

명함 이미지 회전 판단을 위한 딥러닝 모델 비교

경지훈*

Comparison of Deep Learning Models for Judging Business Card Image Rotation

Ji-Hoon Kyung*

*Assistant Professor, Department of Industrial & Management Engineering, Hannam University, Daejeon, 34430 Korea

요약

고객이 온라인으로 요청한 명함을 자동으로 명함을 인쇄하는 스마트 명함 인쇄 시스템이 활성화되고 있다. 이때, 문제는 고객이 시스템에 제출한 명함이 비정상일 수 있다는 것이다. 본 논문에서는 인공 지능 기술을 도입하여 명함의 이미지가 비정상적으로 회전됐는지 여부를 판정하는 문제를 다룬다. 명함은 0도, 90도, 180도, 270도 회전한다고 가정하였다. 특별한 인공신경망을 설계하지 않고 기존의 VGG, ResNet, DenseNet 인공신경망을 적용하여 실험하였는데 모든 신경망이 97% 정도의 정확도로 이미지 회전을 분별할 수 있었다. DenseNet161은 97.9%의 정확도를 보였고 ResNet34도 97.2%의 정밀도를 보였다. 이는 문제가 단순할 경우, 복잡한 인공신경망이 아니어도 충분히 좋은 결과를 낼 수 있음을 시사한다.

ABSTRACT

A smart business card printing system that automatically prints business cards requested by customers online is being activated. What matters is that the business card submitted by the customer to the system may be abnormal. This paper deals with the problem of determining whether the image of a business card has been abnormally rotated by adopting artificial intelligence technology. It is assumed that the business card rotates 0 degrees, 90 degrees, 180 degrees, and 270 degrees. Experiments were conducted by applying existing VGG, ResNet, and DenseNet artificial neural networks without designing special artificial neural networks, and they were able to distinguish image rotation with an accuracy of about 97%. DenseNet161 showed 97.9% accuracy and ResNet34 also showed 97.2% precision. This illustrates that if the problem is simple, it can produce sufficiently good results even if the neural network is not a complex one.

키워드: 명함, 이미지 회전, 인공 신경망, 스마트 프린팅 시스템

Keywords: business cards, image rotation, artificial neural networks, smart printing system

Received 26 November 2022, Revised 2 December 2022, Accepted 27 December 2022

* Corresponding Author Ji-Hoon Kyung(E-mail:kjh@hnu.kr, Tel:+82-42-629-8150)

Assistant Professor, Department of Industrial & Management Engineering, Hannam University, Daejeon, 34430 Korea

Open Access <http://doi.org/10.6109/jkiice.2023.27.1.34>

print ISSN: 2234-4772 online ISSN: 2288-4165

© 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.
Copyright © The Korea Institute of Information and Communication Engineering.

I. 서론

4차 산업혁명의 진전에 따라 제조 현장에 IT를 접목하여 생산성을 향상시키고 공정을 자동화하려는 스마트 팩토리가 활성화되고 있다[1,2]. 스마트 팩토리에서는 센서를 통하여 수집된 데이터를 인공지능 기술 등을 도입하여 분석함으로써 제조 공정상의 불량률을 최소화한다. 또한, 설비 이상을 사전에 감지해 고장을 예방하고 로봇을 도입하여 생산을 자동화하는 등의 통합 제조공정을 구축하고 있다[3,4].

기존 인쇄 산업에서는 생산자가 고객의 인쇄물을 직접 접수 받아 처리했는데 고객이 생산자의 현황을 파악하여 직접 섭외해야 하는 문제를 갖는다. 이를 해결하기 위하여 스마트 인쇄 시스템에서는 생산자 전면에 위치한 온라인 주문 시스템이 고객의 주문을 접수하고 이를 후면의 생산자에게 전달하는 형태를 갖는다[5]. 생산 공정에서도 다양한 센서와 인공지능 기술을 활용하여 생산성을 높인다. 문제는 온라인 주문시스템에서 고객의 인쇄물에 오류가 있음을 발견하지 못하면 이후의 생산 과정이 모두 무효화된다는 것이다. 특히, 명함의 경우, 주요한 오류는 명함 이미지 회전 오류인데 고객이 온라인 접수할 때 이를 자동으로 검출하는 것이 시스템의 생산성을 크게 높일 수 있다.

이미지 인식 기술은 인공지능 기술 분야에서 빠르게 발전한 분야 중의 하나이다[6]. 인공지능망 모델 중에서도 합성곱 신경망(convolutional neural network)은 이미지로부터 특징을 추출하고 정보를 처리하는데 있어서 높은 성능을 보여 주었다[7]. CNN은 1990년대에 필기 숫자 인식에서 시작하여 현재 자연 이미지에서도 사물을 인식할 수 있는 높은 수준에 이르고 있다[8]. 하지만 CNN은 픽셀 값을 추출하는 원리이므로 이미지가 회전하거나 평행이동하는 경우, 이를 새로운 이미지로 판단하여 인식 성능이 떨어지는 단점이 있다[9].

이를 극복하기 위하여 다양한 심층 신경망(deep neural network) 모델 등이 개발되었다. 인공지능을 활용한 문서 인식 분야에 대한 연구도 활발하게 진행되어서, 문자와 구조의 분리를 통한 문서 인식, 문자 오류 검출, 영상 분류 등에 관한 연구가 진행되었다[10, 11]. 문서 인식, 광학 문자 인식 기술을 활용하여 명함을 자동으로 인식하는 다양한 명함 인식 시스템이 개발되어서 상용화되었다. 이들 명함 인식 시스템은 명함의 이름 및

관련 정보를 인식하여 연락처를 자동 검출하는 것이 주요 목표이므로 본 연구 주제에 적용하기는 어렵다.

본 논문에서는 스마트 명함 인쇄 시스템에서 명함 회전 오류를 검출하는 것을 연구한다. 다양한 심층 신경망 모델을 구현하고, 실험을 통하여 그 성능을 검증한다. 별도의 인공지능 모델을 개발하는 것은 아니고 이미 개발된 심층 신경망 모델들을 활용하여 명함 회전 오류를 검출 가능함을 확인한다. 본 논문에서 사용하는 명함 이미지와 오류인식률은 실제 시스템을 상용화할 중소기업체와 협력하여 구성한다.

2장에서는 스마트 명함 인쇄 시스템의 구조를 설명하고 명함 회전 오류 및 명함 이미지의 특성에 관하여 기술한다. 3장에서는 명함 이미지 데이터 셋에 대하여 기술한다. 4장에서는 명함 회전 오류 검출을 위하여 구성된 심층 신경망 들에 대하여 설명한 후에 실험 결과를 기술한다. 그리고 5장에서 결론을 기술하고 기대 효과를 서술한다.

II. 배경 지식

2.1. 스마트 인쇄 시스템

본 논문의 배경인 스마트 인쇄 시스템의 구조도를 그림 1에서 보여준다. 고객은 클라우드 내의 파일 접수 체계에 인쇄를 원하는 PDF 형식의 파일을 접수한다. 이 문서를 스마트 인쇄 시스템의 관리자가 적절한 인쇄 장치에 배정하고 이를 관리하여 최종적으로 출력된 산출물을 사용자에게 배달한다. 이 시스템을 통하여 인쇄 장치의 생산성을 높이고 사용자에게 고도의 서비스를 제공할 수 있다.

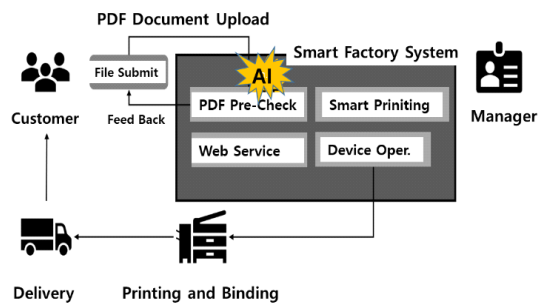


Fig. 1 Architecture of a smart printing system

문제는 그림 1에서 표시한 바와 같이 PDF Pre-check Stage에서 자동으로 오류를 검출해야 하는 것이다. 이미지 파일의 크기나 해상도 등은 PDF 파일의 파라미터 분석을 통하여 가능하다. 하지만, 이미지 자체의 회전 오류는 파라미터 분석으로 감지할 수 없다. 회전 오류가 발생했을 때 최종 결과물이 반전되어 출력되므로 사전에 회전 오류를 검출하는 것이 중요하다.

2.2. 명함 인쇄 시스템의 회전 오류

명함 이미지는 그림, 문자와 더불어 전체적인 구성 디자인이 주요한 요소이다. 명함 이미지는 한글과 영문, 가로와 세로 명함을 모두 포함한다. 명함 이미지의 특징은 명함 템플릿이 존재하지만 글자와 그림의 크기와 배치가 자유롭다는 것이다. 글자가 이미지로 표현되기도 한다. 두 번째의 특징은 명함의 크기가 규격화 되어 있지만 크기가 고유 디자인 등의 이유로 동일하지는 않다. 세 번째로 주제에 따라서 명함 이미지는 그 구성이 자유롭다. 예로서 회사 명함이라면 회사 로고가 들어가고, 식당 명함이라면 음식 사진이 명함의 대부분을 차지한다. 그림 2에 명함 이미지의 예가 보인다.

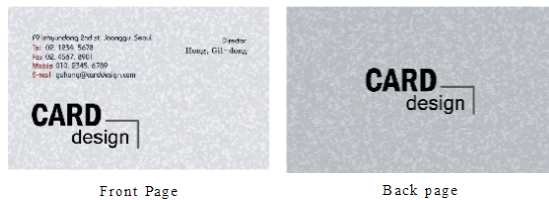


Fig. 2 The example of business card Images

스마트 명함 인쇄 시스템에서 발생하는 명함 이미지 오류는 고객이 인쇄 대상 이미지를 올릴 때에 회전 오류를 포함하는 이미지 회전 오류이다. 정상 명함(0도)에 대하여 90도, 180도, 270도 회전 이미지가 존재하는데, 0도 명함과 180도 회전 명함을 적절하게 올려야 함에도 불구하고 이를 잘못하여서 잘못된 결과를 가져오는 경우가 발생한다. 이를 검출하기 위하여 명함 이미지가 0도, 90도, 180도, 270도 회전 이미지임을 판정하는 것이 필요하다. 이는 기존의 명함 인식 시스템이 명함 소유자 정보를 인식하기 위하여 명함의 문자를 주로 인식하는 것과 다르게 전체적인 구성의 회전 유무를 판단하는 것이 중요하다. 그림 3에 명함 회전 오류 이미지의 예가 보인다. 각 명함에 대하여 앞면과 뒷면의 2개의 이미지가

보이는데, 각각 회전 오류를 포함하고 있다.



Fig. 3 The example of business card Images with a rotation error

인공신경망 모델을 통한 이미지 인식에서 좋은 결과를 얻기 위해서는 충분한 양질의 데이터를 확보하여 훈련하는 것이 필요하다. 명함 이미지는 그 구성이 자유롭기 때문에 단순히 이미지로만 구성되는 경우도 많으며, 회전에서 차이가 보이지 않는 이미지를 사용하는 경우도 흔하게 발생한다. 아래 그림 4에서와 같이 180도 회전에서 차이가 없거나 무의미한 경우가 발생한다. 전체적인 검출 성능 향상을 위해서는 동일한 회전 이미지를 보이는 데이터를 제외하여야 한다.

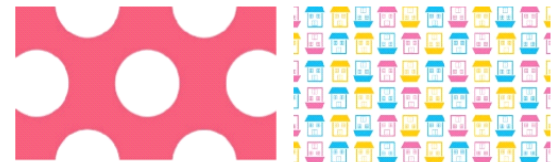


Fig. 4 The example of business card Images with identical images in 180 degree rotation

2.3. 실험에 사용한 심층 신경망 모델들

VGGnet(Visual Geometry Group)은 합성곱 신경망과 ReLU를 활용한 인공신경망이다[12]. 3*3 크기의 합성곱 층을 많이 쌓고 ReLU를 추가해 non-linearity를 높였다. VGGnet은 학습이 쉽고 성능이 우수하여 자주 사용된다. 하지만 파라미터 수가 너무 많아 메모리 사용량이 많다. 본 실험에서는 VGG-16과 VGG-19를 사용하였다. 그림 5은 VGG-16의 구조를 보여준다.

ResNet은 CNN 신경망의 깊이가 깊어짐에 따라서 발생하는 기울기 소실과 폭발 문제를 해결하기 위하여 고

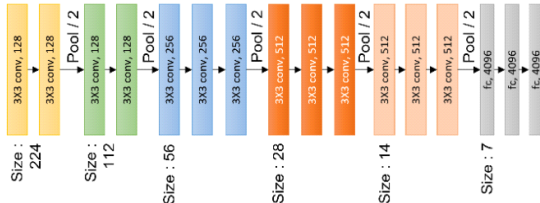


Fig. 5 VGG-16의 구조

안해낸 Residual Learning을 이용하는 마이크로소프트에서 개발한 인공신경망 모델이다[12]. 그림 6은 ResNet 34와 ResNet 50의 구조를 보여준다. 본 실험에서는 ResNet 34와 ResNet 50이 사용되었다.

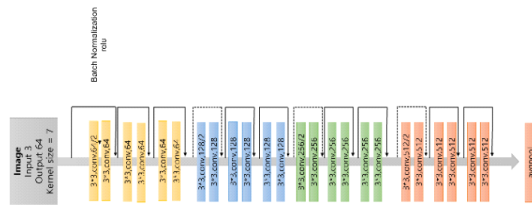


Fig. 6 ResNet-34와 ResNet-50의 구조

DenseNet은 중요한 Feature만 가진 매우 깊은 네트워크를 학습하여 성능 상승을 꾀하는 모델이다. ResNet의 Shortcut 개념을 더 확장하여 CNN 구조를 바꾸는 시도를 하였는데 입력 값을 Summation 하지 않고 Concatenation 하였다. DenseNet은 네트워크에 추가된 정보와 보존된 정보를 명시적으로 구분한다. DenseNet의 layer는 very narrow(e.g. 12 filters per layer)하고 네트워크의 “collective knowledge”에 적은 수의 feature-map set만 추가하며 나머지 feature-map은 변경하지 않는다. 또한, 최종 classifier는 네트워크의 모든 feature-map에 기반하여 결정한다. 전체적인 DenseNet의 구조는 그림7와 같다[13, 14].

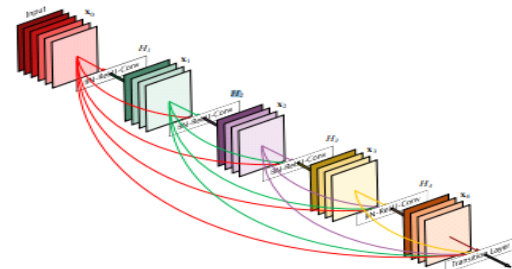


Fig. 7 DensNet의 구조

III. 딥러닝 데이터 준비

3.1. 이미지 수집 및 생성

스마트 명함 인쇄 시스템에서 고객으로부터 접수된 명함 이미지를 적절하게 배치하는 것이 필요하다. 특히 잘못된 명함 이미지가 입력되면 잘못된 명함이 인쇄된다. 이를 해결하기 위하여 명함 회전 유무를 검출하기 위한 전처리 기능을 인공신경망 기반으로 구현한다.

딥러닝 기술로 명함 회전을 인식하려면 훈련 및 시험을 위하여 많은 양의 데이터가 필수적이다. 이를 위하여 스마트명함인쇄시스템을 상용화한 중소기업으로부터 5000장의 실제 명함 이미지를 전달받았다. 이를 이미지 처리 프로그램을 이용하여 90도, 180도, 270도 회전에 따른 4개의 클래스로 명함 이미지를 구성하였다.



Fig. 8 The example of Horizontal English business card rotation



Fig. 9 The example of Vertical English business card rotation

각각의 명함 이미지에 대하여 그림9에 세로 이미지의 명함 90도, 180도, 270도 회전하여서 추가적인 훈련 및 시험 이미지를 확보한다.

본 연구에서는 훈련을 위하여 각 클래스 별로 20,000장의 이미지 데이터를 이용하였으며, 시험을 위하여 각 5,000장의 이미지를 생성하였다.

3.2. 구조적 유사성을 가진 이미지 제거

명함 이미지는 단순히 이미지만을 포함하는 경우도 있는데 이들 이미지들은 그림 4에서와 같이 회전되더라도 차이점을 찾을 수 없다. 이들 이미지들은 훈련과 시험을 위한 데이터베이스에서 제거되어야 한다. 본 연구에서는 Structural Similarity 방식[15]을 이용하여 두 이미지의 유사도를 비교하기 위하여 scikit-image 라이브러리를 사용하였고 유사도 결정 척도는 통계적 유의성을 고려하여 0.95를 사용하였다. 처음 얻어진 80,000개의 훈련 데이터와 20,000개의 시험 데이터에서 유사도 비교를 통하여 동일 이미지 데이터를 제외한 후 60,000개(각 클래스별 15,000개)의 훈련 이미지와 18,000개(각 클래스별 4,500개)의 시험 이미지를 사용하여 실험을 진행하였다.

IV. 실험 및 평가

4.1. 실험 구성

인공신경망의 학습을 위하여 표 1과 같은 시험 환경을 구축하였다.

Table. 1 Experimental environments

Unit	Specification
OS	Ubuntu 18.04.5 LTS
GPU	GeForce RTX 3080 1대
DL	python 3.7.10, pytorch 1.8.0

먼저, 명함 이미지를 224*224*3의 크기로 변환한다. 그리고 다양한 합성곱 신경망(convolution neural network) 모델인 VGG, ResNet, DenseNet 모델을 이용하여 그 성능을 비교한다. 최적화 함수는 Adam(adaptive moment assesment) 최적화 알고리즘을 사용하며 손실함수는 다중 분류를 위한 범주형 교차 엔트로피를 사용한다. 하이퍼 파라미터를 결정하기 위하여 learning rate는 0.001~

0.01, epoch는 50~100, 으로 변경하여 실험하였는데 learning rate, 0.01, epoch, 50에서 안정적으로 수렴하였다.

최종 분류를 위하여 3 계층의 fully connected layer를 사용하며 activation function으로 ReLU를 사용한다. 각각의 심층 신경망 모델에서 제공하는 pre-trained weight 값을 사용하였다.

4.2. 실험 결과

합성곱 신경망 모델과 심층 신경망 모델의 성능을 비교, 분석하기 위하여 각 클래스별 20,000장에 대하여 합성곱 신경망(convolution neural network)과 ResNet-34에 대하여 학습한 후, 0도 이미지 및 180도 이미지, 각 5,000장에 대하여 시험을 진행하였다. 성능 지표로는 accuracy를 사용하였으며 아래 표 2에 그 결과를 표시하였다. CNN에 비교하여서 ResNet-34가 보다 좋은 성능을 보이는 것을 확인하였다.

Table. 2 Performance comparison of CNN and ResNet-34

	Model	Accuracy
1	CNN	92.3%
2	ResNet-34	93.4%

Structural Similarity 비교를 통하여 동일한 회전 이미지를 제거하여 얻어진 60,000개(각 클래스별 15,000개) 이미지에 대하여 3 종류의 심층 신경망에 대하여 학습을 진행한 후에 각 클래스별 4,500개 이미지에 대하여 실험하였다. 표 3에서와 같이 DenseNet-161에서 가장 좋은 성능을 얻었으나, 모델의 복잡성을 감안하면 ResNet-34에서도 충분한 성능을 확보한다고 할 수 있다. 표2에서의 결과와 비교하면 동일한 회전 이미지를 제거함으로써 성능이 향상되는 결과를 가져옴을 알 수 있다.

Table. 3 Accuracy by model

Model		Accuracy(%)
VGG	16	97.6
	19	97.2
ResNet	34	97.2
	50	97.1
DenseNet	121	97.8
	161	97.9

V. 결 론

명함 회전 오류의 검출을 통하여 스마트 명함 인쇄 시스템의 전반적인 생산성 향상을 가져올 수 있다. 온라인 스마트 명함 인쇄 시스템은 고객이 명함 이미지를 업로드하면서 시작한다. 이때, 부적절하게 회전된 명함 이미지를 업로드하여 잘못된 최종 산출물을 출력하는 오류가 발생한다. 이를 제거하기 위하여 업로드한 명함 이미지의 회전 상태 분류가 필요하다.

명함 이미지는 이미지, 문자 등으로 자유롭게 구성되어 있어서 회전 상태를 파악하기 위하여 전반적인 구성을 인식하는 것이 필요하다. 이를 위하여 본 논문에서는 새로운 인공 지능 모델을 개발하지 않고 이미 개발된 심층 신경망 모델들을 사용하여도 충분한 유용성을 확보할 수 있었다.

본 논문은 중소기업이 상용화한 스마트 인쇄시스템에 적용한 것으로 90% 이상의 인식률을 요구하였으나 이를 충분히 만족하였다. 구현한 명함 회전 오류 검출 시스템은 스마트 명함 인쇄 시스템에 실제적으로 적용되어서 그 효과를 검증할 수 있었다.

향후연구방향은 좀 더 간단한 심층 신경망 모델을 사용하여 명함 회전 오류를 검출할 수 있는지를 검증하는 것이다. 이는 본 논문의 연구 과정에서 간단한 모델이 학습 과정에서나 적용 과정에서 더 짧은 시간안에 처리가 가능하다는 것을 체감함에 따른 것이다.

ACKNOWLEDGEMENT

This work was supported by 2021 Hannam University Research Fund.

REFERENCES

- [1] E. J. Park and B. H. Ha, "A Formal Framework for Analyzing Performance of Container Terminal Operations," *Journal of Society for e-Business Studies*, vol. 18, no. 2, pp. 191-203, May 2013. DOI: 10.7838/jsebs.2013.18.2.191.
- [2] R. Seiger, L. Malburg, B. Weber, and R. Bergmann, "Integrating process management and event processing in smart factories: A systems architecture and use cases," *Journal of Manufacturing systems*, vol. 63, pp. 575-592, Apr. 2022. DOI: 10.1016/j.jmsy.2022.05.012.
- [3] H. U. Park, "Trends in production and manufacturing technology related to smart factories," *Information and Communications Magazine*, vol. 33, no. 1, pp. 24-29, Dec. 2015.
- [4] K. S. Ko, J. J. Huh, and J. I. Oh, "A Study on the Factors that Affect the Adoption of a Smart Factory - Focusing on the Comparison between Customers and Suppliers," *Korea Business Review*, vol. 25, no. 3, pp. 129-151, Aug. 2021. DOI: 10.17287/kbr.2021.25.3.129.
- [5] D. Y. Son and K. K. Lee, "A Study on the Recognition of Face Based on CNN Algorithms," *Korean Journal of Artificial Intelligence*, vol. 5, no. 2, pp. 15-22, Dec. 2017. DOI: 10.24225/kjai.2017.5.2.15.
- [6] Y. S. Kwon, D. J. Shin, and J. J. Kim, "A Study on Application Method of Contour Image Learning to improve the Accuracy of CNN by Data," *The Journal of The Institute of Internet, Broadcasting and Communication (IIBC)*, vol. 22, no. 4, pp. 171-176, Aug. 2022. DOI: 10.7236/IIBC.2022.22.4.171.
- [7] J. W. Kim, H. Pyo, J. Ha, C. Lee, and J. Kim, "Various deep learning algorithms and applications," *Communications of the Korean Institute of Information Scientists and Engineers*, vol. 33, no. 8, pp. 25-31, Aug. 2015.
- [8] B. M. Kim, "Trend of image classification technology based on deep learning," *Korea Institute of Communication Sciences*, vol. 35, no. 12, pp. 8-14, Nov. 2018.
- [9] D. Lee, and et al., "CNN-based Image Rotation Correction Algorithm to Improve Image Recognition Rate," *The Journal of The Institute of Internet, Broadcasting and Communication*, vol. 20, no. 1, pp. 225-229, Feb. 2020.
- [10] A. Das, S. Roy, U. Bhattacharya, and S. K. Parui, "Document Image Classification with Intra-Domain Transfer Learning and Stacked Generalization of Deep Convolutional Neural Networks," in *Proceedings of 2018 24th International Conference on Pattern Recognition (ICPR)*, Beijing, China, pp. 3180-3185, 2018. DOI: 10.1109/ICPR.2018.8545630.
- [11] X. Deng, Y. Zhang, S. Yang, P. Tan, L. Chang, Y. Yuan, and H. Wang, "Joint Hand Detection and Rotation Estimation Using CNN," *IEEE transactions on image processing*, vol. 27, no. 1, pp. 1888-1900, Apr. 2018. DOI: 10.1109/TIP.2017.2779600.
- [12] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," in *Proceedings of conference paper at ICLR 2015*, San Diego:

- CA, USA, pp. 7-9, 2015.
- [13] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas: NV, USA, pp. 770-778, Jun. 2016.
- [14] G. Huang, Z. Liu, L. and K. Q. Weinberger, "Densely Connected Convolutional Networks," in *Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu: HI, USA, pp. 4700-4708, Jul. 2017.
- [15] D. Brunet, E. R. Vrscay, and Z. Wang, "On the Mathematical Properties of the Structural Similarity Index," *IEEE transactions on image processing*, vol. 21, no. 4, pp. 1488-1499, April. 2012.



경지훈(Ji-Hoon Kyung)

2004 한남대학교 산업공학과 학사
2006 한남대학교 산업공학과 공학석사
2014 한남대학교 산업공학과 공학박사
2021-현재 한남대학교 산업경영공학과 조교수
※관심분야 : 데이터사이언스, 딥러닝, 정보처리