

# 딥러닝 기반 실시간 이미지 검출 시스템에 의한 한국 전통회화 객체 인식 연구

박수영<sup>1</sup>, 박찬우<sup>2</sup>, 김희권<sup>2</sup>, 이재호<sup>2\*</sup>

<sup>1</sup>성결대학교 정보통신공학과 학생, <sup>2</sup>한국전자통신연구원 홀로그래픽콘텐츠초연구실 연구원

## A Study on the Recognition of Korean Traditional Picture Objects by Deep Learning-Based Real-Time Image Detection System

Soo-Young Park<sup>1</sup>, Chan-Woo Park<sup>2</sup>, Hee-Kwon Kim<sup>2</sup>, Jae-Ho Lee<sup>2\*</sup>

<sup>1</sup>Student, Dept. of Information and Communication, Sungkyul University

<sup>2</sup>Researcher, Lab. of Holographic Contents, ETRI

**요약** 인류의 새로운 생활 형식으로 각광받고 있는 가상 공간에서의 활동을 위해, 전 세계 기업과 연구기관에서는 가상 공간 내부에 존재하는 사물을 생성 및 동식물 환경을 조성하고 설정하는 것에 초점을 두고 있다. 가상 공간의 데이터 객체(Objects)들은 인간의 오감을 극대화할 수 있도록 보다 사실적인 모델링과 피드백 및 이를 통한 후속적인 무브먼트가 향상될 필요성이 있다. 현실의 물리적 객체를 가상의 디지털 데이터로 전사하는 연구의 전처리 단계로 2차원 특정 객체 인식을 선행으로 요구한다. 본 연구는 딥러닝 기반 실시간 이미지 검출 시스템을 이용하여 객체 인식이 향상될 필요가 있는 회화 분야에서의 선행 작업 퍼포먼스를 향상시키는 것을 의도하였다. 현대 미술품들의 파라미터들과 차이가 있어 상대적으로 시행 데이터가 부족한 한국 전통 미술품들을 타깃으로 하여 미술품 내 객체 인식을 향상을 확인하였다. 입체적인 모델링을 형성하는 데에 있어 정확성을 높이는 연구가 추후 진행되어야 할 것이다.

**키워드** : 디지털 콘텐츠, 홀로그래픽, 딥러닝, 이미지 검출, 객체 인식, 한국 전통회화

**Abstract** For activities in virtual space, which is in the spotlight as a new lifestyle of mankind, companies and research institutes around the world focus on creating objects inside the virtual space and creating and setting up animal and plant environments. Data objects in virtual space need to improve more realistic modeling and feedback and subsequent movements to maximize human five senses. Require two-dimensional specific object recognition as a preprocessing step for research to transfer real physical objects to virtual digital data. This study intended to improve the performance of prior work in the field of conversation, which needs to improve object recognition, by using a deep learning-based real-time image detection system. It was confirmed that the object recognition rate in the artwork was improved by targeting traditional Korean art that lacked relatively implementation data due to differences from the parameters of modern art. Research to increase accuracy in forming three-dimensional modeling should be conducted later.

**Key Words** : Digital Content, Holographic, Deep Learning, Image Detection, Object Recognition, Korean Traditional Picture

\*Corresponding Author : Jae-Ho Lee(jhlee3@etri.re.kr)

Received October 20, 2022

Accepted December 24, 2022

Revised November 11, 2022

Published December 28, 2022

## 1. 서론

증강 현실과 가상 현실 기술이 대중화되고 있는 시점에서 해당 기술을 응용한 미디어 플랫폼과 콘텐츠 프로그램 수요가 높아지고 있다. 여러 산업 분야에서 미디어와 콘텐츠를 서비스하기 위한 시도를 지속 중에 있으며, 그중에서도 미술 분야에 대한 접목이 주목 받고 있는 추세이다. 하지만 서구권 근현대 미술품들에 대하여 활발한 확장 현실 접근이 이루어지고 있는 흐름에서 한국의 전통회화 분야에 대한 접근은 부족한 상황이다. 대체적으로 한국 전통회화 미술품들은 서구권 근현대 미술품과는 미술적 파라미터들에서 차이가 있기 때문에 컴퓨팅 테크놀로지를 이용한 접근에 어려움이 있다. 점, 선, 면에서 비롯되는 명암, 색채, 질감과 더 나아가 형태, 입체를 이루는 공간적인 조형 요소가 서구권 근현대 미술품과는 성격이 다르기에 이것을 보완하여 컴퓨터가 한국 전통회화의 접근을 용이할 수 있도록 해야 한다.

현실의 물리적 사물을 가상의 데이터로 전사하는 작업은, 먼저 그 사물의 전반적인 정보들을 인식하고 종합하는 것을 요구한다. 한국의 전통회화들은 종이와 비단 등 2차원 평면 상에 그려진 미술품들이 대부분이므로 스캔과 촬영을 통해 이미지 파일을 확보한 후, 인공지능을 사용한 딥러닝 기반 실시간 이미지 검출 시스템을 통하여 회화 상에 그려진 특정 객체들을 인식할 수 있다. 이런 객체 인식 작업을 진행하여 추출한 특정 객체들은 후처리 과정을 거쳐 증강 및 가상 현실 환경에 하나의 그래픽 모델로 생성되어 사용자들에게 서비스될 수 있다. 처음 데이터가 생성되는 미술품 스캔과 촬영이 초고화질로 이루어진 후, 전처리 단계인 객체 인식 과정이 정확하게 진행되어야 후에 그래픽 작업과 모델링 생성 작업이 원활하게 이루어진다. 처음 단계 과정들의 퀄리티들을 기반으로 결과적인 콘텐츠의 퍼포먼스가 결정되기 때문에 정확하고 일정한 객체 인식이 이루어져야 한다[3].

따라서 본 연구에서는 딥러닝 기반 실시간 이미지 검출 시스템을 활용하여 한국 전통회화 객체 인식 시뮬레이션을 진행하고 결과 비교 분석을 통해 인식률 향상 방안을 제시하고자 한다.

## 2. 관련연구

### 2.1 심층학습

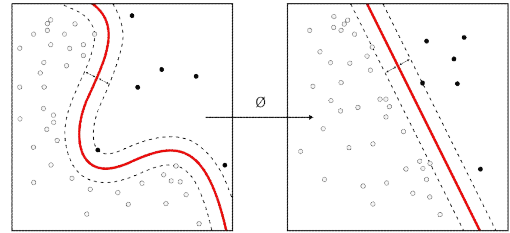


Fig. 1. Boundary scatterplot of linear support vector systems representing machine learning and data mining

심층학습 또는 딥러닝은 여러 비선형 변환기법의 조합을 통해 높은 수준의 추상화를 시도하는 기계 학습 알고리즘의 집합으로 정의되며, 큰 틀에서 사람의 사고방식을 컴퓨터에게 가르치는 기계 학습의 한 분야라고 할 수 있다. 어떠한 데이터가 있을 때 이를 컴퓨터가 알아 들을 수 있는 형태로 표현하고 이를 학습에 적용하기 위해 많은 연구가 진행되고 있으며, 이러한 노력의 결과로 다양한 기법들과 컴퓨터 비전, 음성인식, 자연어 처리, 음성/신호처리 등의 분야에 적용되어 최첨단의 결과들을 보여주고 있다[1].

### 2.2 회화 문화재 객체 검출



Fig. 2. Korean-Chinese-Japanese traditional paintings of the same object but different styles

딥러닝 기반 객체 검출 분야에서 대용량의 데이터를 통해 높은 성능을 지닌 모델을 생성할 수 있다. 전통회화에서 표현된 객체를 탐지하는 경우에는 동일한 객체 클래스여도 작품에 따라서 상이한 화풍과 형태로 표현되어 있으므로, 다양한 범위를 취급하는 데이

터 확보가 필요하다. 동양화는 서양화와 달리 객체가 표현된 전경과 배경의 경계가 확실하지 않고, 시간이 지남에 따라 일부분이 훼손되거나 형상이 불확실하게 표현된 경우가 다수 존재한다. 이러한 특성을 지닌 전통회화 내의 객체를 탐지하기 위해 자연 이미지 기반 데이터셋을 활용하여 모델 구축을 진행하거나 객체 다양성을 고려한 검출을 할 수 있다. 해당 특성을 반영한 새로운 종류의 이미지 데이터셋을 확보하여 새로운 영상 이미지 기반 생성 기법이 필요하다[4].

### 2.3 실시간 객체 탐지

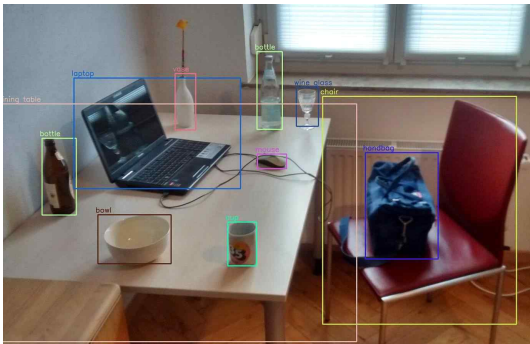


Fig. 3. YOLO model used in OpenCV

객체 탐지는 컴퓨터 비전과 이미지 처리와 관련된 컴퓨터 기술로서, 디지털 이미지와 비디오로 특정한 계열의 객체를 감지하는 일을 다룬다. 얼굴 검출, 보행자 검출 및 영상 복구, 비디오 감시를 포함한 수많은 컴퓨터 비전 분야에 응용되고 있다. 객체 탐지를 위한 방식은 일반적으로 기계 학습 기반 접근 또는 딥러닝 기반 접근으로 분류된다. 기계 학습 접근의 경우 서포트 벡터 머신 등의 기법을 사용하여 분류하는 일이 필요하다[2].

### 2.4 YOLOv5

기존의 실체를 가지고 있는 물리적인 사물을 실시간으로 인식하는 것에 초점이 맞춰져 있는 YOLO(You Only Look Once) 모듈은 기본적으로 카메라에 포착된 사물의 영역 표시와 인식된 사물이 무엇인지, 어떤 것인지 알려주는 모델이다[5].

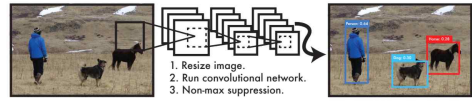


Fig. 4. YOLO's detection system

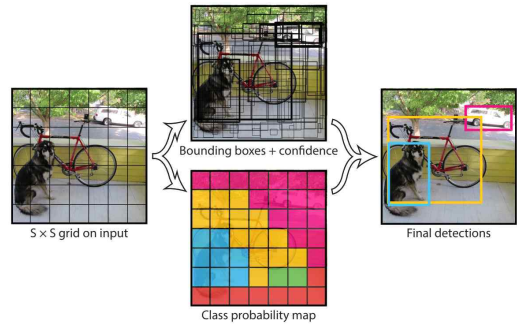


Fig. 5. YOLO's Detection Model

YOLO는 이미지 내에서 특정 객체의 경계 범위와 클래스 연관성을 단일 회귀 문제로 간주하여, 이미지를 한 번 보는 것으로 객체의 종류와 위치를 추측한다. 사람이 어떤 이미지를 보았을 때, 이미지 내부에 있는 객체들의 디테일을 한 눈에 파악할 수 있는 의식적 사고의 개입을 전산적으로 구현한 것이라고 할 수 있다. Fig. 4. 및 Fig. 5.와 같이 단일 합성곱 네트워크를 통하여 다수의 객체 경계 범위에 대한 클래스 연관성을 계산하는 방식이다[6-8].

## 3. YOLO를 이용한 회화 객체 인식

### 3.1 소규모 객체 인식

전통회화의 다양한 객체들을 인위적으로 설정하여 학습을 진행한 YOLOv5 모듈을 통해 인식과 검출을 진행한다. 앞서 언급했듯이 명확한 특징이 필요하므로 객체들은 사람과 동물들을 위주로 선정하였으며, 이와 유사한 객체들을 태깅(Tagging)하여 학습을 진행한다.

선정한 회화의 객체들을 학습을 거치지 않은 초기 상태로 인식한 결과와 유사한 객체들을 학습한 상태의 인식 결과는 Fig. 6.과 같았다.

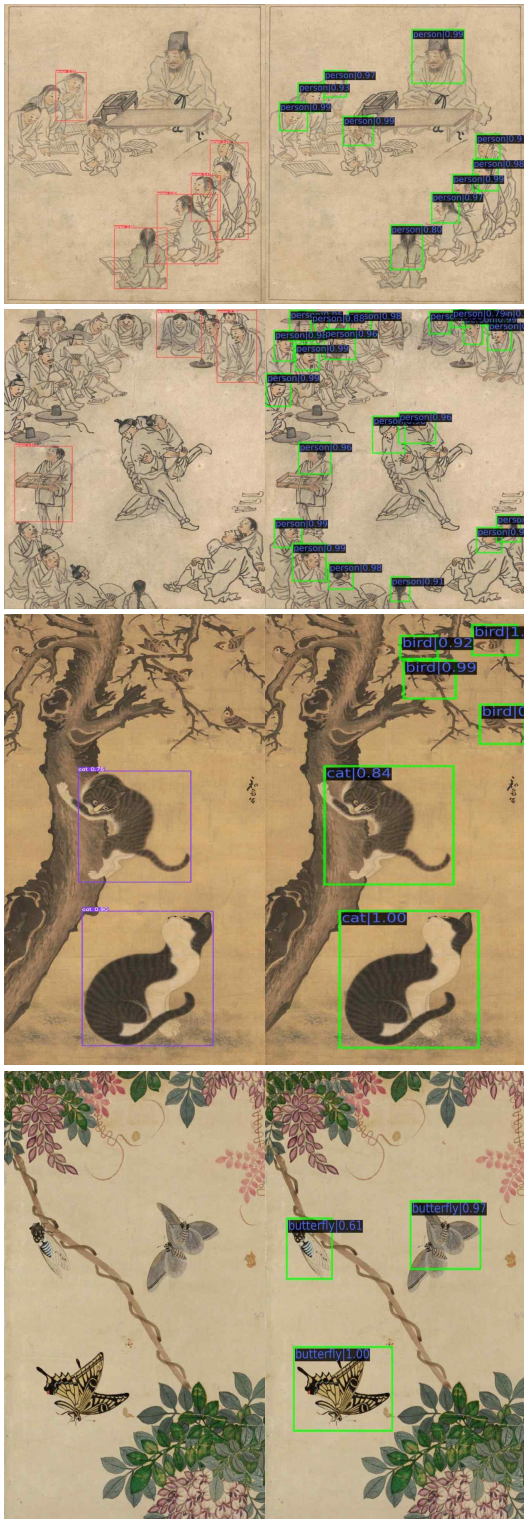


Fig. 6. YOLO object recognition results in initial state (left) and results in learned state (right)

### 3.2 대규모 객체 인식

YOLOv5가 이미지 내부 객체를 인식하게 되면 기본적으로 지니고 있는 클래스 명칭을 통하여 해당 객체의 클래스를 표시하게 된다. 여기서 Table 1.과 같은 자체적인 클래스를 설정하고 반복적인 학습을 통한 인식을 향상 정도를 관측하기 위해 다수의 객체가 나타난 사진 내부의 세부 객체들을 이미지 라벨링 (Labeling)하여 학습을 진행한다[9].

Table 1. Self-setting class

Carrier (말의 기수)	Soldier (병사)	Horse (기병)	Writer (문관)
Fighter (무관)	Woman (궁녀)	Dancer (기생)	Gama (침의 가마)
HFlag (기수 기병)	Flag (기수 병사)	SFlag (기수 하인)	Sunbi (선비)
Politic (관료)	Musician (군악병)	King (왕)	Queen (정실 가마)

자체적인 클래스를 설정하여 총 2차에 걸친 학습을 통해 인식된 객체는 Fig. 7, Fig. 8, Fig. 9과 같았으며 모듈 내부의 변화는 Fig. 10과 같았다.

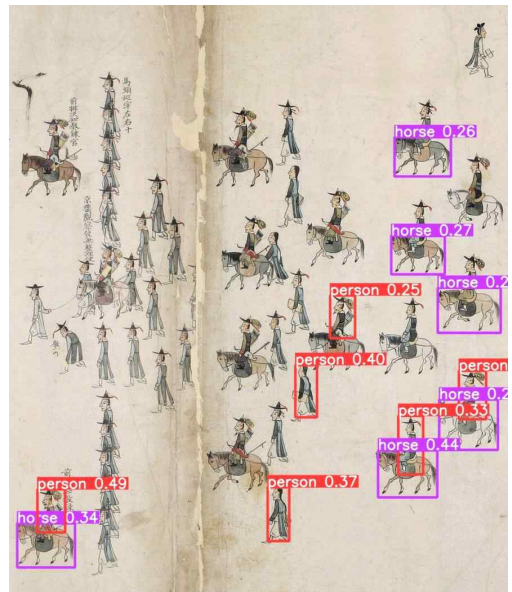


Fig. 7. YOLO Object Recognition Results in Initial State



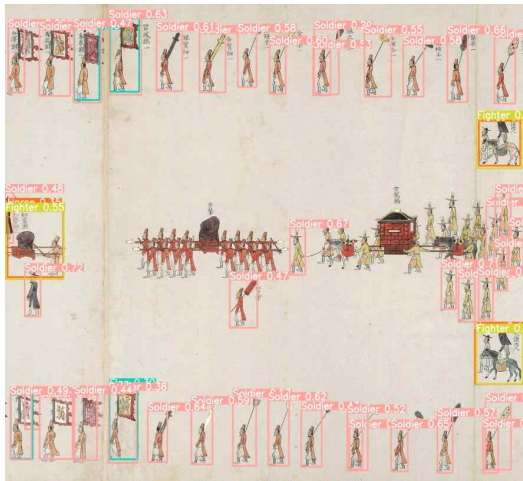


Fig. 8. Object Recognition Results in 1st Learning State

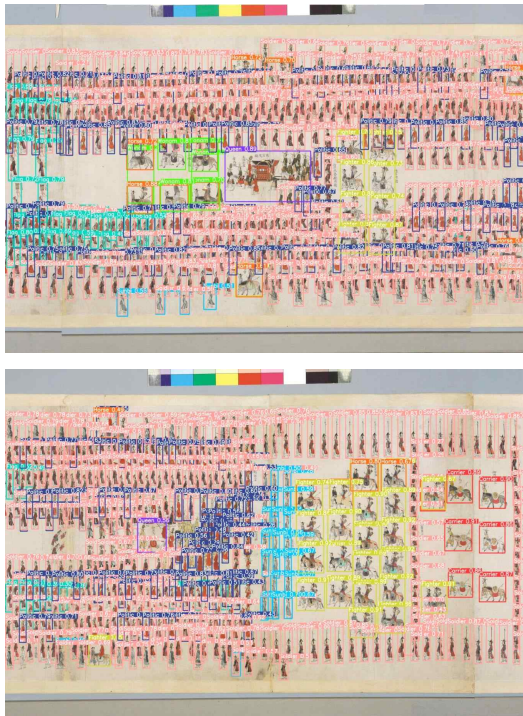


Fig. 9. Object Recognition Results in 2nd Learning State

#### 4. 결과 및 분석

실험 결과, 일부의 사람 객체, 일부의 동물 객체만 인식이 되었으며 전체적으로 인식률이 저조하거나 인식 자체가 불가능한 결과를 확인하였다. 동물 객체들이 그려진 회화는 대부분 인식하였으나, 사람

객체들이 그려진 회화는 사람의 전체 부분이 아닌, 특정 부분만이 검출된 것을 확인할 수 있었다. 이는 동물 객체의 전체적인 모습에서 특징을 포착하는 것과 달리 사람 객체는 얼굴이라는 특정 부분이 가장 효과적으로 그 객체를 설명하는 특징이 되므로 해당 결과가 도출된 것을 짐작할 수 있다.

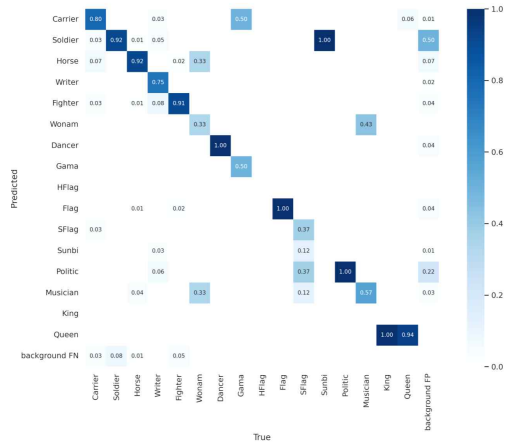


Fig. 10. Changes in internal modules that have undergone learning

자체적인 클래스를 설정한 환경에서는 일부 의도한 클래스대로 인식했으나 절반에 가까운 객체는 인식하지 못한 것을 확인하였고, 회화에 다수를 차지하는 'Soldier' 클래스와 'Horse' 클래스 위주로 객체 인식이 된 것을 추가로 확인할 수 있다. 이를 바탕으로 1차 학습을 거친 결과, 대다수를 차지하던 'Soldier', 'Horse' 클래스를 좀 더 자세하게 분류하여 'Flag', 'Fighter'와 같은 클래스를 인식할 수 있게 된 것을 확인할 수 있었다. YOLO를 통해 출력된 콘볼루션 매트릭스와 결과 그래프를 통하여 일부 클래스가 객체에 대하여 향상된 인식률을 보여준다. 이어서 2차 학습을 거친 결과, 회화 내의 모든 객체를 인식하는 것을 확인하였고, 객체마다 학습된 데이터를 바탕으로 어느 클래스를 부여하는 것이 적절한지 정확도가 향상된 것을 알아볼 수 있다. 하지만 전통회화의 특성 상 색감의 구분이 명확하지 않다는 것에 의해 'King' 클래스와 'Queen' 클래스 주변의 'Soldier', 'Politic' 클래스가 확실하게 구분되지 않았다는 점이 있었고, 화성인 행반차도 내부에서 단 하나만의 고유 객체인 'King'과

'Queen'의 인식과 구분이 되지 않았다는 점을 확인할 수 있었다. 'Queen' 클래스의 경우에는 가마와 주변의 하인들을 통해 다른 객체들보다 구분이 쉬운 특징을 가지고 있지만 'King' 클래스는 회화에 표시하지 않는 전통적 사유와 말의 탑승자 유무 및 양산의 색감에서 다른 객체들과의 차별성을 확보하지 못하였기에 이러한 결과가 도출되었다는 것을 짐작할 수 있다.

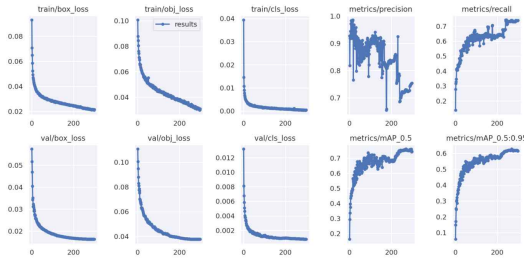


Fig. 11. Check of recognition rate based on convolution matrix

1차와 마찬가지로 콘볼루션 매트릭스와 결과 그래프를 통해 확인된 데이터로는 모든 클래스가 객체에 대하여 고르게 분포된 인식률을 확인할 수 있었고, 결과적으로 태깅과 이미지 라벨링을 통해 입력 데이터인 회화의 객체들을 표시하고 학습을 거칠수록 인식률이 향상된다는 것과 잘못된 탐지 정도를 줄일 수 있다는 것을 출력 데이터 및 실험을 통해 확인하였다.

## 5. 결론

기존 YOLO 시리즈 모듈 및 모델과 인공지능을 이용한 학습 기반 실시간 이미지 처리 시스템들은 현실의 실체를 가진 물리적 사물을 인식하고 탐지하는 것에 초점을 맞추고 있기 때문에, 현실 객체와 비교적 유사한 전통회화 내의 객체를 인식 및 탐지하는 것을 의도해보았다. 학습을 거치면서 객체의 인식률이 향상되고 탐색 로스가 감소되는 것을 통하여 인공지능의 반복 학습을 통해 전통회화 측면의 접근 및 응용 가능성을 확인할 수 있었다. 본 연구에서는 이러한 인공지능의 특성을 기반으로 실시간적인 이미지 처리 시스템을 통해 한국 전통회화라는 특정 분야에서의 객체 인식 및 검출에 접근해보았다. 그러나 현대 미술품과 달리 실제 사물을 비교적 이질적으로 표현한 과거 전통 미술품 특성으로 인해 정확한 객체의 인식 및

검출, 클래스 지정은 한계점으로 확인하였으며, 추후 이와 같은 한계점을 개선하기 위한 많은 양의 회화 데이터 학습과 다양한 클래스 지정을 통한 연구가 진행되어야 할 것이다.

## ACKNOWLEDGMENTS

본 연구는 문화체육관광부 문화기술 연구개발 지정 과제인 CHIC(Cultural Heritage Intelligent Curation) 실감형 문화유산 체험을 위한 애셋 기반 지능형 큐레이션 및 서비스 운영기술 개발 과제의 일환으로 수행되었습니다. (과제번호 : R2020040045)

## REFERENCES

- [1] Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. *IEEE transactions on pattern analysis and machine intelligence*, 35(8), 1798–1828. DOI: 10.1109/TPAMI.2013.50
- [2] Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 580–587). DOI: 10.1109/CVPR.2014.81
- [3] D. H Kwon, & J. M Yu. (2021). A Study on the Construction of Image Datasets for Object Detection of Painting Cultural Heritage. *Proceedings of the Korea Information Processing Society Conference*, 28(2), 853–855. DOI: 10.3745/PKIPS.y2021m11a.853
- [4] C. W Park., H. K Kim., & J. H Lee. (2022). A Study on Text Mining for Generating Knowledge-Based Relational Networks in Cultural Heritage Domain. *Proceedings of the Korea Contents Association Conference, 2022*, 479–480.
- [5] H. J Mun. (2020). Design for Safety System get On or Off the Kindergarten Bus using User Authentication based on Deep-learning. *Journal of Convergence for Information Technology (JCIT)*, 10(5), 111–116. DOI: 10.22156/CS4SMB.2020.10.05.111

- [6] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition. DOI: 10.1109/CVPR.2016.91
- [7] Redmon, J., & Farhadi, A. (2017). YOLO9000: better, faster, stronger. In Proceedings of the IEEE conference on computer vision and pattern recognition. DOI: 10.1109/CVPR.2017.690
- [8] YOLOv5. <https://github.com/ultralytics/yolov5>
- [9] OpenLabeling. <https://github.com/Cartucho/OpenLabeling>
- [10] Viraktamath, S. V., Katti, M., Khatawkar, A., & Kulkarni, P. (2013). Face detection and tracking using OpenCV. *The SIJ Transactions on Computer Networks & Communication Engineering (CNCE)*, 1(3), 45-50. DOI: 10.1109/ICECA.2017.8203730
- [11] Mun, H. J., & Kim, G. H. (2019). A Survey on Deep Learning based Face Recognition for User Authentication. *Journal of Industrial Convergence*, 17(3), 23-29. DOI: 10.22678/JIC.2019.17.3.023
- [12] B. R Jung. (2020). A Study on Video Caching Control Protocol for Deduplicate Transmission. *Journal of Next-generation Convergence Technology Association*, 4(6), 589-594. DOI: 10.33097/JNCTA.2020.04.06.589
- [13] J. S Song, & Y. S Ra. (1989). A Study on Input Output System for PC Operating System. *Proceedings of the Korean Information Science Society Conference*, 16(2), 545-548.
- [14] J. S Song, (2005). Frontline Activity Knowledge Discovery Modeling Based on Spatiotemporal Attributes, Chungbuk National University.
- [15] S. H Gwak, (2022). Touring Performance Measurement of Pin-point Surveillance System, Proceedings of the Korea Contents Association Conference, 2022, 93-94

**박수영(Soo-Young Park)**

**[학생회원]**



- 2022년 8월 : 한국전자통신연구원 통신미디어연구소 콘텐츠연구본부 홀로그래픽콘텐츠연구실(연구연수생)
- 2023년 2월 : 성결대학교 정보통신공학과(학생)
- 관심분야 : 하드웨어, 전자회로, 논리회로, 물리소자, 반도체, 디스플레이

· E-Mail : parksuyeong@naver.com

**박찬우(Chan-Woo Park)**

**[정회원]**

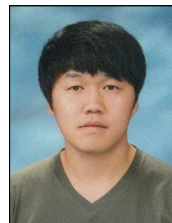


- 한국전자통신연구원 통신미디어연구소 콘텐츠연구본부 홀로그래픽콘텐츠연구실
- 관심분야 : 전자/정보통신공학, 컴퓨터학, 복합학, 학제간연구

· E-Mail : gamer@etri.re.kr

**김희권(Hee-Kwon Kim)**

**[정회원]**



- 한국전자통신연구원 통신미디어연구소 콘텐츠연구본부 홀로그래픽콘텐츠연구실
- 관심분야 : 전자/정보통신공학, 컴퓨터학, 복합학, 학제간연구

· E-Mail : hkkim79@etri.re.kr

**이재호(Jae-Ho Lee)**

**[정회원]**



- 한국전자통신연구원 통신미디어연구소 콘텐츠연구본부 홀로그래픽콘텐츠연구실
- 관심분야 : 교육학, 컴퓨터학, 학제간연구, 전자/정보통신공학, 한국어문학

· E-Mail : jhlee3@etri.re.kr