

스포츠 경기에서 지능인식모델을 이용하기 위한 대상체 인식오류 보상방법에 관한 연구

한준수¹, 김종원^{2*}

¹한국기술교육대학교 메카트로닉스공학과, ²한국기술교육대학교 기전융합공학과

A Study on the Compensation Methods of Object Recognition Errors for Using Intelligent Recognition Model in Sports Games

Junsu Han¹, Jongwon Kim^{2*}

¹Graduate school of Mechatronics Engineering, Korea University of Technology and Education.

²Department of Electromechanical Convergence Engineering, Korea University of Technology and Education.

요약 본 논문은 인공지능 모델의 하나인 YOLO(You Only Look Once) 인식모델 기반의 이미지 내 객체인식을 위한 활용 환경에서 딥 러닝 네트워크를 통한 고속 이동 대상체 인식의 가능성 향상과 생활 속에서 쉽게 활용될 수 있도록 2차적인 정보의 가공을 통한 의미적 데이터 수집 방법을 연구하는데 그 목적이 있다. 인식모델에서 이동 대상체 인식오류는 카메라의 프레임 속도와 대상체의 이동속도 차이에서 발생하는 미인식과 대상체와 인접한 환경에서의 유사물체가 존재로 인한 오인식으로 확인되었으며 이를 보상하는 데이터 수집 방법을 제안했다. 실제 유사환경을 대표할 수 있는 스포츠(테니스 경기)를 대상으로 획득된 이미지에서 오류의 원인요소를 비전처리 기술을 적용하여 해당오류를 최소화하기 위한 방법과 처리구조를 연구하여 유효한 2차적인 데이터 수집의 효과성을 향상시켰다. 따라서 본 연구에서 제안된 데이터 수집 방법을 적용함으로써 일반인도 스마트폰 카메라의 간단한 촬영을 통해 스스로 건강 및 경기력 향상을 위한 스포츠 및 건강관련 산업에 적용될 수 있는 데이터의 수집 및 관리가 가능함을 보였다.

Abstract This paper improves the possibility of recognizing fast-moving objects through the YOLO (You Only Look Once) deep learning recognition model in an application environment for object recognition in images. The purpose was to study the method of collecting semantic data through processing. In the recognition model, the moving object recognition error was identified as unrecognized because of the difference between the frame rate of the camera and the moving speed of the object and a misrecognition due to the existence of a similar object in an environment adjacent to the object. To minimize the recognition errors by compensating for errors, such as unrecognized and misrecognized objects through the proposed data collection method, and applying vision processing technology for the causes of errors that may occur in images acquired for sports (tennis games) that can represent real similar environments. The effectiveness of effective secondary data collection was improved by research on methods and processing structures. Therefore, by applying the data collection method proposed in this study, ordinary people can collect and manage data to improve their health and athletic performance in the sports and health industry through the simple shooting of a smart-phone camera.

Keywords : Object Recognition, YOLO, Deep Learning, Recognition Errors, Vision Technology, Data Acquisition

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

email: kamuiaj@koreatech.ac.kr

Received January 27, 2021

Accepted May 7, 2021

Revised February 22, 2021

Published May 31, 2021

1. 서론

본 논문은 인공지능 모델의 하나인 YOLO(You Only Look Once)알고리즘 기반의 이미지 내 객체인식을 위한 활용환경에서 딥 러닝(Deep learning) 네트워크를 통한 고속 이동 대상체 인식의 가능성 향상과 생활 속에서 쉽게 활용될 수 있는 방법으로 2 차적인 정보의 가공을 통한 의미적 데이터 수집 방법 연구에 관한 논문이다.

최근 지능알고리즘을 기반으로 한 전기전자 신호의 인식, 검출, 분석, 보정 및 생성과 관련된 기술과 활용 가능한 모델들이 다양한 방식을 통해 광범위하게 연구개발되고 있다. 이런 모델들은 컴퓨터를 이용한 대량의 연산 과정을 통해 이루어지기 때문에 고성능 하드웨어 시스템이 주로 사용되며, 과도한 연산 및 데이터의 효과적인 활용을 위하여 개발된 기술을 적용한 모바일 디바이스의 활용도 및 기능성이 점차 확장되고 있다. 따라서 이러한 시스템의 효율성 차원에서의 그 효과성을 검증하고자 하는 여러 기법들도 함께 연구되고 있다.

YOLO모델은 이미지의 수천 개의 하위 영역에 대하여 예측을 수행하는 다른 인공지능 네트워크모델과는 달리, 학습된 인식 대상체를 전체 이미지에 적용하고 객체 인식 예측은 글로벌 컨텍스트를 기반으로 하는 특징이 있다. YOLO모델의 아키텍처는 일련의 잔여 블록과 컨볼루션 계층으로 구성되며, 잔여 블록은 기본적으로 단축 경로가 있는 여러 컨볼루션 레이어로 구성되어 딥 러닝을 효과적으로 구현한 특징이 있다[1].

기존의 YOLO기반 인식모델은 고정된 인식대상 물체(대상체)가 획득 가능한 이미지영역 내에 존재하였을 때 인식 대상체의 인식영역에 대한 좌표 등을 데이터로 획득할 수 있고, 이를 동일 이미지에 겹침(overlapping)을 통해 표시함으로써 인식결과를 대상 이미지 내에 표시할 수 있는 구조로 구성되어 있다. 하지만 대상체가 이미지 내에 존재하지 않거나, 시간차원에서 특정 속도를 가지고 이동 또는 변형이 일어나는 특징을 갖는다면, 특정 시간 간격을 갖는 프레임(frame)을 통해 입력된 이미지에 대하여 해당 이미지 내에 존재하는 대상체의 이동 및 변형 등에 따른 인식의 오류를 나타낼 수 있다. 이는 실제로 변형되지 않는 대상체가 이동할 때, 이미지를 획득하는 장치(카메라)의 프레임 속도와 대상체의 이동속도의 차이에서 발생하는 문제이다. 이런 문제는 이미지 처리 영역에서 장치들의 특성에 따른 잔상(afterimage)에 기인한 것으로 정의되어있다. 또한 이미지 내의 대상체 인식 관점에서 본다면 특정 모양과 색을 갖는 대상체는 그 대

상체가 위치한 주변 환경이 갖는 이미지의 특징(대상체와 동일한 색상과 모양을 갖는 주변 환경의 특징)과 카메라의 촬영 각도에 따라서 인식이 되지 않거나 오인식 되는 경우가 발생한다.

이러한 일련의 과정을 살펴보면 대상체의 이동은 실질적으로 대상체가 변형되지는 않지만 시간에 따라 획득된 이미지에 대해서 인식모델의 관점에서 보면 변형과 이동이 동시에 일어나고 있으므로 미인식(unrecognition), 오인식(misrecognition) 오류가 Fig. 1 과 같이 발생할 수밖에 없다.



(a) Correct Recognition (b) Unrecognition (c) Misrecognition

Fig. 1. Two types of Recognition Errors

또한 연속된 이미지의 입력과정을 통해 시간에 따른 대상체의 이동 또는 변형과 관련된 추가적인 정보(대상체의 이동속도, 이동위치 등)의 획득에 있어서도 오류가 포함되기 때문에 생성된 추가정보에 대한 신뢰도가 낮아진다. 본 논문에서는 기존의 인식모델을 이용한 실제 유사환경을 대표할 수 있는 스포츠(테니스 경기)의 결과를 기반으로 본 연구를 통해 구현된 방법의 결과를 비교하여 인식오류를 최소화하기 위한 방법과 처리구조의 유용성을 검증하고 이를 이용한 2 차적인 데이터 수집의 효과성을 향상시켰다.

2. 본론

2.1 이동 대상체의 인식에 따른 오류

고속 이동체의 이동경로 파악은 스포츠 영역에서는 매우 중요한 요소로 활용되고 있다. Fig. 2 와 같이 사용되는 공(ball)의 인식에 필요한 고속의 하드웨어 환경이 갖추어 졌을 때 이동경로에 대한 보다 정확한 인지가 가능하다. 고속카메라(Hawk-eye)를 이용하여 160 FPS(frame/second)의 속도로 이미지를 획득했을 때와 30 FPS로 획득된 이미지에서 추출한 공의 경로 모형을 확인해 보면 큰 차이를 볼 수 있다. 공의 이동경로는 선수들의 행동에 따른 바운드, 히트, 스매시 등의 특성에 관

한 판정 및 선수들의 체력, 활동패턴 등의 2 차적인 경기력 분석 및 스포츠 과학 분야의 기초자료로 활용된다. 하지만 Fig. 2 와 같은 영상의 획득은 결국 사람의 육안인식을 통해 공의 이동경로를 예측하고 판단한다. 따라서 인식모델을 적용한 공의 이동경로 탐색방법의 구현은 스포츠산업에서 활용될 기초자료 획득을 위해 매우 유용하다.

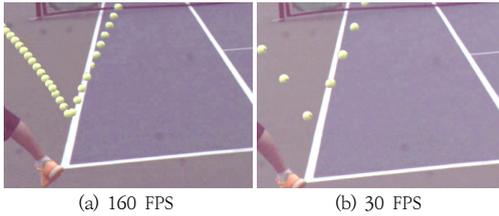


Fig. 2. Trajectory Difference according to the number of Frames Per Second

최신의 인식모델인 YOLO모델도 획득된 이미지를 기반으로 인식하기 때문에 대상체가 공과 같이 작은 물체인 경우에 카메라로부터 획득된 이미지 내에서는 유사물체가 인접한 환경(라인에 겹쳐져 보이는 공, 강한 타격에 의한 잔상, 촬영 각도에 따른 선수행동의 간섭 등)에 따른 다양한 요소에 기인한 인식오류가 발생하게 된다.

테니스 경기를 기반으로 발생할 수 있는 여러 인식오류 중 중요한 인식오류는 크게 3 가지 형태로 분석되는데 Fig. 3 과 같이 네트나 라인에 공이 겹쳐짐으로 발생하는 인식오류(overlap error), Fig. 4 와 같이 공의 고속이동에 따른 잔상효과로 인한 인식오류(afterimage error), Fig. 5 와 같이 촬영 각도에 따른 공의 실제위치 오류(position error) 등이 경기 중 발생하는 공의 이동경로 인식에 영향을 주는 중요 요인으로 작용한다.



Fig. 3. Recognition Error due to Line Overlap



Fig. 4. Recognition Error due to Afterimage



Fig. 5. Position Error according to the Shooting Angle

2.2 인식오류 요소분석을 통한 보상

2.2.1 경기분석 데이터 수집을 위한 방법

인식오류 보상 및 경기분석을 위한 데이터 수집과 관련된 인식모델의 동작은 Fig. 6 과 같다.

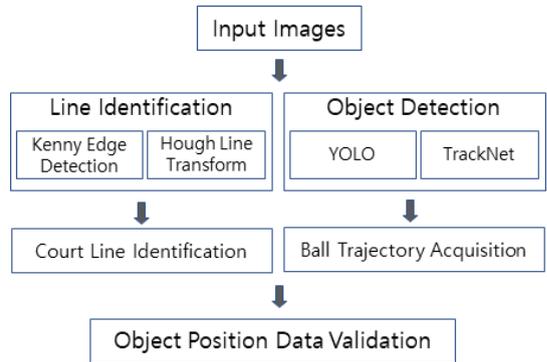


Fig. 6. Recognition Model System for Data Collection

2.2.2 코트라인 식별

코트라인(court line)은 공의 이동경로에 따른 인식에 있어서 선수의 움직임과 경기의 흐름에 지대한 영향을 주는 요소이다. 이는 식별이 가능할 정도로 선명하며, 잔디, 모래, 하드 등 다양한 코트표면 상태에 따라 식별 정도는 차이가 있을 수 있고, 공의 경로추적을 기준으로 코트좌표 계산에 있어서 카메라의 급격한 움직임이 최소화되어야 하는 조건이 수반되어야 한다. 또한 해당 코트가 획득하고자 하는 영상범위 내에 존재하지 않거나, 조도가 매우 높거나 낮은 경우 또는 선수가 획득영상 범위 내의 코트 대부분을 가리는 등의 영상획득 조건은 제외되어야 한다. 일반적인 밝기의 기상 상태와 라인 구분이 명확한 코트표면과 카메라 움직임이 고정된 1280x720, 60 FPS 화질의 경기 영상을 기준으로 촬영된 실제 경기영상(청주 국제 테니스장, 일반인 경기 복식)을 이용하여 중요 요소를 분석하였다.

테니스 코트라인을 감지하는 알고리즘은 흰색 픽셀 추출, Hough 라인 변환, 이미지에서 찾은 라인을 전체 코

트 라인에 연결 순으로 진행하였다[2]. 코트에서 어두운 픽셀과 밝은 픽셀의 평균 밝기로 두 픽셀을 구분하는 기준값을 설정하고 흑백 이미지를 Gaussian blur 처리로 노이즈를 줄인 후 Canny edge[3] 검출을 통해 Fig. 7 (a)와 같이 코트의 경계를 명확히 하였다. 다음으로 Fig. 7 (b)의 Hough 라인 변환[4] 에서 계산속도와 인식타이밍을 위해 5 프레임마다 흰색 라인의 교차점을 찾은 뒤 테니스 코트에서 모든 모서리 조합 쌍과 비교하여 가장 많은 흰색 픽셀과 일치하는 것을 선택한다. 이렇게 검색된 모서리 좌표로 코트라인을 연결하여 기준선을 확보한다.

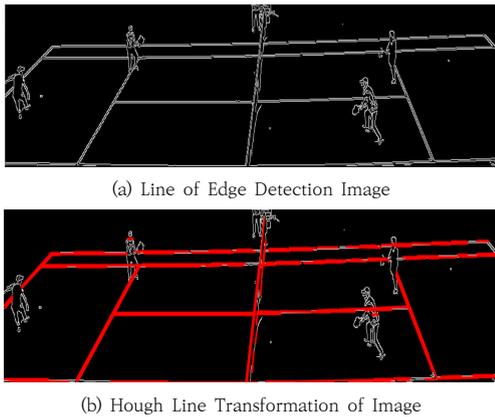


Fig. 7. Court Line Detection Results

2.2.3 공 위치좌표 추적

YOLO 인식모델과 고속 대상체 인식에 대한 성능 보완으로 딥 러닝 네트워크 모델을 이용하여 공을 탐지하고 공이 베이스라인과 사이드라인 안쪽의 인 플레이 여백 내에 착지하였는지 여부를 추론하여 코트에서 공의 착지 좌표를 추출한다. 우선, 공의 녹색 RGB 값을 선택하여 주변 노이즈에서 걸러내어 공을 식별하고, 주변 환경과 거의 같은 색상을 갖는 프레임에서 오인식을 개선하였다. 특히 고속의 작은 물체를 추적하기 위해 Fig. 8과 같은 TrackNet 모델을 적용하여 좌표를 탐지할 수 있다.

촬영한 영상 이미지 프레임을 네트워크 입력 매개 변수로 처리하여 전반부의 컨볼루션 신경망 모델(VGG16)에서 분류된 객체의 궤적 패턴을 추적하고 학습하여 입력 프레임에 대한 높은 테스트 정확도로 동일한 크기 해상도의 이동 대상체 인식 히트 맵(Heat map)을 출력할 수 있다[5].

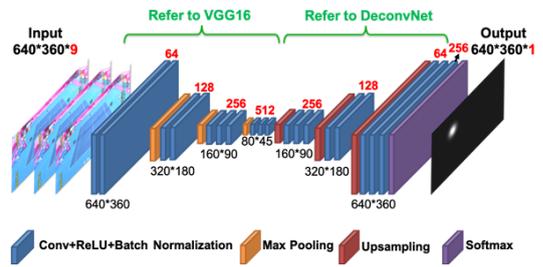


Fig. 8. TrackNet for Tracking High-speed and Tiny Objects [5]

다수의 촬영 경기 영상을 TrackNet을 통해 훈련하여 특히 작은 객체(공)에 대한 동적 객체 추적 기능을 더욱 향상시킬 수 있었으며 공 궤적 패턴을 학습을 통해 추론한 정확한 좌표 정보가 되었다.

좌표가 탐지된 공의 바운스(bounce) 상태를 식별하기 위해 3 개의 히트 맵 프레임마다 공의 기하학적 거리가 위치 변화로서 인식할 수 있는 결정값 범위 내에 존재할 때 동일한 공으로 간주하고, 그 중 가장 작은 거리 좌표를 찾는다. 후속 프레임에서 공의 인식범위 내에 좌표가 없거나 영상의 끝(final frame)에 도달 할 때까지 이를 반복한다. 최종적으로 전후 이미지 내에 존재하는 공의 두 지점 사이의 평균거리가 공의 정지로 판단할 결정값(Threshold)으로 정의되어 공의 바운스 상태로 판정한다.

2.2.4 코트 이미지의 역 원근 매핑

Fig. 9는 카메라 좌표계에서 한 평면의 픽셀 좌표(x, y)가 실제 좌표계(world coordinate)로 변환된 좌표(X, Y) 관계를 나타낸다. 이때, Eq. (1)과 같은 호모그래피(Homography) 3x3 행렬(H)을 통해 변환할 수 있다[6].

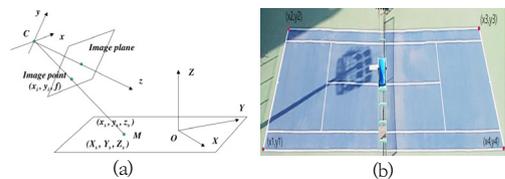


Fig. 9. Transformation From Camera Coordinates(a) to World Coordinates(b)

$$\begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} \simeq H \begin{bmatrix} X_i \\ Y_i \\ 1 \end{bmatrix} \quad H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \quad (1)$$

Where, (x, y) pixel coordinate, (X, Y) world coordinate, H denotes Homography Matrix

코트에서 4 개의 고정된 포인트의 픽셀 좌표를 통해 H의 추정치를 계산하기 위해 OpenCV 의 findHomography 라이브러리를 사용하여 구하고 이를 역변환(H-1)을 통해 가상의 코트 좌표로 Eq. (2)와 같이 매핑한다[7].

$$\begin{bmatrix} X_i \\ Y_i \\ 1 \end{bmatrix} = H^{-1} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} \quad \begin{aligned} x_i &= \frac{h_{11}X_i + h_{12}Y_i + h_{13}}{h_{31}X_i + h_{32}Y_i + h_{33}} \\ y_i &= \frac{h_{21}X_i + h_{22}Y_i + h_{23}}{h_{31}X_i + h_{32}Y_i + h_{33}} \end{aligned} \quad (2)$$

$$\sum_i \left(\left(x_i - \frac{h_{11}X_i + h_{12}Y_i + h_{13}}{h_{31}X_i + h_{32}Y_i + h_{33}} \right)^2 + \left(y_i - \frac{h_{21}X_i + h_{22}Y_i + h_{23}}{h_{31}X_i + h_{32}Y_i + h_{33}} \right)^2 \right) \quad (3)$$

최적화 파라미터(h) 결정 방법으로 Eq. (3)을 통해 최소제곱평균(Least Mean Square) 에러를 최소화하도록 평가한다. 카메라에서 본 3 차원 영상 정보에 적용된 호모그래피 변환 파라미터를 계산하여 2 차원 평면도에서의 공의 경로좌표를 누적하여 추적할 수 있다.

2.3 영상처리를 통한 공의 상태정보 확인

2.3.1 경기 중 공의 인과 아웃 판정

호모그래피 변환을 통해 영상 프레임의 모든 공의 바운스 지점을 코트의 실제 좌표로 변환한 후 바운스 된 공이 코트의 득점 영역의 임계값 이내의 좌표에 해당하는지 비교하여 공의 인과 아웃 여부를 Fig. 10 과 같이 결정한다. 이 경우 호모그래피 변환에 절대적인 영향을 미칠 수 있는 코트라인 상태는 명확히 구분 가능한 상태를 전제로 한다. 공이 라인, 네트, 라켓, 선수 등에 가려 바운스 감지가 어려운 상황에서 Eq. (4)와 같은 보간법을 활용한다.

n-th coordinate: {x1, y1}, null,, null,

k-th coordinate: {x2, y2}

(n+i)-th interpolation coordinate: {xi, yi}

$$x_i = x_1 + \frac{x_2 - x_1}{k - n} \times i, \quad y_i = y_1 + \frac{y_2 - y_1}{k - n} \times i \quad (4)$$

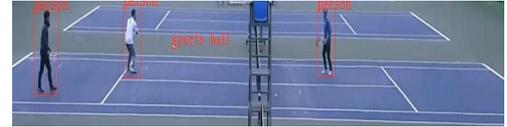
where $1 \leq i \leq (k - n - 1)$

공을 탐지하지 못한 N번째 프레임은 n-1 번째 프레임과 n+1 번째 이후 프레임내 공의 좌표를 활용하여 두 지

점과의 직선거리에 따라 선형적으로 누락된 좌표를 연결하여 결정한다.



(a) Ball Inside the court line



(b) Ball Outside the court line

Fig. 10. Video frame for determining Ball In and Out

2.3.2 좌표 변환을 통한 데이터 확보 및 검증

촬영한 영상의 좌표변환을 위해서 카메라를 통해 촬영한 공의 위치와 해당 프레임의 시간정보, 선수의 위치정보를 좌표 데이터로 변환하여 출력하도록 프로그래밍 하고 이를 비전처리 기법이 결합된 시뮬레이션을 진행하였다. 1 차적으로는 영상 안의 3 차원 정보를 2 차원의 평면상으로 보정하여 바운드 전후의 영상 정보를 바탕으로 새로운 2 차원 평면상으로 변환하여 촬영된 영상으로부터 공의 이동 정보를 볼 수 있도록 구성하였다. 변환을 위한 기준 정보는 영상의 사이드라인과 베이스라인이 만나는 4 개의 점을 통해 인식좌표를 만들고 바운스가 되는 시점의 베이스라인 측의 공의 정보를 좌표로 표시하는 방법으로 진행하였다. 본 연구에서는 Fig. 6 을 통해 나타난 방법으로 대상 테니스공의 실시간 바운스 위치정보를 인식모델을 통해 2 차 데이터로 저장하고 획득된 이미지의 FPS 정보를 이용하여 데이터 가공이 가능한 기초자료를 확보하였다.

3. 결론

본 연구를 통해 적용된 테니스 경기에서 공을 인식하기 위한 인식모델의 오류를 개선하고 실제 육안으로 판단했을 때의 일치도와 불필요하게 획득되거나 오인식 되어 수집된 데이터를 비교 하였다.

Table 1. Recognition Results of Ball Bounce Detection

	Original Image	Model Data	Un-Recognition	Mis-Recognition
Total frame	68	94	38	12
Total bounce	10	14	6	2
Bounce errors (overlap)	36	48	18	6
Bounce errors (after image)	12	20	11	3
Bounce errors (in / out)	10	12	3	1

인식모델을 이용한 오류의 보정내용을 결과표로 비교 함으로서 경기의 내용을 데이터화하기 위해 불필요하게 발생하는 오류를 보상하고 선수들의 움직임, 공의 중요 인식 좌표데이터를 확보하여 경기력 분석을 위한 기초 데이터 수집의 정확도를 높일 수 있었고, 선수들의 움직임, 공의 중요 좌표 내용을 해당 이미지의 프레임에 맞추어 수집하였다. 스마트폰이 보편화된 현대 사회에서 일반 인도 생활체육 차원에서의 스마트폰 카메라를 통한 간단한 촬영을 통해 별도의 장치를 갖추지 않아도 스스로 건강 및 경기력 향상을 위한 데이터의 수집 및 관리가능성을 보였다고 판단된다.

References

- [1] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, 2016, pp. 779-788. DOI: <https://www.doi.org/10.1109/CVPR.2016.91>
- [2] S. V. Mora and W. J. Knottenbelt, "Deep Learning for Domain-Specific Action Recognition in Tennis," *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Honolulu, HI, 2017, pp. 170-178. DOI: <https://www.doi.org/10.1109/CVPRW.2017.27>
- [3] J. Canny, "A Computational Approach to Edge Detection," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 6, pp. 679-698, Nov. 1986. DOI: <https://www.doi.org/10.1109/TPAMI.1986.4767851>
- [4] OpenCV 2.4.13.7 documentation. https://docs.opencv.org/2.4/doc/tutorials/imgproc/imgtrans/hough_lines/hough_lines.html (accessed Nov. 16, 2020)
- [5] Y. Huang, I. Liao, C. Chen, T. Ik and W. Peng,

"TrackNet: A Deep Learning Network for Tracking High-speed and Tiny Objects in Sports Applications*," *2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, Taipei, Taiwan, 2019, pp. 1-8.

DOI: <https://www.doi.org/10.1109/AVSS.2019.8909871>

- [6] Criminisi A. (2002) *Single-View Metrology: Algorithms and Applications (Invited Paper)*. In: Van Gool L. (eds) *Pattern Recognition. DAGM 2002. Lecture Notes in Computer Science*, vol 2449. Springer, Berlin, Heidelberg. DOI: https://www.doi.org/10.1007/3-540-45783-6_28
- [7] OpenCV 2.4.13.7 documentation. https://docs.opencv.org/2.4/modules/calib3d/doc/camera_calibration_and_3d_reconstruction.html#highlight=findhomography#findhomography (accessed Nov. 16, 2020)

한 준 수(Junsu Han)

[정회원]



- 2009년 2월 : KOREATECH 메카트로닉스공학과 공학석사
- 2019년 2월 : KOREATECH 메카트로닉스공학과 박사과정

<관심분야>

지능형시스템, 머신러닝, 전자제어

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, 산업응용 시스템