

## A model to secure storage space for CCTV video files using YOLO v3

Seong-Ik Kim\*

\*Professor, Dept. of Smart IT, U1 University, Asan, Korea

### [Abstract]

In this paper, we propose a CCTV storage space securing model using YOLO v3. CCTV is installed and operated in various parts of society for disasters, disasters and safety such as crime prevention, fire prevention, and monitoring, and the number of CCTV is increasing and the quality of the video quality is improving. Due to this, as the number and size of image files increase, it is difficult to cope with the existing storage space. In order to solve this problem, we propose a model that detects specific objects in CCTV images using YOLO v3 library and deletes unnecessary frames by saving only the corresponding frames, thereby securing storage space by reducing the size of the image file, and thereby Periodic images can be stored and managed. After applying the proposed model, it was confirmed that the average image file size was reduced by 94.9%, and it was confirmed that the storage period was increased by about 20 times compared to before the application of the proposed model.

▶ **Key words:** CCTV, YOLO, Computer Vision, Object Detection, Object Recognition

### [요 약]

본 논문에서는 YOLO v3를 이용한 CCTV 저장공간 확보 모델을 제안한다. CCTV는 방범, 화재 예방, 감시 등 재난·재해 및 안전을 위해 사회 곳곳에 설치·운영되며, 개수의 증가와 화질의 개선이 함께 이뤄지고 있다. 이로 인해 영상파일의 개수와 크기가 증가하면서 기존의 저장공간으로는 이를 감당하기 어려움을 느끼고 있다. 이를 해소하기 위해 CCTV 영상 속의 특정 객체를 YOLO v3를 이용하여 탐지하여 해당 프레임만을 저장하여 불필요한 프레임을 삭제하는 모델을 제안하여 영상파일의 크기를 줄임으로써 저장공간을 확보하고, 그로 인해 더 오랜 기간 영상을 저장·관리할 수 있도록 하였다. 제안 모델 적용 후 평균 94.9% 영상파일의 크기가 절감됨을 확인하였으며, 제안 모델을 적용하기 전보다 약 20배의 보관 기간이 증가했음을 확인할 수 있었다.

▶ **주제어:** CCTV, YOLO, 컴퓨터 비전, 객체 탐지, 객체 인식

- 
- First Author: Seong-Ik Kim, Corresponding Author: Seong-Ik Kim
  - \*Seong-Ik Kim (jaraka@u1.ac.kr), Dept. of Smart IT, U1 University
  - Received: 2022. 11. 21, Revised: 2023. 01. 04, Accepted: 2023. 01. 05.
  - This paper is an extension of the paper "Proposal of CCTV Storage Space Securing Model using YOLO v3 Library" presented at the 66th Summer Conference of the Korea Society of Computer and Information in 2022.

## I. Introduction

4차 산업혁명의 시대에 빠르게 발전한 과학기술로 우주의 인공위성부터 극소형의 IoT까지 다양한 정보의 수집과 분석을 할 수 있어졌음에도 불구하고 지진, 태풍, 해일 등 다양한 자연 재해를 겪고 있다. 그뿐만 아니라 화재 및 폭발 등의 인적 재난, 그리고 금융, 통신, 교통 등의 사회적 재난까지 다양한 재난·재해 및 안전 분야의 위협에 노출되어 있다. 전 세계적으로 이러한 재난·재해로부터 국민의 재산과 생명을 지키기 위해 노력하고 있다. 이러한 문제들을 해결하기 위해 인공지능(Artificial Intelligence), 빅데이터(Big-Data) 등 다양한 첨단 기술이 활용되고 있다. 인공지능 분야 중 하나인 영상분석을 이용한 컴퓨터 비전이 주목받고 다양한 분야에서 활용되고 있다. CCTV에 컴퓨터 비전 기술을 결합한 지능형 CCTV는 실시간 이벤트를 판단하는 목적으로 도입되어, 연평균('12~'17) 10% 이상의 시장 규모가 성장하면서 빠른 발전 속도를 보인다[1].

사회안전망 확보를 위한 일원으로 CCTV를 활용한 관제 시스템이 다양한 방법으로 운용 및 확대되고 있는 가운데 범죄 예방에 대한 CCTV 활용은 효과적이라는 사회적 인식이 점점 증가하고 있다. 이와 함께 앞서 말한 지능형 CCTV 및 고해상도 고성능의 CCTV 카메라의 설치, 이동식 CCTV 카메라의 설치 등 다양하고 고도화된 방법으로 CCTV가 운영되고 있다[2-7]. CCTV는 정부 기관, 경비보안업체, 사설업체, 개인 등 다양한 방법으로 곳곳에 설치되고 있으며, 방법, 화재 예방, 감시 등 재난·재해 및 안전을 위해 설치한다. 사람과 반려동물, 자동차 등이 지나가는 곳과 범죄가 발생할 수 있는 우범지역, 건물 내부와 외부, 엘리베이터 등에 CCTV를 설치하여 그곳의 상황을 즉시 확인하거나, 이후에도 확인이 필요할 경우 녹화된 영상을 볼 수 있다. CCTV가 곳곳에 설치됨에 따라 CCTV 영상파일의 개수가 증가함으로 인해 저장공간에 제약이 발생된다. 그에 따라 저장공간을 확보하기 위해 오래된 영상들이 삭제하고 새로운 영상을 저장하는 프로세스가 반복적으로 이뤄지고 있다. 이러한 이유로 저장공간의 크기에 따라 시간이 지난 영상을 확인해야 하면 이미 삭제되어 볼 수 없는 상황도 발생하고 있다.

본 논문에서는 YOLO v3를 이용하여 CCTV 영상 속 사람을 탐지해 해당 프레임만 획득하여 저장함으로써, CCTV 영상파일 크기를 줄여 저장공간을 확보하는 모델을 제안한다. 본문에서는 제안 시스템 모델의 구성도를 통해 CCTV와 서버, PC의 구성을 표현하였고, 흐름도에서는 영상파일이 작업스케줄러에 의해 실행된 파이썬 실행 파일

이 YOLO v3 라이브러리를 이용하여 탐지된 객체를 찾고 저장하는 처리 과정을 나타내고 있다.

## II. Related works

### 1. YOLO(You Only Look Once)

2015년 Joseph Redmon이 YOLO v1 논문 발표를 시작으로 2018년 YOLO v3까지 개발하였다. 딥러닝을 기반으로 한 Object Detection 모델은 1-stage detector와 2-stage detector로 구분할 수 있으며, YOLO는 그중 1-stage detector 모델이다[8-10].

다음의 Fig. 1.은 YOLO v1의 모델로 입력 이미지를 7x7 그리드 셀(Grid Cell)로 나누고 각 셀(Cell)이 하나의 객체(Object)에 대한 탐지를 수행한다. 이는 탐지 시간이 빠르지만 겹친 객체에 대한 탐지 성능이 떨어진대[11][12].

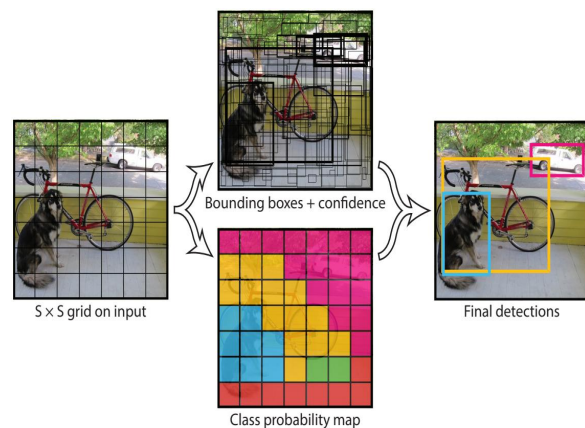


Fig. 1. YOLO v1 Model

YOLO v2의 원래 이름은 YOLO 9000으로 완전 연결 계층(Fully Connected Layer)을 없애 각 셀(Cell)에 대해 여러 객체(Object)에 대한 탐지가 가능하여 속도와 정확도를 개선했다. YOLO v3는 서로 다른 크기의 스케일(Scale)을 가진 앵커 박스(Anchor Box)를 이용해 객체 탐지하여 정확도가 향상되었다[11-13]. 그리고 YOLO v3는 Darknet-53을 사용하여 Backbone의 성능을 향상시켰으며, 이는 ResNet과 유사한 형태로 개발되었다. 또한 예측한 Bounding box에서 Sigmoid를 기반으로 하는 Logistic classifier를 통해 각 객체들의 개별 위치에 대해 Multilabel Classification을 수행한다. 그리고 서로 다른 크기와 scale을 가지는 Anchor box로 객체를 탐지하여 기존의 v1, v2 대비 향상된 성능을 보여 준다[9].

YOLO v3는 사용이 용이하며, 기존 버전보다 뛰어난 이미지 인식 성능으로 빠른 처리시간이 필요한 시스템에 최적화되어 있다[14]. 기존 R-CNN은 Proposal 수가 많으며 그로 인한 오버헤드가 많이 발생하기 때문에 이미지나 영상을 처리하면서 상당히 많은 시간이 소요되기 때문에 시간적 소모가 많았다. YOLO v3는 이러한 시간적 소모를 줄이기 위해 제안되었다[14-16].

### III. The Proposed Scheme

본 논문에서 제안된 모델은 딥러닝 기반의 객체 탐지 알고리즘이 탑재된 YOLO v3를 기반으로 진행하였다. 데이터셋의 경우 깃허브(GitHub)의 데이터[17]를 참조했고, 그 중 Person Object에 대한 데이터를 별도로 추출하여 본 논문에 적용하였다. 다음의 Fig. 2와 Table. 1은 본 논문에서 제안한 시스템 모델의 구성도와 세부 내용이다.

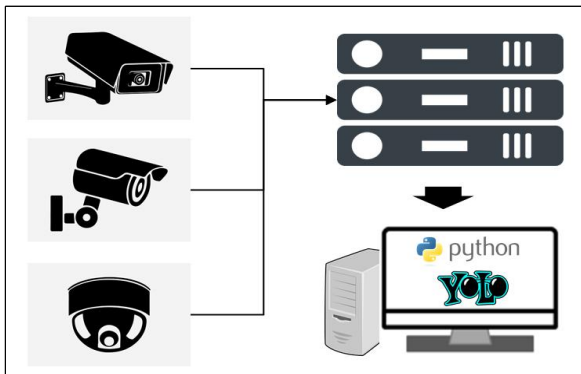


Fig. 2. System Model Architecture

Table 1. System model contents

CCTV	Count	12 Ea
Server	disk capacity	1 TByte
OS	Version	Windows 10

CCTV는 총 12대를 사용했고, CCTV 영상파일을 저장하는 서버의 용량은 1 TByte이다. 도심에서 약간 떨어진 중형 마트에 설치했다. CCTV의 영상은 1시간 단위로 녹화되며, 녹화된 파일은 서버에 저장되고, 서버에 저장된 영상파일은 매 시 30분이 되면 자동으로 PC에 전송하도록 시스템을 구성하였다. PC의 OS는 Windows 10이며, 작업 스케줄러를 통해 Person Object만 탐지하게 수정된 데이터셋에 YOLO v3 라이브러리를 포함한 파이썬으로 프로그래밍 파일을 실행하도록 구성하였다.

본 논문에서는 각 영역에 따라 세 가지의 프로세스로 나뉘며, 영역에 따른 프로세스는 다음과 같다. 다음의 Fig. 3은 CCTV 영상파일을 처리하는 프로세스 과정이다.

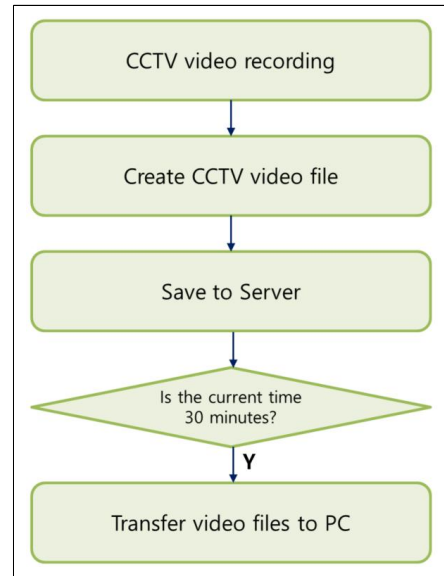


Fig. 3. CCTV video file processing process

곳곳에 설치된 CCTV를 통해 영상을 녹화하고, 녹화된 영상파일을 생성한다. 영상파일은 1시간 단위로 생성되며, 파일명은 '연월일\_시\_CCTV Name'의 형태로 저장하고, 확장자는 avi를 사용한다. CCTV 녹화로 생성된 영상파일은 서버에 저장하고, 매 시 30분에 PC로 저장된 영상파일을 전송하는 프로세스로 설계하였다.

다음의 Fig. 4는 작업스케줄러에 의해 파이썬으로 구현된 실행 파일이 실행되는 처리 과정이다.

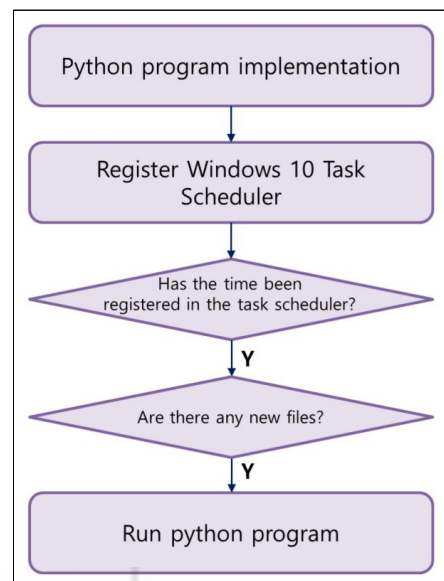


Fig. 4. Python program runs process

프로그램은 파이썬으로 YOLO v3 라이브러리를 이용하여 코딩했으며, 데이터 셋은 범범의 목적을 두어 사람 객체만 탐지하기 위해 Person Object만 탐지하도록 수정하였다. 해당 프로그램은 Windows 10의 작업스케줄러에 등록했으며, 작업스케줄러는 정시일 경우 실행되도록 설정하였다. 서버로부터 전송받은 파일이 있는지 여부를 체크한 뒤, 전송받은 파일이 있을 경우 파이썬 프로그램을 실행하도록 설정하였다.

다음의 Fig. 5는 파이썬 프로그램이 실행되어 CCTV 영상파일에서 YOLO v3 라이브러리를 이용하여 탐지된 객체를 찾고 저장하는 처리 과정이다.

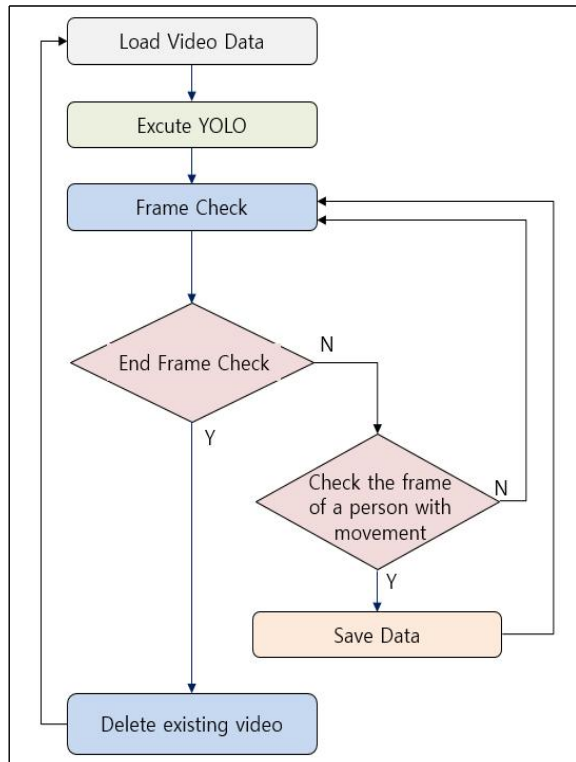


Fig. 5. Object detection process

작업스케줄러에 의해 파이썬 프로그램이 실행되면 서버로부터 전송받은 영상파일을 로드한 뒤 YOLO v3 라이브러리를 이용한 Person Object를 탐지하게 된다. 영상 프레임을 체크하여 움직임이 있는 프레임 즉, Person Object 프레임을 찾을 경우 해당 프레임을 시, 분, 초의 워터마크와 함께 저장하고 그렇지 않을 경우 skip 하는 과정을 반복한다. 마지막 프레임까지 확인을 마치면 저장된 프레임을 하나의 파일로 합치는 인코딩 과정을 진행하고, 기존 영상파일을 삭제한다.

#### IV. Test and Result

CCTV 영상은 도심에서 약간 떨어진 중형 마트에서 추출하였으며, 각 CCTV는 마트의 입구와 내부, 창고 등에 설치되어 운영되었다.

다음의 Fig. 6은 본 논문에서 제안한 모델을 적용하기 전 서버에서 저장하고 있던 파일을 PC로 옮겨 확인한 파일 리스트의 일부이다.

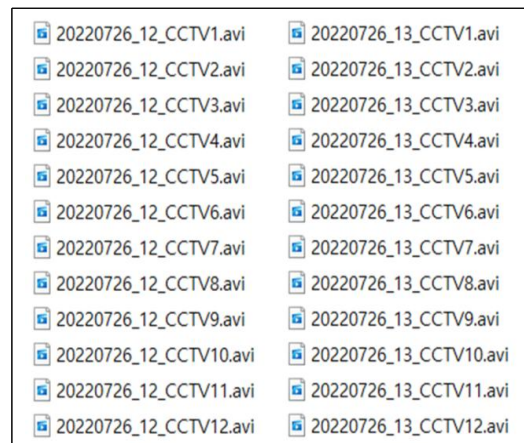


Fig. 6. CCTV video file

다음의 Table. 2는 본 논문에서 제안한 모델을 적용하기 전 서버에서 저장하고 있던 각각의 CCTV 영상파일의 크기이며, 이는 하루 24시간을 기준의 크기를 나타낸다.

Table 2. Data before model application

CCTV Name	File Size
CCTV1	5.96 GByte
CCTV2	6.02 GByte
CCTV3	5.96 GByte
CCTV4	6.02 GByte
CCTV5	5.90 GByte
CCTV6	5.96 GByte
CCTV7	5.90 GByte
CCTV8	5.98 GByte
CCTV9	5.92 GByte
CCTV10	6.03 GByte
CCTV11	5.97 GByte
CCTV12	6.09 GByte

12대의 CCTV에서 하루에 저장되는 파일 크기는 작게는 5.90GByte에서 크게는 6.09GByte로 생성되었으며, 서버에 저장된 파일을 확인했을 때 13일 동안의 파일이 유지되고 있었다. 서버의 저장공간은 1TByte 이며, 약 950GByte를 사용 중으로 확인되었다.

다음의 Table. 3은 본 논문에서 제안한 모델을 적용 전과 후를 비교한 표이다. 24시간 저장된 영상파일의 크기를 비교하였다. 제안 모델 적용 전후를 비교했을 때, 평균 94.9% 절감되었으며, 최소 84.0%에서 최대 99.2%까지 영상파일의 크기가 절감됨을 확인할 수 있었다. 이 과정에서 Person Object의 인식률은 98.6%로 나타났으며, 프레임 단위로 인식을 진행했기 때문에 객체 탐지에는 영향이 없는 것으로 볼 수 있다.

Table 3. Comparison before and after model application

CCTV Name	File Size		savings rate
	Before	After	
CCTV1	5.96 GByte	0.95 GByte	84.0%
CCTV2	6.02 GByte	0.54 GByte	91.0%
CCTV3	5.96 GByte	0.66 GByte	89.0%
CCTV4	6.02 GByte	0.18 GByte	97.0%
CCTV5	5.90 GByte	0.41 GByte	93.0%
CCTV6	5.96 GByte	0.09 GByte	98.5%
CCTV7	5.90 GByte	0.06 GByte	99.0%
CCTV8	5.98 GByte	0.05 GByte	99.2%
CCTV9	5.92 GByte	0.18 GByte	97.0%
CCTV10	6.03 GByte	0.42 GByte	93.0%
CCTV11	5.97 GByte	0.09 GByte	98.5%
CCTV12	6.09 GByte	0.06 GByte	99.0%

모델 적용 전의 경우 71.70GByte의 용량을 사용하고 있었으나, 모델 적용 후의 경우 3.69GByte로 줄어들었다. 모델 적용 전의 경우 13일이 지난 시점부터 삭제되고 있었으나, 모델 적용 후의 경우 약 20배의 보관 기간이 증가한 것으로, 이는 약 260일까지 보관할 수 있을 것으로 예측할 수 있다.

## V. Conclusions

사회안전망의 일원인 CCTV는 방범, 화재 예방, 감시 등 재난·재해 및 안전을 위해 사회 곳곳에 설치되어 운영되고 있으며, CCTV 개수는 점점 증가하고 있다. 영상은 더 선명하게 개체를 인식하기 위해 고화질, 초고화질 등 화소(Pixel)가 많아지면서 화질 또한 개선되고 있다. 개수의 증가와 화질의 개선으로 비례적으로 영상파일의 크기도 증가하고 있다. 이에 따라 영상파일을 저장하고 있는 디스크를 업그레이드하거나, 추가 디스크를 배치하는 등의 노력이 이뤄지고 있다. 그렇지 않을 경우 기존 영상을 삭제하고 새로운 영상을 저장하고 있는 방식을 이용하고 있으나, 이러한 이유로 지난 영상을 확인하기가 어려운 상황이 발생하기도 한다.

본 논문에서는 인공지능 분야 중 하나인 영상분석을 이용한 컴퓨터 비전을 이용한 것으로, CCTV 영상 속의 특정 객체를 YOLO v3 라이브러리를 활용하여 탐지하여 해당 프레임만을 저장하여 불필요한 프레임을 삭제하는 모델을 제안하여 영상파일의 크기를 줄임으로써 저장공간을 확보하고, 그로 인해 더 오랜 기간 영상을 저장·관리 할 수 있도록 하였다.

제안 모델 적용 후 최소 84.0%에서 최대 99.2%까지 영상파일의 크기가 절감됨을 확인하였으며, 저장공간이 추가로 확보됨에 따라 제안 모델을 적용하기 전보다 약 20배의 보관 기간이 증가했음을 확인할 수 있었다. 또한, 프레임 추출 후 재 인코딩하는 과정에서 별도의 변환 없이 원본 그대로를 유지했기 때문에 영상의 품질은 기존 영상의 품질과 동일하게 유지할 수 있었다.

CCTV 설치 위치에 따라 객체 존재 여부에 따른 탐지가 다르기 때문에 CCTV별 감소율의 차이가 있을 수 있으며, 사람이 지속적으로 통행하는 곳에서는 감소율이 현저히 낮을 것으로 예상된다. 이러한 이유로 본 제안 모델은 사람의 통행이 적거나 심야 시간 등 인적이 드문 곳에 활용하면 감지되는 객체가 적기 때문에 불필요한 데이터의 삭제가 많아지게 되어 저장되는 영상파일의 용량이 적어 더 많은 저장공간을 확보할 수 있을 것으로 기대할 수 있다.

본 논문에서 제안한 모델을 이용하여 CCTV 뿐만 아니라 차량용 블랙박스의 영상파일이나, 촬영용 드론으로 얻은 영상파일에서 특정 객체를 탐지하여 영상파일의 크기를 줄일 수 있는 응용도 가능하다.

## REFERENCES

- [1] S. H. Choi, D. H. Kim, H. H. Kim, Y. Kim. "Disaster warning system using Convolutional Neural Network - Focused on intelligent CCTV", Journal of the Korea Society of Computer and Information, Vol. 24, No. 2, pp. 25-33, 2019.
- [2] Y. G. Kwak, T. H. Lim, "A Study on the CCTV Effective Utilization Method for the Crime Prevention and Action", Journal of Korean Public Police and Security Studies, Vol. 8, No. 2, pp. 119-144, 2011. DOI : 10.25023/kapsa.8.2.201108.119
- [3] I. S. Kim, J. D. Yoo, B. H. Kim, "A Monitoring Way and Installation of Monitoring System using Intelligent CCTV under the u-City Environment", The Journal of The Korea Institute of Electronic Communication Sciences, Vol. 3, No. 4, pp. 295-303, 2008.
- [4] K. T. Kim, D. S. Seong, K. B. Lee, "A Smart Remote Surveillance System using GPS Correction and Smart Device", Journal of

Korean Institute of Information Technology, Vol. 10, No. 11, pp. 181-191, 2012.

- [5] I. S. Kim, H. S. Shin, "A Study on Development of Intelligent CCTV Security System based on BIM", The Journal of The Korea Institute of Electronic Communication Sciences, Vol. 6, No. 5, pp 789-795, 2011.
- [6] Y. H. Kim, J. H. Kim, "Development of Real-Time Face Region Recognition System for City-Security CCTV", Journal of Korea Multimedia Society, Vol. 13, No. 4 pp. 504-511, 2010.
- [7] D. S. Seong. "Improvement of Safe Return Service by CCTV Camera Sharing", The Journal of Korean Institute of Next Generation Computing, Vol. 13, No. 1, pp. 65-73, 2017.
- [8] J. H. Kim, H. S. Moon, "A model development for object detection and classification at sea using modified YOLO v4", Journal of the Military Operations Research Society of Korea, Vol. 47, No. 2, pp. 13-25, 2021.
- [9] Redmon, J., & Farhadi, A. YOLO9000: better, faster, stronger. In Proceedings of the IEEE conference on computer vision and pattern recognition, 7263-7271. 2017.
- [10] Redmon, J., & Farhadi, A. Yolov3: An incremental improvement. arXiv preprint, arXiv:1804.02767. 2018.
- [11] H. J. Shin, S. M. Oh, S. W. Oh, Y. G. Kim, T. W. Um, J. S. Kim. "Implementation of Mobility Detection Model in Media using YOLO v3", Proceedings of Symposium of the Korean Institute of communications and Information Sciences. pp. 967-968, 2022.
- [12] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. "You only look once: Unified, real-time object detection". In Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 779-788. 2016. DOI : 10.48550/arXiv.1506.02640
- [13] Y. J. Lee, H. Y. Son, K. W. Choi, S. K. Lee, J. H. Park, "Object Detection Based on Deep Learning Using Port CCTV Video", Proceedings of Symposium of the Korean Institute of communications and Information Sciences. pp. 796-797, 2021.
- [14] G. H. Jo, H. W. Cho, Y. J. Song. "Implementation of 360-degree Image Recognition System Using Lidar Sensors and Yolo v3 Libraries in Cloud Server Environment", The Journal of Korean Institute of Information Technology. Vol. 18, No. 2, pp.1-7. 2020. DOI : 10.14801/jkiit.2020.18.2.1,
- [15] S. H. Shim, S. I. Choi, "Development on Identification Algorithm of Risk Situation around Construction Vehicle using YOLO-v3", Journal of the Korea Academia-Industrial cooperation Society, Vol. 20, No. 7, pp. 622-629, Jul. 2019.
- [16] Tian, Yunong, Yang, Guodong, Wang, Zhe, Wang, Hao, Li, En, and Liang, Zize, "Apple detection during different growth stages in orchards using the improved YOLO-V3 model", Computers and electronics in agriculture, Vol. 157, pp. 417-426, Feb. 2019.
- [17] YoloV3 Implemented in TensorFlow 2.0, <https://github.com/zzh8829/yolov3-tf2>

## Authors



Seong-Ik Kim received the B.S. degree in Multimedia Engineering from Kongju National university, Korea in 2011. The M.S. and Ph.D. degrees in Computer Science from Kongju National university, Korea, in 2017,

2021, respectively. Dr. Kim joined the faculty of the Smart IT at U1 University, Asan, Korea, in 2021. He is currently a Professor in the Smart IT, U1 University. He is interested in IT convergence, artificial intelligence, bigdata, computer vision, image processing.