

## 딥러닝 모델을 이용한 비전이미지 내의 대상체 분류에 관한 연구

조영준<sup>1</sup>, 김종원<sup>2\*</sup>

<sup>1</sup>한국기술교육대학교 전기전자정보통신공학과 대학원, <sup>2</sup>한국기술교육대학교 기전융합공학과

### A Study on The Classification of Target-objects with The Deep-learning Model in The Vision-images

Youngjoon Cho<sup>1</sup>, Jongwon Kim<sup>2\*</sup>

<sup>1</sup>Graduate school of Electrical, Electronics and Communication Engineering, Korea University of Technology and Education.

<sup>2</sup>Department of Electromechanical Convergence Engineering, Korea University of Technology and Education.

**요약** 본 논문은 Deep-learning 기반의 검출모델을 이용하여 연속적으로 입력되는 비디오 이미지 내의 해당 대상체를 의미별로 분류해야하는 문제에 대한 구현방법에 관한 논문이다. 기존의 대상체 검출모델은 Deep-learning 기반의 검출 모델로서 유사한 대상체 분류를 위해서는 방대한 DATA의 수집과 기계학습과정을 통해서 가능했다. 대상체 검출모델의 구조개선을 통한 유사물체의 인식 및 분류를 위하여 기존의 검출모델을 이용한 분류 문제를 분석하고 처리구조를 변경하여 개선된 비전처리 모듈개발을 통해 이를 기존 인식모델에 접목함으로써 대상체에 대한 인식모델을 구현하였으며, 대상체의 분류를 위하여 검출모델의 구조변경을 통해 고유성과 유사성을 정의하고 이를 검출모델에 적용하였다. 실제 축구경기 영상을 이용하여 대상체의 특징점을 분류의 기준으로 설정하여 실시간으로 분류문제를 해결하여 인식모델의 활용성 증진을 통해 산업에서의 활용도를 확인하였다. 기존의 검출모델과 새롭게 구성된 인식모델을 활용하여 실시간 이미지를 색상과 강도의 구분이 용이한 HSV의 칼라공간으로 변환하는 비전기술을 이용하여 기존모델과 비교 검증하였고, 조도 및 노이즈 환경에서도 높은 검출률을 확보할 수 있는 실시간 환경의 인식모델 최적화를 위한 선행연구를 수행하였다.

**Abstract** The target-object classification method was implemented using a deep-learning-based detection model in real-time images. The object detection model was a deep-learning-based detection model that allowed extensive data collection and machine learning processes to classify similar target-objects. The recognition model was implemented by changing the processing structure of the detection model and combining developed the vision-processing module. To classify the target-objects, the identity and similarity were defined and applied to the detection model. The use of the recognition model in industry was also considered by verifying the effectiveness of the recognition model using the real-time images of an actual soccer game. The detection model and the newly constructed recognition model were compared and verified using real-time images. Furthermore, research was conducted to optimize the recognition model in a real-time environment.

**Keywords** : Deep Learning Recognition Model, Image Processing, Target-object Recognition, Classification, Real-time Recognition

본 논문은 2020년도 한국기술교육대학교 교수 교육연구진흥과제 지원에 의하여 연구되었음.

\*Corresponding Author : Jongwon Kim(Korea University of Technology and Education.)

email: kamuiaj@koreatech.ac.kr

Received November 2, 2020

Revised December 23, 2020

Accepted February 5, 2021

Published February 28, 2021

## 1. 서론

인간은 경험을 통해 학습된 내용을 기억하여 기억속의 다양한 정보의 조합을 통해서 특정물체를 구분할 수 있다. 특히 시각을 통해서 전달된 정보는 학습된 물체의 특징점이 존재한다면 즉시 시각정보 속에서 다양한 대상물체 중 특징요소를 기준으로 대상체를 구분할 수 있는 능력을 가지고 있다. 이러한 시각정보 메커니즘의 기계적 구현을 위하여 카메라와 같은 기계적인 요소와 이를 이미지정보로 구현하는 소프트웨어로 시각정보를 형상화하고 있다. 하지만 이렇게 인공적으로 형상화된 정보는 이미지(Image)로 정의되어 실제 이미지 내에 존재하는 대상물체는 인식을 위한 대상체와의 연관성을 찾을 수 없기 때문에 정보화 되었다고 볼 수 없으며, 시각 메커니즘과 같이 즉각적인 응답이 필요한 환경에서 활용될 수 있는 연속된 이미지의 물체 검출기법 연구[1], 검출대상물체의 특징점 축출을 위한 방법[2]과 검출기준 설정 방법[3] 등의 연구가 이루어지고 있고 이를 이미지 처리 기술(Vision technology)로 활용할 수 있는 거시적인 결과가 특정 환경에서는 활용되고 있다[4,5]. 따라서 본 논문에서는 인간과 같은 실시간 대상체인식의 분류목적을 구현하기 위하여 딥러닝을 기반으로 한 이미지 내의 대상체 인식모델을 이용한 검출(Detection)기반 실시간 대상체(Target-Object)의 분류방법(Classification)에 대해 연구하였다.

## 2. 물체검출모델의 구조개선을 통한 유사 물체의 인식 및 분류

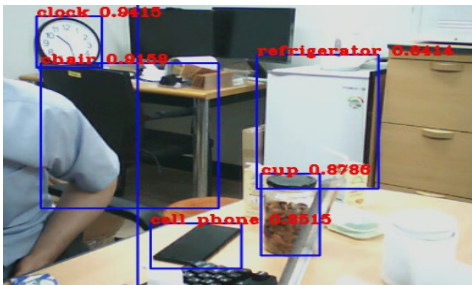


Fig. 1. Recognition of independent objects in a image

Fig. 1과 같이 카메라(범용 CMOS 카메라, 30fps, VGA)를 통해 획득된 이미지 내에는 독립된 의미를 갖는 여러 종류의 물체들이 존재하고 있다. 이런 여러 물체는

단위 이미지 측면에서는 모두 화소로 구분되며 이는 이미지 내에서 2차원의 특정 위치 정보와 색 정보로 구성된다. 하지만 정보의 입장에서 보면 물체는 이런 화소정보의 그룹으로 표현될 뿐 물체 자체가 갖는 의미를 구분하는 정보는 가지고 있지 않다.

비전기술은 화소 정보로부터 물체에 해당하는 고유한 의미를 부여하기 위하여 많은 방법과 기술들이 개발 중에 있다[6]. 특히 최근 연구되고 있는 Deep learning 기반의 YOLO 모델은 검출대상 물체들에 해당하는 다양한 이미지를 이용하여 사전학습과정을 통해 물체인식모델을 구현하고, 대상체의 추가적인 인식확대를 위한 추가학습을 통해 입력된 물체들의 인식이 가능하도록 연구가 진행되고 있다[7,8]. 하지만 대상체가 가지고 있는 고유성(Identity)이 모호하고, 고유성이 불분명한 환경조건에서 고유성 확보를 위한 데이터의 의미적 분류(Classification)를 위해 많은 시간과 노력이 필요한 구조의 인식모델이라면, 이런 인식모델을 이용하여 해당 이미지 내의 대상체에 대한 분류문제의 실시간 구현은 매우 어렵다. 따라서 본 논문에서는 기존의 인식모델의 구조변경과 비전기술의 결합을 통해 실시간으로 입력되어지는 동영상 이미지 내에 존재하는 대상체들의 특징을 살려 대상체 분류를 실시간으로 구현할 수 있는 방법을 연구하였다.

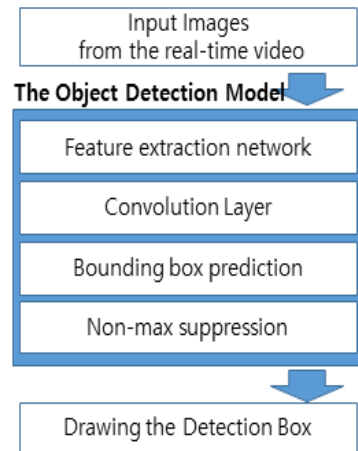


Fig. 2. Process of the recognition model

### 2.1 기존 물체검출모델의 한계

YOLO알고리즘을 이용한 대상체 인식모델은 기존 인식모델에 비해 검출속도가 빠르고 2배 정도의 mAP를 갖는 장점이 있다. Fig. 2와 같은 구조와 기능을 통해 이미지 내의 학습된 80종류의 물체를 검출할 수 있으며 ,

확률적 영역을 해당 이미지 내에 표시함으로써 사람으로 하여금 대상체가 인식되었음을 단순히 표시하는 검출 모델이다.

또한 대상체가 존재하는 주변 환경의 다양성에 따른 인식의 오류를 최소화하기 위해서 인식모델 개발 과정에서 해당 물체에 대한 방대한 데이터의 수집이 요구되는 특징을 가지고 있다. 따라서 인식을 위한 데이터의 다양성은 대상체에 대한 종류의 구분을 가능하게 하지만, 대상체가 가지고 있는 고유성은 Fig. 3과 같이 별도로 구분하기 어렵다는 문제점을 가지고 있다. 이는 해당 대상체의 의미적 세분화를 통한 재학습 방법을 이용하여 해결할 수 있지만, 기존의 학습모델을 이용하기 위해서는 인간의 대상체 구분을 위한 판단에 해당하는 기능을 포함하고 있어야 한다.

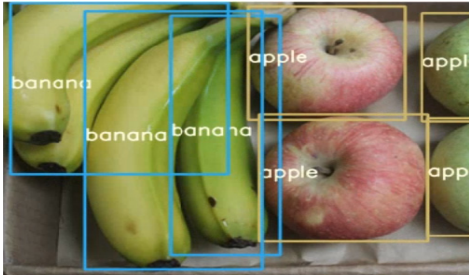


Fig. 3. Uniqueness of the objects

## 2.2 물체검출모델의 분류 문제 분석

기존의 인식모델은 Fig.3에서 보여지는바와 같이 과일류(Fruit)에 해당하는 분류로 구분되지 않고 고유한 물체로 사과(apple), 바나나(banana)로 인식한다. 이는 사과와 바나나가 과일의 분류에 속하지만 그 형태 및 색깔 등과 같이 확연히 구분될 수 있는 특징(고유성)을 갖고 있기 때문에 기계학습 메커니즘을 적용한 인식모델의 입장에서는 모양과 형태가 유사한 다른 물체들보다 구분이 용이한 특징을 보인다. 반면에 Fig. 4와 같이 사람(person)의 경우, 사람으로서 갖출 수 있는 유사성(similarity)은 머리, 다리, 팔의, 눈 등의 존재와 같은 공통점과 해당 사람이 착용하고 있는 옷, 장신구, 환경 등에 따른 고유한 특징을 고려하여 사람1과 사람2로 구분할 수는 있으나, 의미적 사람 1, 2로 구분할 수 있는 분류 문제에 있어서는 별도의 구분기준과 판단근거가 필요함을 알 수 있다.



Fig. 4. Similarity of the objects

## 2.3 검출 대상체의 분류를 위한 정의

기존의 검출모델을 그대로 사용하여 검출 대상체 인식을 위한 고유성과 유사성을 동시에 고려할 수 있도록 검출모델의 구조를 변경하여 Fig. 5와 같은 새로운 인식모델을 구현하였다.

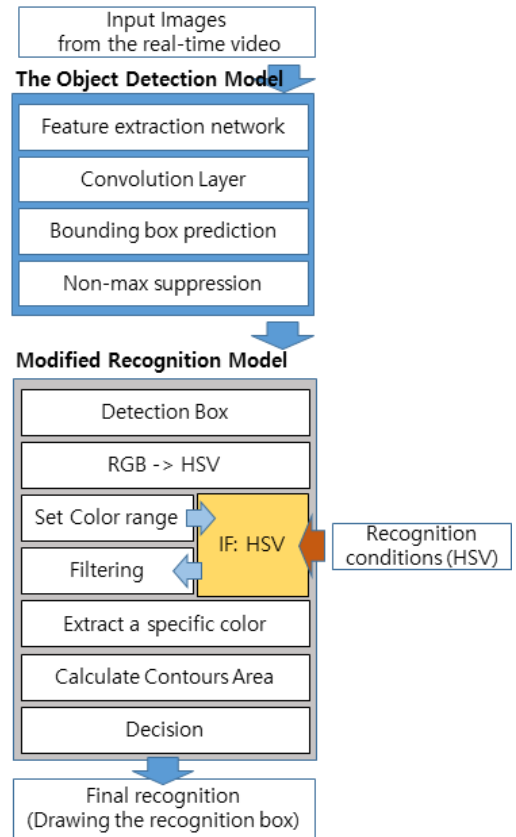


Fig. 5. Modification of the Object Recognition Model

이미지 정보가 가지고 있는 특정 화소(pixel) 정보는 RGB 칼라 영역으로 구성되어 있는데, 단순히 화소에 대해 RGB가 가지고 있는 컬러 정보를 가지고 인식을 위한 기초자료로 이용하는 것보다, 색상과 강도를 구분하는 인간의 시각 시스템과 유사한 HSV 칼라 공간 변환을 통해 구현하였다[9]. 고유성 확보를 위한 인식 조건으로는 HSV 칼라 공간(H:0~360, S:0~1, V:0~1)에 대해 각 필드별로 구분기준에 맞는 범위를 지정하였고, 지정된 영역의 최대, 최소값을 임계값으로 정의하여 고유성 정보를 정의 하였으며, 검출된 영상의 면적에 필터링을 적용하여 고유성 조건의 임계값을 중심으로 검출대상 물체에 대한 A, B, C, D 등의 그룹으로 분류가 가능하도록 구현하였다. 따라서 임계값 설정은 대상체의 분류 문제에 있어서 고유성 및 유사성 결정을 위한 인식조건(Recognition condition)으로 정량화 하여 Fig.5와 같이 인식모델에 적용 하였다.

## 2.4 변경된 대상체 인식모델의 적용

대상체의 고유성과 유사성을 모두 고려하여 분류문제를 적용할 수 있는 환경조건은 매우 많으나, 본 논문에서는 실시간으로 분류가 필요한 축구경기 영상을 이용하여 그 성능을 검증하였다. 축구경기는 각 팀에 소속된 선수와 심판이 해당 팀과 심판을 구분하기 위한 유니폼을 착용하고 하나의 경기장(이미지 내의 공존환경)에 혼재되어 경기를 실시하므로 동일 환경내의 대상체 분류문제에 대한 특이점(고유성, 유사성)이 모두 존재한다. 이때 해당 이미지 내의 사람은 Fig. 4를 통해 사람이라는 의미로 분류된 동일한 객체로 인식되고, Fig. 6과 같이 사람이 동일한 유사성만 고려된 학습결과에 기반 한 각각의 여러 객체들로 검출됨을 볼 수 있다.

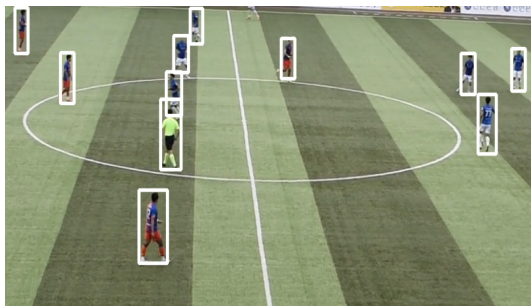


Fig. 6. The Recognition result of persons in the image

동일한 이미지 내에 개별적으로 인식된 선수들을 의미

적으로 분류해 보면, '심판(referee)들', '팀A의 선수(player)들', '팀B의 선수들'로 그 의미가 구분될 수 있다. 이때 실시간 이미지는 1개의 이미지가 시간의 흐름(frame)에 따라 지속적으로 획득되고, 다음 이미지 프레임 사이에 분류문제를 해결해야 한다. 이는 이미지 데이터 내에서 분류문제 해결을 위한 정보를 최소화 할 필요성이 있으며, 지속적으로 입력되는 각각의 이미지마다 '사람'이라는 분류로 검출된 대상체들을 '심판들', '팀A의 선수들', '팀B의 선수들'로 구분하여 분류될 수 있는 판단근거가 적용되어야 하며, 이를 위하여 '심판들', '팀A', '팀B'의 의미를 포함할 수 있는 분류기준을 정의하고, 이런 기준을 인식모델에 적용해야 한다. 이는 방대한 양의 데이터 확보를 통해 '심판들', '팀A', '팀B'의 의미를 갖는 재학습 모델을 적용하여 새롭게 검출대상자를 추가한 검출모델을 개선하는 것 보다 간단한 분류기준을 적용한 기능의 추가와 처리구조의 변경을 통해 기존 인식모델의 기능을 이용하면서 실시간으로 분류문제를 해결할 수 있는 인식모델의 기능을 향상시킬 수 있다.

본 논문에서 이용된 이미지는 연구목적으로 승인된 실제 축구경기를 촬영한 영상에서 추출한 이미지(K3\_양평\_16리그)로 앞서 설명한 조건과 환경을 적절히 갖추고 있다. 해당 이미지에서 분류의 형태는 3가지 경우(심판, 팀A, 팀B)로 정하였다. 분류기준 조건으로, 실제 선수들과 심판이 착용하고 있는 유니폼을 특징점 조건요소로 정하고, 각 유니폼에 해당하는 HSV의 임계값을 설정하였으며 고유성과 유사성을 적용하기 위한 인식조건으로 활용하였다. 또한 다양한 축구장과 팀별 유니폼의 변화에 따른 환경변화를 고려하여 임계값을 조절할 수 있도록 설정하여 인식모델을 구현하였다.

Fig. 7과 같이 유니폼이 갖는 고유성을 이용하여 '심판'(녹색/굵은 사각형모양으로 표현)을 인식한 것을 볼 수 있다. 이는 심판이 갖는 고유성의 일치도(심판이 착용하고 있는 유니폼의 HSV의 임계값 적용)에 따라 '선수'와 '심판'을 구분하였다. 동일한 방법으로 '팀A', '팀B' 고유성을 이미지 내에 적용하여 Fig. 8, 9와 같이 '팀A'(파란색/좌-사선/사각형 모양으로 표현)와 '팀B'(빨간색/우-사선/사각형 모양으로 표현)의 선수들이 분류됨을 볼 수 있었다.

이런 분류방법은 구조가 변경된 대상체 검출모델에 분류조건 3가지를 동시에 적용함으로써 Fig. 10과 같이 하나의 이미지 내에서 '심판', '팀A' 와 '팀B'의 선수들을 실시간으로 분류 할 수 있음을 보였다.



Fig. 7. Classification for the referees in the image

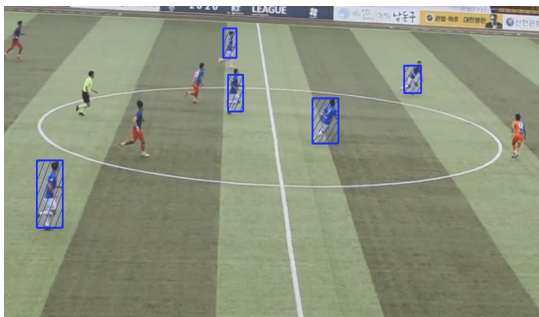


Fig. 8. Classification for the players of team 'A' in the image



Fig. 9. Classification for the players of team 'B' in the image



Fig. 10. Classification of the 3 sets in the image

### 3. 결론

본 논문은 Deep-learning 기반 물체검출 모델을 이용하여 기존의 모델이 수행하지 못했던 분류 문제를 구조변경 및 비전처리 모듈을 이용하여 실시간으로 해결할 수 있는 방법을 구현 하였다.

실세계에서는 분류 및 인식을 실시간으로 필요로 하는 다양한 요구가 존재하며, 이를 만족시키기 위하여 해당 조건에 부합되는 방대한 데이터 수집을 통해 학습모델을 이용한 해결책을 선택하기에는 많은 시간과 노력이 낭비될 소지가 있다.

본 논문에서는 특징점 또는 판단근거를 정량화하여 분류 및 인식문제 해결의 기준을 마련하여 쉽게 문제를 해결하였다. 제시될 수 있는 모든 환경에서 유사한 분류 및 인식을 위한 문제해결의 조건을 정의함에 있어서 자료 수집을 통한 학습모델의 안정적인 효과성을 검증을 위한 학습결과를 분석하고 모델을 재구성 하는 것 보다는 인간의 판단근거와 기준을 명확히 정의하여 기계학습 메커니즘에 적용하는 것이 매우 합리적임을 보였다.

또한 동영상 이미지의 실시간 처리를 위한 모듈의 구성은 기계학습에 방해되지 않는 형태와 구조를 고민하여 실시간 이미지 처리 조건에 부합하도록 연구하였기 때문에 향후 학습모델을 이용한 문제해결 방법도 변경된 인식모델에 적용할 수 있는 유연성도 확보됨을 볼 수 있었다.

본 논문의 사례와 같이 이미지 내의 대상체에 대한 분류 및 인식 위한 판단기준설정에 있어서 각 '심판', '선수'의 유니폼을 판단근거로 정의한 것처럼 일반적인 사람의 사교체계를 기반으로 어떻게 기계학습의 정의에 해당하는 정보로 판단근거를 정의하였을 때 기계학습이 가능하도록 구현 할 것인지에 대한 다양한 아이디어와 방법의 연구를 통해 인식모델의 발전이 이루어져야 한다. 이때 본 논문에서 사용한 HSV 칼라 공간의 활용은 화소의 RGB 칼라 공간으로 구성되어 있는 정보를 대상체 인식을 위한 판단근거로 직접 활용하기 보다는 색상과 강도를 구분하여 표현된 HSV 공간이 인간의 시각정보처리 메커니즘에 근접하다는 점을 이용하였고, HSV 변환을 통해 색상 단위의 범주가 아닌 채도, 명도의 영역도 포함되기 때문에 조도 및 노이즈 환경에서도 높은 검출률을 나타내며[10], 대상체 주변의 그림자, 조명, 바탕면의 무늬, 이미지 전반의 명암 변화 등에 따른 오인식 (Miss-recognition), 인식손실 (Recognition loss)의 문제에 효과적임을 간접적으로 확인하였다[11].

본 논문의 선행연구를 바탕으로 향후 기존의 인식모델

의 인식을 증가, 입력 영상의 전체적인 히스토그램 변화에 대한 연구를 통해 다양한 환경에 적응할 수 있도록 HSV 칼라 공간의 지능적으로 범위를 조절함으로써 대상체 인식의 기준설정방법에 대한 다양한 연구와 이를 기존의 인식모델과 결합하는 구조변경의 연구가 진행된다면 인식, 구분, 분류 등의 문제해결 영역으로의 기계학습 기반의 인식모델 구현이 가능할 것으로 생각한다.

## References

- [1] Liang Wang, "Moving Object Detection and Extraction in serial Image", *4<sup>th</sup> IEEE International Congress on Image and Signal Processing*, Vol. 1, pp. 928~931, 2011.  
DOI: <http://dx.doi.org/10.1109/CISP.2011.6100284>
- [2] Cyganek, Boguslaw, "Object Detection and Recognition in Digital Image: Theory and Practice", Wiley-Blackwell, 2013.  
DOI: <https://doi.org/10.1109/MIE.2014.2388012>
- [3] Seo, W. C, "Image and Voice Processing : Dominant Point Detection Algorithm on Digital Contours with Constrained Number of Points", *he KIPS Transactionsty*, Vol. 4, No. 9, 1997.  
<http://kiss.kstudy.com/thesis/thesis-view.asp?key=324529>
- [4] Lili Huang, Hefeng Wu, Guanbin Li, and Qing Wang. "Instruction-guided Object Detection", In *ACM Turing Celebration Conference (ACM TURC 2019)*, Chengdu, China, Vol. 1, pp. 6~12, May 17-19, 2019.  
DOI: <https://doi.org/10.1145/3321408.3322632>
- [5] H. Tian, T. Wang, Y. Liu, X. Qiao, Y. Li, "Computer vision technology in agricultural automation —A review", *Information Processing in Agriculture*, Vol. 7, Issue 1, pp. 1-9, March. 2020.  
DOI: <https://doi.org/10.1016/i.inpa.2019.09.006>
- [6] D. K. Shetty, U. Dinesh Acharya, N. Malarout, R. Gopakumar and P. P.J., "A Review of Application of Computer-vision for Quality Grading of Food Products", *2019 International Conference on Automation, Computational and Technology Management (ICACTM)*, London, UK, pp. 297~303, Apr. 2019.  
DOI: <https://doi.org/10.1109/ICACTM.2019.8776763>.
- [7] Y. Tian, G. Yang, Z. Wang, H. Wang, E. Li, "Apple detection during different growth stages in orchards using the improved YOLO-V3 model", *Computers and Electronics in Agriculture*, Vol. 157, Pages 417~426, Feb. 2019.  
DOI: <https://doi.org/10.1016/j.compag.2019.01.012>
- [8] A. Wong, M. J. Shafiee, F. Li, B. Chywl, "Tiny SSD: A Tiny Single-Shot Detection Deep Convolutional Neural Network for Real-Time Embedded Object Detection", *2018 15th Conference on Computer and Robot Vision (CRV)*, pp. 95-101, May. 2018.  
DOI: <https://doi.org/10.1109/CRV.2018.00023>.
- [9] Lee. S. K, Park. Y. S, Lee. G. S, Lee. J. Y, Lee. S. H, "An Automatic Object Extraction Method Using Color Features Of Object And Background In Image", *The Journal of Digital Policy & Management*, Vol. 11, No. 11, pp 459~465, 2013  
DOI: <https://doi.org/10.14400/JDPM.2013.11.12.459>
- [10] Chae. S. H, Jun. K. K, "HSV Color Model based Hand Contour Detector Robust to Noise", *Journal of Korea Multimedia Society* Vol. 18, No. 10, pp. 1149~1156, 2015.  
DOI: <https://doi.org/10.9717/kmms.2015.18.10.1149>
- [11] Na. L, Jiajun. B, Chun. C, "Real-time video object segmentation using HSV space", *Image Processing. International Conference on*, Vol. 2, pp. 85~88. Sep. 2002.  
DOI: <https://doi.org/10.1109/ICIP.2002.1039893>

### 조 영 준(Youngjoon Cho)

[정회원]



- 2000년 2월 : KOREATECH 전기전자공학과 공학석사
- 2020년 8월 ~ 현재 : KOREATECH 전기전자통신공학과 대학원

<관심분야>

지능제어, 머신비전, 로봇

### 김 종 원(Jongwon Kim)

[정회원]



- 2006년 8월 : KOREATECH 전기전자공학과 (공학박사)
- 2009년 8월 ~ 2016년 9월 : 개도국기술이전연구소 책임연구원
- 2016년 9월 ~ 현재 : KOREATECH 기전융합공학과 조교수

<관심분야>

지능제어, 융합기술, 산업용용 시스템