

# 대형 이미지 데이터셋 구축을 위한 이미지 이진화 기반 데이터 증강 기법

## Data augmentation technique based on image binarization for constructing large-scale datasets

이주혁\*, 김미희\*

Lee-JuHyeok\*, Kim-Mi Hui\*

### Abstract

Deep learning can solve various computer vision problems, but it requires a large dataset. Data augmentation technique based on image binarization for constructing large-scale datasets is proposed in this paper. By extracting features using image binarization and randomly placing the remaining pixels, new images are generated. The generated images showed similar quality to the original images and demonstrated excellent performance in deep learning models.

### 요약

딥러닝은 다양한 컴퓨터 비전 문제를 해결할 수 있지만, 대량의 데이터셋이 필요하다. 본 논문에서는 대형 이미지 데이터셋을 구축하기 위해 이미지 이진화 기반 데이터 증강 기법을 제안한다. 이미지 이진화를 사용하여 특성을 추출하고 추출된 나머지 픽셀을 랜덤하게 배치하여 새로운 이미지를 생성한다. 생성된 이미지는 원본 이미지와 유사한 품질을 보여주며, 딥러닝 모델에서도 뛰어난 성능을 보였다.

*Key words : data preprocessing, data augmentation, image processing, deep learning, image classification*

### 1. 서론

딥러닝을 통해 컴퓨터 비전 분야에서 이미지 분류, 객체 탐지, 세분화 등 다양한 문제를 해결할 수 있게 되었다[1]. 이러한 기술들은 대량의 데이터셋을 필요로 하므로, 데이터의 양이 많을수록 성능이 향상되는 경향을 보

인다. 하지만, 현실적으로 대량의 데이터를 수집하기 어려울 때가 많기 때문에, 데이터 증강 기술이 주목받고 있다.

최근 데이터 증강 기술에 관한 많은 연구가 진행되었으며, 이러한 연구들은 크게 다음과 같은 방향으로 진행되고 있다.

\* School. of Computer Engineering & Applied Mathematics, Hankyong National University

★ Corresponding author

Email : mhkim@hknu.ac.kr, Tel : +82 31-670-5167

※ Acknowledgment

This research was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No.2018R1A2B6009620)

Manuscript received Feb. 27, 2023; revised Mar. 15, 2023; accepted Mar. 20, 2023.

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.

첫째, 데이터 증강 기술의 성능 개선을 위한 기술적 연구가 이루어지고 있다. 예를 들어, 데이터 증강 기술 중 하나인 데이터 증강 기술은 기존의 이미지 데이터를 변형시켜 새로운 데이터를 생성하는 기술로, 이를 통해 딥러닝 모델의 성능을 향상하는 방법이 연구되고 있다.

둘째, 데이터 증강 기술의 다양한 활용 분야가 연구되고 있다. 예를 들어, 의료 이미지에서 병변 영역을 자동으로 검출하거나, 자율주행 자동차에서 주변 환경을 인식하는 데에 활용될 수 있다.

셋째, 데이터 증강 기술을 활용한 다양한 응용 프로그램이 개발되고 있다. 예를 들어, 스마트폰 카메라 어플리케이션에서 얼굴 인식 기술과 함께 이모티콘을 삽입하는 기술이 이미 상용화되어 있다[1].

데이터 증강 기술이 발전하면서, 학습 모델에 대한 데이터의 양과 질이 중요한 역할을 한다. 특히 딥러닝 모델의 경우, 수많은 이미지 데이터셋으로부터 다양한 특성들을 학습하여 모델의 정확도를 향상한다.

데이터 증강 기술은 기존의 이미지 데이터셋에서 새로운 이미지 데이터를 생성하거나, 기존 이미지를 변형하여 새로운 데이터를 만들어내는 방식이다. 이를 통해 데이터의 양을 증가시키고, 모델의 일반화(generalization) 성능을 향상할 수 있다[2].

본 논문에서는 이미지 이진화를 통한 최대 빈도 측 기반 랜덤 재배치 이미지 생성 알고리즘을 통해 컴퓨터 비전 분야에서의 학습 모델 성능을 향상하는 방법을 제안한다. 이를 통해 데이터 증강 기술이 학습 모델 성능에 미치는 영향을 분석하고, 데이터 증강 기법에 대한 유효성을 증명한다. 또한 데이터셋의 크기를 증가시켜 성능 향상을 도모할 수 있다.

## II. 배경 지식

### 1. 이미지 증강 기술

이미지 증강 기술은 이미지를 수정 혹은 변형하여 원본 이미지보다 더 많은 이미지를 생성하는 기술이며, 대형 데이터셋이 필수적으로 필요한 딥러닝 분야에서 활발하게 사용 중이다. 이미지 증강은 데이터의 크기, 색상, 밝기, 대비, 회전 등으로 이미지를 수정하여 생성한다. 다른 이미지 증강 기술들의 예시로는 랜덤 크롭(Random Crop), 랜덤 플립(Random Flip), 컬러 제로(Cutout), 블러링(Blur), 색조 변경(Hue Shift) 등 기술들이 있다 [3]~[6].

각 이미지 증강 기술들은 서로 다른 방식으로 이미지

데이터를 변형한다. 예를 들어, 랜덤 크롭은 이미지에서 랜덤한 위치를 선택하여 해당 영역을 잘라내는 기술이다.

이미지 증강 기술은 이미 다양한 연구에서 다뤄졌다.. 이전 연구에서는 회전, 크기 조정, 밝기 조절, 색감 변화 등을 사용해 이미지 데이터를 증강하는 다양한 방법이 연구되어왔다. 최근에는 GAN(Generative Adversarial Networks)과 같은 새로운 기술이 등장하면서, 더욱 다양한 방법으로 이미지 증강이 가능해졌다[7].

### 2. 기존 기술

Pixel-Level Transform, PatchShuffle Regularization, Random Erasing Data Augmentation 등[8] 여러 이미지 증강 기술이 있다.

Pixel 단위로 변환을 시키는 Pixel-Level Transform은 대표적으로 Blur, Jitter, Noise 등을 이미지에 적용하는 기법이다. Gaussian Blur, Motion Blur 등 다양한 기법이 Pixel 단위에 사용해 이미지 데이터를 증강한다.

PatchShuffle Regularization은 feature 값들을 랜덤하게 섞는 기법을 제안했으며, 임의의 크기를 지정하여 그 크기만큼 랜덤하여 섞기 때문에 임의의 크기에 따라 성능이 달라진다.

Random Erasing Data Augmentation은 Patch-Shuffle과는 다르게 랜덤한 바운딩 박스를 만들어 그 안을 임의의 값으로 채우거나, 원본 이미지의 평균, 0 또는 255 등으로 바운딩 박스를 채우는 채워서 이미지를 증강하는 기법이다.

앞서 소개한 방법들처럼 본 논문에서 원본 이미지의 특성을 이용해 새로운 이미지를 생성하는 기법을 제안한다.

## III. 제안 증강 시스템

### 1. 이미지 이진화(Image binarization)

이미지 이진화는 임의의 값인 임계값을 지정해 임계값보

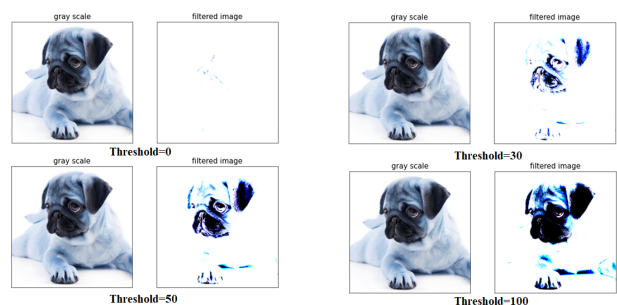


Fig. 1. Image Binarization based on Thresholding.

그림 1. 임계값에 따른 이미지 이진화

다 밝은 픽셀들은 모두 흰색으로, 그렇지 않은 픽셀들은 모두 검은색으로 바꾸는 기법이다[9]. Fig. 1는 임의의 임계값을 기준으로 새로 생성된 이미지들의 예시이다.

이미지 이진화를 통해서 사물 혹은 물체의 특성을 추출 및 검출에 사용이 될 수 있다. 이는 컴퓨터 비전, 의료 비전, 자동차 번호판인식 등 영상처리에 기초가 되는 기법의 하나다. 이진화를 통해서 이미지 내의 특성을 추출 및 검출 할 수 있으므로, 추출된 나머지 이미지들을 랜덤 배치함에 따라, 새로운 이미지를 만들어 낼 수 있다.

하지만 이 임계값은 Fig. 1.처럼 수작업으로 임의의 임계값을 하나하나 정할 수 없다. 이를 해결하기 위해 오츠 알고리즘(Otsu's algorithm)[10]을 이용해 해결한다. 오츠 알고리즘은 이미지 이진화 알고리즘 중 하나로 이미지의 히스토그램 분석을 통해서 자동으로 이진화 임계값을 계산하는 기법이다. Fig. 2의 오츠 알고리즘에 사용되는 히스토그램의 예시이다.

본 논문은 이 아이디어를 바탕으로 새로운 기법을 제시한다.

## 2. 제안 증강 시스템

Fig. 3은 제안하는 증강 시스템의 알고리즘이다. 크게 6단계를 통해서 새로운 이미지를 생성한다.

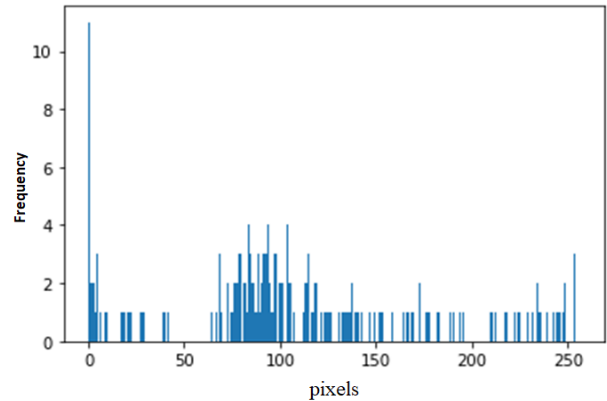


Fig. 2. Example of Otsu's algorithm histogram.

그림 2. 오츠 알고리즘 히스토그램 예시

1. 이미지를 입력받은 후 이미지 행렬인  $M$ 을 이진화를 한다.
2.  $M$ 의 크기를 계산한다.  $M$ 의 크기를 계산하는 이유는 최대 빈도 축에 따라서 정해진 이미지 내의 특성을 고정하기 위해서  $M$ 의 크기를 계산하고, 계산된 특성이 원본 이미지에 벗어나는지 판단하기 위해  $M$ 의 크기를 계산한다.
3.  $M$ 의 각행과 열에서 255의 개수가 가장 많은  $x$ 축과  $y$ 축을 찾는다. 255를 기준으로 행렬에서 빈도수를 체크하는 이유는 이진화를 통해서 255나 0으로 구분되기 때문에 255를 통해서 이미지내의 특성을 추출하기 위함이다.
- 3-1.  $M$ 의 각 행과 열에서 255의 개

---

### Algorithm 1: Random rearrangement image generation technique based on maximum frequency axis

---

| Input: Image matrix  $M$  |

| Output: Randomly shuffled image matrix  $R$  |

1.Convert  $M$  to binary image.

2.Compute the size of  $M$ .

3.Find the  $x$ -axis and  $y$ -axis with the most 255s.

3-1. Compute the number of 255s in each row and each column of  $M$ .

3-2. Select the  $x$ -axis and  $y$ -axis with the highest count of 255s.

4.Define a bounding box around the selected  $x$ -axis and  $y$ -axis.

4-1. Compute the coordinates of the bounding box as follows:  $x\_min$  = the minimum index of the selected  $x$ -axis with a 255 value  $x\_max$  = the maximum index of the selected  $x$ -axis with a 255 value  $y\_min$  = the minimum index of the selected  $y$ -axis with a 255 value  $y\_max$  = the maximum index of the selected  $y$ -axis with a 255 value

4-2. Define the bounding box  $R$  as a submatrix of  $M$  with the coordinates computed above.

5.Shuffle the values of  $M$  outside of the bounding box  $R$ .

5-1. Create a copy of  $M$  called  $R$ .

5-2. for  $i=0$  to  $M$ 's row size-1: for  $j=0$  to  $M$ 's column size-1: if  $(i,j)$  is not inside  $R$ : randomly shuffle the value of  $R[i,j]$  with another value outside of  $R$

6.Return the shuffled image matrix  $R$ .

---

Fig. 3. Random rearrangement image data generation technique based on maximum frequency axis.

그림 3. 이미지 이진화를 통한 최대 빈도 축 기반 랜덤 재배치 이미지 데이터 생성 기법

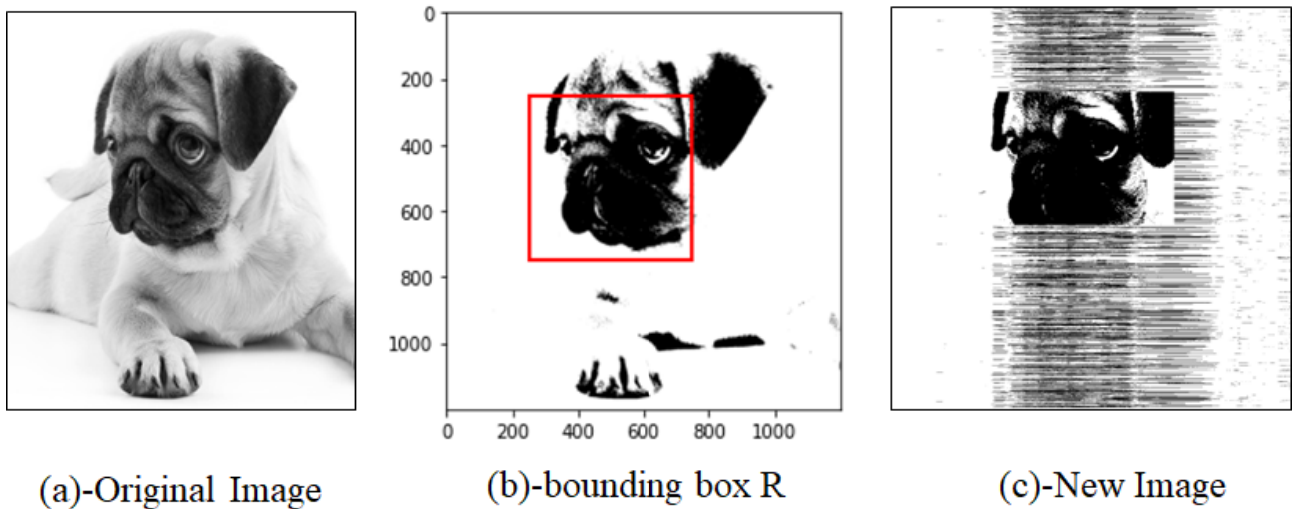


Fig. 4. Example of proposed system.  
그림 4. 제안 시스템 적용 예시

수를 계산 후, 3-2.255의 개수가 가장 많은 x축과 y축을 계산하여 선택한다.

4. 선택된 x축과 y축을 기준으로 바운딩 박스를 정의한다. 4-1. 축의 바운딩 박스의 좌표는 255값이 가장 많은 축의 최소인덱스와 최대인덱스를 비교한다. 4-2. 4-1.에서 계산된 좌표를 바탕으로 바운딩 박스 R을 정의한다. 5. 바운딩 박스 R을 제외한 M의 행렬값을 랜덤하게 섞는다. 6. 랜덤하게 섞인 새로운 이미지 행렬인 R을 반환한다.

#### IV. 실험 결과 및 분석

##### 1. 이미지 증강

이미지 증강에 필요한 이미지 데이터셋은 ImageNet dataset[11] 중에서 강아지 데이터를 이용해 이미지 증강을 진행했다. 이미지 데이터는 강아지 데이터 셋 중 500장을 랜덤으로 선택했다.

이미지 생성은 Fig. 3의 과정을 거쳤으며, Fig.4-(a)가 원본 이미지 Fig.4-(c)가 새로 생성된 이미지로 이미지 특성을 추출하여 특성을 제외한 나머지 픽셀들을 랜덤 배치했다. 500장과 이미지 증강을 거쳐 총 1,000장의 이미지데이터 셋을 구축했다.

##### 2. 이미지 품질평가

이미지 증강을 거친 이미지가 실질적으로 사용이 가능한지 판단하기 위해서 이미지 품질평가를 실시한다. 이미지 품질 평가는 SSIM, MS-SSIM, VIF[12]-[14]을 이용해 품질평가를 진행한다.

SSIM, MS-SSIM, VIF로 비교하려는 이유는 Table. 1 처럼 원본이미지와외의 구조적 유사성과 이미지의 해상도가 어떻게 변했는지 확인하기 위해서다.

Table 1. SSIM, MS-SSIM, VIF Comparison.

표 1. SSIM, MS-SSIM, VIF비교

	overview	Calculation Method
SSIM	comparison of structural similarity	structural similarity
MS-SSIM	Improve the shortcomings of SSIM, calculate structural similarity for various scales	Adding a process of applying SSIM calculation to various scales
VIF	Compare image quality before and after applying image quality improvement and compression	Calculation of variance ratio between pixels

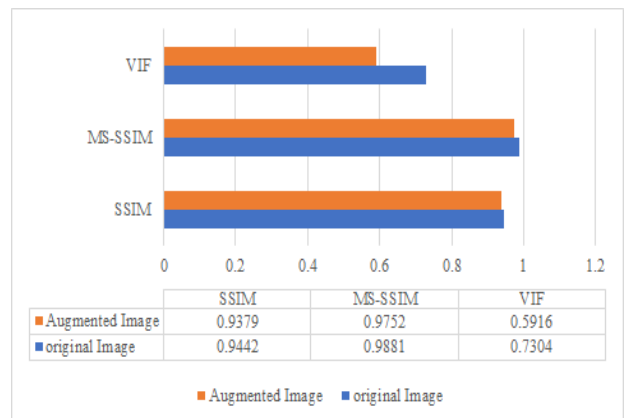


Fig. 5. image quality assessment results.

그림 5. 이미지 품질 평가 결과

Fig. 5처럼 원본 이미지와 증강된 이미지에 대한 SSIM과 MS-SSIM의 값은 유사한 결과를 보였다. 하지만 VIF 값은 원본 이미지보다 품질이 떨어진다고 판단하였고, 원본 이미지가 증강 이미지보다 더 세부 정보를 가지고 있다는 의미이다. 하지만 SSIM, MS-SSIM, VIF 비교 결과 원본 이미지와 품질 평가값이 유사하기 때문에 원본 이미지와 차이가 존재는 하나, 이미지의 특성은 살려 이미지 증강을 한 것을 알 수 있다.

### 3. 딥러닝 모델

이미지 품질평가를 통해 원본이미지와는 차이가 있지만, 원본 이미지의 특성과 같은 특성이 증강된 이미지 데이터에도 포함된 것을 파악했다. 이를 바탕으로 원본 이미지와 증강이미지로 데이터셋을 구축하여 딥러닝 모델에 적용했다.

딥러닝은 대형 데이터셋이 필수적이다. 딥러닝 모델을 통해 증강된 데이터를 포함시켜 실험을 진행하는 경우 원본이미지 데이터만 사용했을 경우 모델의 성능을 비교한다. 딥러닝 모델 중 이미지 분류 모델인 CNN과 객체 탐지 모델인 YOLOv5[15]를 이용해 실험을 진행했다. CNN 모델은 ResNet-50[16]의 구조를 이용해 모델을 설계 한 후 실험을 진행하였다. 실험 결과 Fig. 6처럼 원본이미지와 증강된 이미지가 섞인 데이터셋으로 약 97%의 정확도로 원본이미지만 사용했을 때보다 약 2% 상승된 효과를 볼 수 있었다.

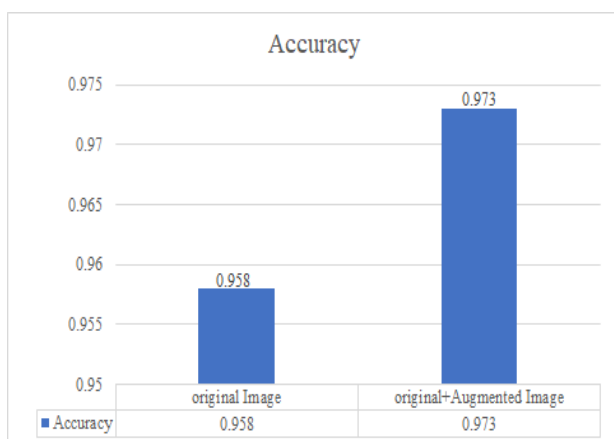


Fig. 6. Image classification result.

그림 6. 이미지 분류 결과

객체 탐지 결과 역시 이미지분류모델과 비슷하게 원본 이미지와 증강된 이미지가 합쳐진 데이터셋이 더 좋은 결과를 보여준다.

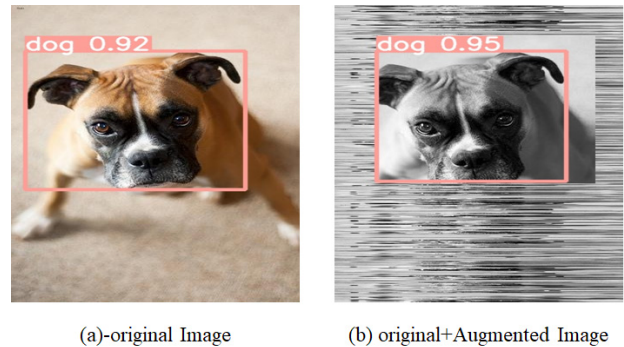


Fig. 7. object detection result.

그림 7. 객체 탐지 결과

## V. 결론 및 향후 연구

본 논문에서는 대형 데이터 셋 구축을 위한 이미지 이진화 기반 이미지 증강 기법을 제안 했다.

실험 결과 제안 기술을 통해 생성된 이미지는 원본 이미지와 품질평가에서 비슷한 결과를 보여줬고, 이미지 분류 모델과 객체 탐지 모델 즉 딥러닝 모델에서도 좋은 성능을 보였다.

향후 연구에서는 최대 빈도 축을 기반으로 바운딩 박스를 생성하는데 바운딩 박스의 개수를 늘려 이미지 내의 특성을 최대한 살리고, 바운딩 박스의 개수를 효율적으로 찾는 방안으로 연구하고자 한다. 또한 컬러이미지와 특정 이미지에 따라 본 논문 기법을 적용하지 못하는 경우에 대한 방안을 연구하고자 한다.

## References

- [1] Tang, H., Xu, D., Sebe, N., & Wang, Y., "A survey on multimodal deep learning for image synthesis: Applications, methods, datasets, evaluation metrics, and results comparison," *IEEE Access*, Vol.8, pp.919-145, 2020. DOI: 10.1145/3461353.3461388
- [2] Shorten, C., & Khoshgoftaar, T. M., "A survey on image data augmentation for deep learning," *Journal of Big Data*, Vol.6, No.1, pp.1-48, 2019. DOI: 10.1186/s40537-019-0197-0
- [3] Lin, T.-Y., Goyal, P., Girshick, R., He, K., & Dollar, P. "Focal loss for dense object detection," *Proceedings of the IEEE international conference on computer vision*, pp.2980-2988, 2018.

DOI: 10.48550/arXiv.1708.02002

[4] Zhong, Z., Zheng, L., Kang, G., Li, S., & Yang, Y. "Random erasing data augmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.43, No.2, pp.564-578. 2018.

DOI: 10.48550/arXiv.1708.04896

[5] Gao, Y., Feng, J., & Krahenbuhl, P. "Augmenting supervised neural networks with unsupervised objectives for large-scale image classification," *Proceedings of the IEEE International Conference on Computer Vision*, pp.2624-2633, 2019.

DOI: 10.48550/arXiv.1606.06582

[6] Zhao, H., Zhang, L., Liu, M. Y., Shao, J., & Kautz, J. Differentiable "augmentation for data-efficient GAN training," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp.6489-6498, 2020.

DOI: 10.48550/arXiv.2006.10738

[7] Siddiquee, M. R., et al., "AugmentGAN: Deep Learning Augmentation for Medical Image Segmentation," *IEEE Access*, 7, 113305-113314. 2019.

[8] Shorten, C., & Khoshgoftaar, T. M. "A survey on image data augmentation for deep learning," *Journal of Big Data*, Vol.6, No.1, pp.1-48. 2019.

DOI: 10.1186/s40537-019-0197-0

[9] Niblack, W., & Sauvola, J. "Adaptive thresholding using the integral image," *Proceedings of the workshop on Document Analysis Systems*, pp.21-28, 1994.

[10] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol.9, no.1, pp. 62-66, 1979. DOI: 10.1109/TSMC.1979.4310076

[11] Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. "ImageNet: A Large-Scale Hierarchical Image Database," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009.

DOI: 10.1109/CVPR.2009.5206848

[12] Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P., "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, Vol.13, No.4, pp.600-612. 2004. DOI: 10.1109/TIP.2003.819861

[13] Wang, Z., Simoncelli, E. P., & Bovik, A. C.,

"Multiscale structural similarity for image quality assessment," *Proceedings of the Thirty-Seventh Asilomar Conference on Signals, Systems & Computers*, Vol.2, pp.1398-1402, 2003.

DOI: 10.1109/ACSSC.2003.1292216

[14] Sheikh, H. R., & Bovik, A. C. "Image information and visual quality," *IEEE Transactions on Image Processing*, Vol.15, No.2, 2006.

DOI: 10.1109/TIP.2005.859378

[15] Jocher, Glenn, et al. "YOLOv5: A Universal Object Detector Made for Production," *arXiv preprint arXiv:2104.02123*, 2021.

[16] He, Kaiming, Zhang, Xiangyu, Ren, Shaoqing, Sun, Jian. "Deep Residual Learning for Image Recognition," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, vol.2016, pp.770-778, 2016.

DOI: 10.1109/CVPR.2016.90

## BIOGRAPHY

### Lee-Ju Hyeok (Member)



2022 : BS degree in Computer Science and Engineering, Hankyong National University

2022~present : MS student in School of Computer Engineering & Applied Mathematics, Hankyong National University

### Kim-Mi Hui (Member)



1997 : BS degree in Computer Science and Engineering, Ewha Womans University.

1999 : MS degree in Computer Science and Engineering, Ewha Womans University.

1999~2003 : Researchers at Switching & Transmission Technology Lab.(ETRI)

2007 : Ph.D. degree in Computer Science and Engineering, Ewha Womans University.

2009~2010 : postdoctoral researcher of the department of computer science, North Carolina State University  
2011~present : School of Computer Engineering & Applied Mathematics, Computer System Institute, Hankyong National University.