 높은 성능을 달성했다. 이는 대부분의 수작업 분류를 능가하는 방법을 보여주었다. 예술 작품 분류 작업에서 가장 큰 문제는 결과를 적절히 비교할 수 있는 일반적으로 통용되는 대규모 데이터셋이 없다는 점이었다. 동일한 분류 과제를 다룬 연구들은 서로 다른 중소규모의 미술품 컬렉션을 사용하거나 임의로 선택한 서로 다른 데이터셋을 사용했다. Eva Cetinic[6] 등은 WikiArt 데이터셋과 WGA 데이터셋에서 장르, 스타일, 아티스트, 아티스트의 국적 총 네 가지 분류 작업을 정의했다. 다양한 예술 관련 작업과 데이터셋에 가장 적합한 훈련 설정을 파악하기 위해 다양한 미세 조정 시나리오를 비교했으며, 특히 도메인별 가중치 초기화의 영향을 살펴보는 데 중점을 두었다. 그리고 이를 바탕으로 다양한 도메인별 이미지 특징의 관계를 분석한 결과를 보여주어 타겟 데

이터셋이 많은 클래스로 구성되고, 클래스 당 이미지 수가 적은 경우 사전 훈련된 모델을 초기화하는 것이 미세 조정 성능에 영향을 미친다는 사실을 보여준다. 이는 미세 조정 에 대한 이해를 높이는 데 유용할 수 있으나, 분류 작업이 일반적으로 최고의 정확도는 달성하지 못했다. 본 연구에서는 예술 작품 아티스트 분류 성능을 극대화하기 위해 위 연구[6]에서 사용한 레이어가 8개인 AlexNet[7]과 같이 비교적 레이어가 얇은 아키텍처 대신 50개의 레이어로 구성되어 더 심층적인 아키텍처인 ResNet50으로 확장하여 분류 작업의 성능을 향상시키고자 한다.

2. Dataset

2.1 WikiArt

가능한 많은 수의 그림을 포함하고 광범위한 분류 작업을 위해 현재 온라인에서 가장 큰 규모의 디지털화된 그림 컬렉션인 WikiArt를 데이터셋으로 사용했다. WikiArt는 작가, 스타일, 장르, 국적, 기법 등과 같은 광범위한 메타 데이터를 통합하는 잘 조직된 컬렉션이다. 특히 19세기와 20세기, 현대 미술에 중점을 둔 광범위한 시대의 예술 작품이 포함되어 있다. 데이터 수집 과정 당시 WikiArt 데이터셋에는 총 133,220개의 예술 작품이 포함되어 있으며, 우리는 작품 수가 많은 아티스트 25명의 예술 작품 데이터만을 사용했다.

2.2 Weighted Cross Entropy

25명의 아티스트의 예술 작품으로 이루어진 훈련 데이터셋은 빈센트 반 고흐의 작품 수가 1,510개, 장바티스트 카미유 코로의 작품 수가 381개로, 일부 몇 개의 클래스에만 샘플들이 몰려 있다. 그리고 대부분의 클래스(Class)에는 매우 적은 수의 샘플이 있는 Long Tailed Dataset의 형태를 띠고 있다. 이 데이터셋으로 효과적으로 학습시키기 위해서 Weighted Cross Entropy[8]를 사용한다. 이는 가중치(weight) 인자에 클래스별 가중치를 전달하여 데이터 비율에 맞게 학습되는 정도를 조정하여, 데이터 불균형을 해결한다.

2.3 Data Augmentation

데이터 증강(Data Augmentation)[9]은 기존 데이터셋에서 새로운 데이터를 생성하여 학습 데이터셋의 크기와 다양성을 늘리는 데 사용되는 기술이다. 학습 데이터셋을 변형하여 모델이 더 다양한 데이터를 학습하도록 함으로써 모델의 일반화 성능을 향상시킨다. 또한, 학습 데이터셋이 적은 경우 모델이 학습 데이터에 과적합되지 않도록

도움을 주기 때문에 딥러닝에서는 다양한 데이터 증강 기법이 사용된다. 이러한 이유로 이미지 분류 분야에서도 다양한 데이터 증강이 사용된다. 본 논문에서는 데이터셋의 크기가 작음을 고려하여 Albumentations 라이브러리를 활용해 크기 조정, 뒤집기, 회전 등의 이미지 증강과 밝기, 대비를 무작위로 변경하도록 하는 이미지 증강을 적용했다. 또한, CLAHE 증강 기법[10]을 이용해 이미지의 대비를 강화하여 예술 작품의 세부 정보가 더욱 뚜렷하게 보이도록 했다. 그리고 패치의 크기와 위치를 무작위로 하여 한 이미지에서 잘라낸 패치를 다른 이미지에 붙이는 방식을 통해 데이터를 증강하는 CutMix[11]기법을 사용했다. CutMix기법은 Mixup[12]이나 Cutout[13]보다 더 나은 성능을 보여주지만 영상을 떼어 붙이는 과정에서 원본 영상이 갖는 의미론적 속성을 잃을 수 있는 단점이 있는 증강 기법이다. 그러나 예술 작품 아티스트 분류를 위해서는 영상의 형태보다도 텍스처에 집중해야 하기에 CutMix기법을 사용한다.

III. The Proposed Scheme

3.1. Resnet50

ResNet은 2015년부터 2017년까지 3년 연속으로 ILSVRC(ImageNet Large Scale Visual Recognition Challenge)에서 우승을 차지한 CNN 모델이다. ResNet이 나오기 전까지의 모델들[14]은 성능을 높이고자 레이어를 깊게 쌓는 데 중점을 두었고, 이는 역전파 과정에서 출력층에서 멀어질수록 기울기 값이 매우 작아지는 기울기 소실 문제를 야기했다. ResNet에서는 스킵 연결(skip connection)을 이용한 잔차 블록(Residual Block)을 사용함으로써 이 문제를 해결했다. 잔차 학습(Residual learning)을 통해 모델의 깊이를 증가시키는 데 성공함으로써 무려 152층까지 레이어를 쌓으며 처음으로 인간 오류(Human error)를 능가하는 3.57%의 top-5 error rate를 달성했다. ResNet은 층의 개수에 따라 ResNet18, 34, 50 등으로 나뉘지는데 본 논문에서는 ResNet50을 이용하여 미세 조정 실험을 진행한다.

3.2 Hyperparameters

두 가지 시나리오에서 다음과 같은 최적화 알고리즘, 손실 함수를 동일하게 사용한다.

```

nn.CrossEntropyLoss()
optim.SGD(resnet.parameters(),
           lr=0.001, momentum=0.9)

optim.lr_scheduler.StepLR(optimizer,
                           step_size=10, gamma=0.5)
    
```

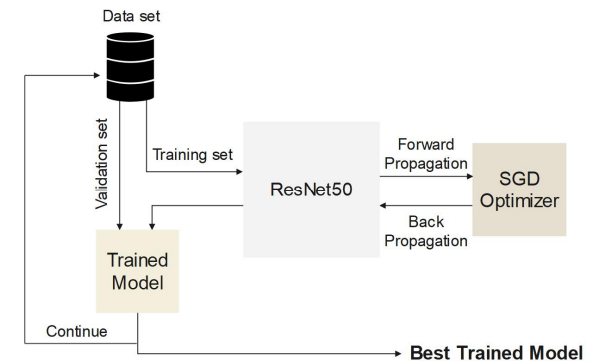


Fig. 1. Training course diagram

손실 함수로 교차-엔트로피 손실함수(Cross-entropy Loss)를 사용하였고, 최적화 알고리즘으로 SGD[15]를 사용했다. 이때 모멘텀을 0.9로 설정해 진동을 줄이고 빠르게 수렴되도록 했다. 그리고 완전연결층 조정 시나리오에서는 StepLR 스케줄러를 사용해 10 epoch마다 0.5의 비율로 lr이 감소되도록 했다. Fig 1은 예술 작품 아티스트 분류를 위한 Resnet50 훈련 과정의 개요이다. Wiki Art 학습 데이터를 활용해 forward propagation과 back propagation을 거치며, 손실함수를 최소화하는 SGD 방법을 이용하여 모델을 훈련시킨다. 이때, validation dataset을 이용하여 훈련된 모델의 성능을 평가하고, 계속해서 학습을 진행하여 최고의 성능을 보이는 Best Trained Model을 찾아낸다.

3.3 Fine-tuning range scenarios

초기 레이어는 다양한 이미지 인식 작업과 관련된 기능을 추출하기 때문에 일반적으로 마지막 또는 마지막 몇 개의 레이어만 미세 조정하면 네트워크가 타겟 도메인의 특성에 적응하고 성능이 향상된다. 분류 문제에 대해 타겟 데이터 세트 크기와 타겟 도메인과 소스 도메인 간의 유사성을 기반으로 가장 효율적인 것을 찾기 위해 다음과 같이 4가지 미세 조정 범위 시나리오를 세웠다.

Table 1. Number of trainable parameters based on the number of convolution layers

	Trainable params	Non-Trainable params	Total params
no_freeze	23,559,257	0	23,559,257
Conv1_freeze	23,549,721	9,536	
Conv1, Conv2_freeze	23,333,913	225,344	
Conv1, Conv2, Conv3_freeze	22,114,329	1,444,928	

1. 모든 컨볼루션 레이어의 가중치를 고정하지 않는다.
(no_freeze)
2. 첫번째 컨볼루션 레이어의 가중치를 고정한다.
(Conv1_freeze)
3. 두번째 컨볼루션 레이어까지 가중치를 고정한다.
(Conv1, Conv2_freeze)
4. 세번째 컨볼루션 레이어까지 가중치를 고정한다.
(Conv1, Conv2, Conv3_freeze)

Table 1을 통해 고정시키는 컨볼루션 레이어 (Convolution layer)의 수가 증가할수록 학습 가능한 모델의 파라미터 수가 줄어드는 것을 확인할 수 있다. 합성곱 층 중 저차원의 특징을 추출하는 초기 레이어는 고정하여 학습이 되지 않도록 하고, 이후 레이어는 학습이 가능하게 두어 새로운 데이터에 대한 특징을 학습하도록 한다.

3.4 Fully Connected Layer Adjustment Scenario

완전연결층은 컨볼루션 레이어의 출력을 가져와 입력 이미지의 클래스 확률을 나타내는 최종 출력을 생성한다. 따라서 적절한 개수의 완전연결층을 사용하는 것 또한 분류 모델의 성능을 높일 수 있는 방법이다. 본 논문에서는 최적의 완전연결층 수를 찾기 위해 세 가지 시나리오를 세웠다. Table 2는 각 완전연결층의 입력 크기와 출력 크기를 나타낸 것이다. 완전연결층을 1개 사용하는 경우 입력 크기는 Resnet50 구조에서 최종적으로 얻어지는 특성 맵의 크기인 2048이고, 출력 크기는 분류 아티스트 수에 해당하는 25이다. 최종 출력 크기를 제외하고 나머지 출력 크기는 입력 크기의 절반이 되도록 했다. 따라서 완전연결층이 2개, 3개인 경우 입력 크기는 2048, 1024, 512로 감소하는 것을 확인할 수 있다.

Table 2. Fully Connected Layer Adjustment Scenario

	Input size	Output size
1 Fully Connected Layer	2048	25
2 Fully Connected Layer	2048	1024
	1024	25
3 Fully Connected Layer	2048	1024
	1024	512
	512	25

Table 3. Loss and accuracy by fine-tuning range scenario

	Train_loss	Val_acc (%)
no_freeze	1.24778	79.60227
Conv1_freeze	1.24702	79.56439
Conv1,2_freeze	1.24843	79.31817
Conv1,2,3_freeze	1.26923	78.73106

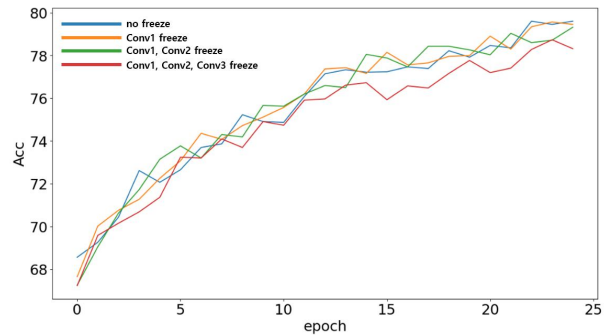


Fig. 2. Accuracy by fine-tuning range scenarios

Table 4. Loss and accuracy by fully connected layer adjustment scenario

	Train_loss	Val_acc(%)
1 Fully Connected Layer	1.16870	79.75378
2 Fully Connected Layer	0.79718	84.98106
3 Fully Connected Layer	1.22217	79.50757

Table 5. Comparison of artwork artist classification accuracy

Reference	Method	Dataset	acc(%)
Our results	CNN fine-tuning, (ResNet50)	WikiArt	84.98
Cetinic et. al	CNN fine-tuning (CaffeNet)	WikiArt	81.94

IV. Experiment Results

4.1 Experimental Environments

본 실험은 프로세서 라이젠 7 3700X 8-Core, 그래픽 카드 GeForce RTX 2070 SUPER, 메모리 16GB 등으로 구성된 장비에 파이썬 3.9 기반의 파이토치 1.9.0으로 신경망을 구축했다.

4.2 Experimental results

미세 조정 범위에 따른 손실(Loss)값과 정확도는 Table 3과 같다. 파라미터 고정 범위가 커질수록 모델의 정확도가 떨어지는 현상이 나타났다. 즉, 모든 층의 가중치가 고정되지 않고 미세 조정 되도록 훈련시킴으로써 예술 작품 아티스트 분류 작업에서 높은 정확도를 얻었다. Fig 2는 미세 조정 범위에 따른 훈련 과정에서의 검증 데이터셋의 분류 정확도를 비교한 그래프이다. 모든 층의 가중치가 고정되지 않고 미세 조정 되도록 훈련 시킨 모델의 정확도가 높은 것을 확인할 수 있다. 완전연결층 수에 따른 정확도는 Table 4와 같다. 이때, 완전연결층 조정 시나리오는 미세 조정 범위 시나리오에서 가장 높은 정확도를 보였던 모든 층의 가중치를 고정하지 않는 방식으로 훈련을 진행했다. 완전연결층 수를 1개로 하여 바로 결과를 출력한 것보다 2개로 늘려 결과를 출력한 것이 더 높은 정확도를 보여주었다. 그리고 완전연결층을 3개 이상으로 늘렸을 때보다도 높은 정확도를 나타내, 2개가 최적의 완전연결층 수인 것을 확인했다.

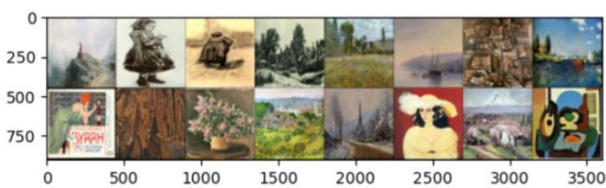


Fig. 3. Part of the validation dataset

실험 결과 최적의 미세 조정 범위였던 모든 층의 가중치를 고정하지 않는 방식으로 2개의 완전연결층을 갖는 구조의 모델을 학습시킨 뒤 Fig 3와 같은 예술 작품을 아티스트 별로 분류하는 작업을 수행한 결과 84.98%의 정확도 값을 나타냈다. Fig 4은 완전연결층 수에 따른 검증 데이터셋의 분류 정확도를 비교한 그래프로, 2개의 완전연결층을 갖는 모델의 정확도가 눈에 띄게 높은 것을 확인할 수 있다.

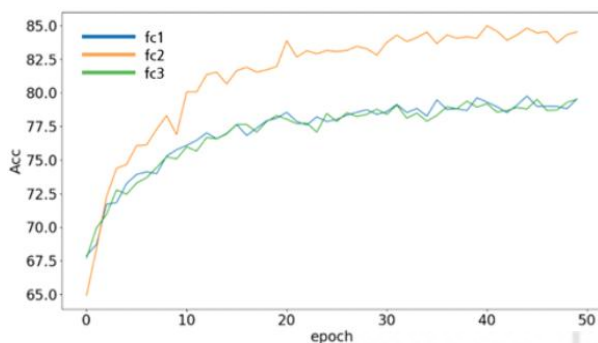


Fig. 4. Accuracy by fully connected layer adjustment scenario

V. Conclusions

본 논문에서는 CNN 기반의 예술 작품 아티스트 분류에서 분류 정확도를 높이기 위해서 2가지 미세 조정 실험을 진행했다. 가장 높은 예술 작품 아티스트 분류 정확도를 보이는 구조는 Resnet50의 5개의 모든 컨볼루션 레이어의 가중치를 고정하지 않은 두 개의 완전연결층을 갖는 구조이며 정확도 값은 84.98%이다. Table 5는 본 논문에서 제안하는 합성곱 신경망 구조와 기존 연구[6]의 AlexNet CNN 구조를 사용한 CaffeNet 모델의 예술 작품 아티스트 분류 정확도를 비교한 것이다. 본 논문에서 제안하는 Resnet50의 모든 컨볼루션 레이어의 가중치를 고정하지 않은 두 개의 완전연결층을 갖는 구조가 약 3.04% 더 높은 정확도 값을 더 나타내고 있다. 8개의 계층을 가진 AlexNet 기반 CaffeNet 대신 50개의 계층을 가진 Resnet50을 사용하여 복잡한 예술 작품 아티스트 분류의 성능을 향상시켰다. 이와 더불어 미세 조정 범위 시나리오와 기존 연구와 차별화된 완전연결층 조정 시나리오를 통해 합성곱 신경망 구조를 최적화함으로써 예술 작품 아티스트 분류의 정확도를 향상시켰다. 그러나 기존 연구와 비교하였을 때 예술 작품 아티스트 분류에서 약 3.04%의 미미한 정확도 향상이 이루어졌다. 이러한 결과로 보아, 예술 작품 아티스트 분류라는 복잡한 작업에서 더 높은 성능을 얻기 위해서는 데이터 전처리와 모델 아키텍처 개선을 위한 추가적인 연구가 필요하다는 한계점이 있다. 일반적으로 타겟 도메인과 소스 도메인 간의 유사성이 작을 때 데이터셋의 크기가 크면 사전 훈련된 네트워크의 전체를 재학습시키는 방법이 유용하고, 크기가 작으면 컨볼루션 레이어 일부분과 완전연결층을 학습시키는 것이 유용하다. 예술 작품은 ImageNet 데이터셋과 유사성이 매우 낮고 실험에서 사용한 학습 데이터셋의 크기가 작기 때문에 컨볼루션 레이어 일부분을 고정시키는 것이 더 좋은 분류 성능을 보일 것으로 예상하였으나, 네트워크 전체를 재학습시키는 방법이 가장 적합함을 입증했다. 분석 결과 Data Augmentation을 통해 데이터셋의 규모를 늘린 것이 이에 영향을 준 것으로 보인다. 향후 연구에서는 데이터 증강 기법에 따른 예술 작품 아티스트 분류의 정확도를 타 모델과 비교해 보고, 예술 작품 아티스트 분류에 적합한 데이터 증강 기법을 연구하고자 한다.

ACKNOWLEDGEMENT

“This work was supported by the 2023 education, research and student guidance grant funded by Jeju National University”

REFERENCES

- [1] A Krizhevsky, I Sutskever, GE Hinton, “ImageNet classification with deep convolutional neural networks”, *Communications of the ACM*, Volume 60, Issue 6, pp 84–90, 2017, DOI: 10.1145/3065386
- [2] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, Li Fei-Fei, “ImageNet: A large-scale hierarchical image database”, 2009 IEEE Conference on Computer Vision and Pattern Recognition, pp. 1063-6919, 2009, DOI: 10.1109/CVPR.2009.5206848
- [3] Angie K. Reyes, Juan C. Caicedo, and Jorge E. Camargo, “Fine-tuning Deep Convolutional Networks for Plant Recognition”, *CEUR Workshop Proceedings*, Volume 1391, 2015
- [4] K. He, X. Zhang, S. Ren, J. Sun, “Deep Residual Learning for Image Recognition”, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770-778, 2016, DOI: 10.1109/CVPR.2016.90
- [5] Sergey Karayev, Matthew Trentacoste, Helen Han, Aseem Agarwala, Trevor Darrell, Aaron Hertzmann, and Holger Winnemoeller, “Recognizing Image Style”, *British Machine Vision Conference*, v3, pp. 11, 2014, DOI: 10.5244/C.28.122
- [6] Eva Cetinic, Tomislav Lipic, Sonja Grgic, “Fine-tuning Convolutional Neural Networks for fine art classification”, *Expert Systems with Applications*, Volume 114, pp. 107-118, 2018, DOI: 10.1016/j.eswa.2018.07.026
- [7] Krizhevsky, Alex, Ilya Sutskever, Geoffrey E. Hinton, "Imagenet classification with deep convolutional neural networks", *Communications of the ACM*, Volume 60, Issue 6, Page, pp. 84-90, 2017, DOI: 10.1145/3065386
- [8] Yuri Sousa Aurelio, Gustavo Matheus de Almeida, Cristiano Leite de Castro, Antonio Padua Braga, “Learning from Imbalanced Data Sets with Weighted Cross-Entropy Function”, *Neural Processing Letters*, volume 50, Issue 2, pp. 1937–1949, 2019, DOI: 10.1007/s11063-018-09977-1
- [9] Agnieszka Mikołajczyk; Michał Grochowski, “Data augmentation for improving deep learning in image classification problem”, *International Interdisciplinary PhD Workshop (IIPHDW)*, 2018, DOI: 10.1109/IIPHDW.2018.8388338
- [10] Agung W. Setiawan, Tati R. Mengko, Oerip S. Santoso, Andriyan B. Suksmono, “Color retinal image enhancement using CLAHE”, *International Conference on ICT for Smart Society*, 2013, DOI: 10.1109/ICTSS.2013.6588092
- [11] Sangdoon Yun, Dongyoon Han, Seong Joon Oh, Sanghyuk Chun, Junsuk Choe, Youngjoon Yoo, “CutMix: Regularization Strategy to Train Strong Classifiers with Localizable Features”, the IEEE International Conference on Computer Vision, Volume 2019-October, pp. 6023-3032, 2019, DOI: 10.1109/ICCV.2019.00612
- [12] Hongyi Zhang, Moustapha Cisse, Yann N. Dauphin, David Lopez-Paz, “mixup : Beyond Empirical Risk Minimization”, *International Conference on Learning Representations, ICLR, Vancouver*, 2018
- [13] Terrance DeVries, Graham W. Taylor, “Improved Regularization of Convolutional Neural Networks with Cutout”, *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.1-8, 2017, DOI: 10.48550/arXiv.1708.04552
- [14] SZEGEDY, Christian, et al, “Going deeper with convolutions”, *The IEEE conference on computer vision and pattern recognition*, 2015, DOI: 10.48550/arXiv.1409.4842
- [15] RUDER, Sebastian. “An overview of gradient descent optimization algorithms”, *ArXiv, Computer Science*, 2016 DOI: 10.48550/arXiv.1609.04747
- [16] Yangqing Jia, Evan Shelhamer, Jeff Donahue, Sergey Karayev, Jonathan Long, Ross Girshick, Sergio Guadarrama, Trevor Darrell, “Caffe: Convolutional Architecture for Fast Feature Embedding”, *ACM Multimedia*, pp 675-678, 2014, DOI: 10.48550/arXiv.1408.5093

Authors



Ji-Seon Park is currently a Student in the Department of Computer Engineering, Jeju National University. She is interested in Computer vision.



So-Yeon Kim is currently a Student in the Department of Computer Engineering, Jeju National University. She is interested in Computer vision.



Yeo-Chan Yoon received BS, MS and Ph.D degrees in computer science and engineering from Korea University, Seoul, Rep. of Korea, in 2004, 2007 and 2020 respectively. Currently, he is an assistant professor in the

Department of Artificial Intelligence at Jeju National University, Jeju-si, Rep. of Korea. His research interests include Deep Learning, Vision, Natural Language Processing and machine learning.



Soo Kyun Kim received Ph.D. in Computer Science & Engineering Department of Korea University, Seoul, Korea, in 2006. He joined the Telecommunication R&D center at Samsung Electronics Co., Ltd., from 2006

Department of Game Engineering at Paichai University until 2020. He is now a professor at the Department of Computer Engineering at Jeju National University, Korea.