



## Journal of Knowledge Information Technology and Systems

ISSN 1975-7700 (Print), ISSN 2734-0570 (Online)

<http://www.kkits.or.kr>

# Real-time Object Recognition and Translation App for Foreign Workers Using Artificial Intelligence

**Bum-Soo Kim, Young-Hyen Kim, Byung-Tae Chun\***

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

### ABSTRACT

The number of foreign workers entering Korea is increasing every year. Accordingly, Korean language education and education on understanding Korean culture are provided at training institutions of government departments. Despite these efforts, it can be seen that foreign workers face many difficulties when they first encounter Korean culture. In this paper, we introduce a public app developed to solve the cultural difficulties of foreign workers. Among the popular objects in Korea, a study was conducted on objects that foreign workers feel somewhat unfamiliar to them. We developed an app that recognizes an object in real time on a smartphone and translates its usage into several languages. Language selection was based on the languages of five countries with a large portion of entry into Korea. 7,900 copies of training images were collected and labeled. A total of 20,000 iterations were trained with the Yolov3-tiny learning model using Darknet, an AI neural network. The developed app can recognize and translate real-time objects on a smartphone with an average accuracy of 91.3%. It is a system that translates and provides the description of the detected object into five languages at the touch of a button. It is expected that this app will be able to bridge cultural and linguistic differences in daily life of foreign workers.

© 2020 KKITS All rights reserved

**KEYWORDS :** Yolov3-tiny training model, Darknet-53 AI neural networks, Android smartphones, Translation data, Real-time detection, OpenCV, Image dataset.

**ARTICLE INFO:** Received 10 November 2020, Revised 27 November 2020, Accepted 11 December 2020.

\*Corresponding author is with the School of Computer Engineering & Applied Mathematics, Hankyong National University, Jungang-ro, Anseong-si, Gyeonggi-do. 17579. Korea

*E-mail address:* [chunbt@hknu.ac.kr](mailto:chunbt@hknu.ac.kr)

## 1. 서론

1988년 서울 올림픽 개최와 2002년 한일 월드컵의 개최로 우리나라의 경제성장이 외국에 알려지면서 동남아시아에서 근로자로 입국하는 사례가 증가하였다[1]. 소득 1만 달러 시대의 ‘국내 3D 업종 기피 현상’은 공단지역, 농지 지역의 외국인 노동자의 유입을 가속화 하였다. 이로 인하여 91년 산업연수생 제도를 도입했고 94년, 이런 영향이 중소기업으로 확대되면서 외국인 근로자 러시가 본격화되었다[2].

현재 국내 거주 중인 외국인 근로자는 국내 총 인구의 4%나 된다[3]. 이들은 국내 입국 후 언어 소통, 문화적 차이로 인해 적응에 어려움을 겪는다[4]. 이를 해결 하고자 ‘한국 산업인력공단’에서 외국인 근로자들을 위한 취업 교육, 한국어 교육, 한국문화의 이해에 대한 강의를 기본적으로 진행하고 그 이외에 각 국가의 정부 부처의 공공직업 훈련기관에서 국내에 입국하기 전에 교육을 실시한다. 또한, 노동자 지원센터에서는 한국에서의 생활이 어려운 근로자들을 위한 도움센터를 운영한다. 이처럼 근로자들의 한국 생활 적응에 크나큰 어려움이 있기에 위와 같은 다양한 기관에서 지원을 아끼지 않는다[5]. 그럼에도 불구하고, 현실적인 해결책이 되지 못하고 있다고 외국인 근로자들은 말한다. 문화적 소외감으로 생긴 스트레스는 업무에 지장을 주고, 결국 외국인 근로자들은 업무에 집중하지 못한다는 고정관념이 생기게 된다[6]. 문화적, 언어적 소외감을 줄이기 위해선 가장 기본적인 일상생활에서의 해결책이 필요하다. 한국에 처음 와서 소외감을 느끼는 경우는 당연히게도 사용하는 언어의 차이와 일상에서 사용하는 물건의 차이이다[7]. 한국에서만 사용하는 다양한 물건들이 있고 이는 외국인들에게 다소 생소하게 느껴진다[8]. 예를들어, ‘이태리 타올’을 수건처럼 사용하

는 일들이 실제로 벌어진다.

기존에 존재하는 사물 인식 앱들은 기본적으로 제공되는 이미지 데이터셋트를 사용하였기에 기본적으로 정해져있는 사물만 인식하였다. 또한, 글자를 번역해주는 앱은 여럿 존재하지만, 사물을 인식하여 사물의 설명을 번역해주는 앱은 찾아볼 수 없다. 이런 문제를 해결하고자 한국 고유의 다양한 물건들을 실시간으로 인식하여 그 물건의 쓰임새를 외국인 근로자들의 언어로 번역하는 앱 개발이 필요하다.

본 논문의 구성은 다음과 같다. 2장에서는 전체 시스템의 구조, 3장에서는 Yolo 딥러닝 모델 및 학습, 4장에서는 실험 및 구현, 마지막으로 5장에서 결론을 기술한다.

## 2. 시스템 구조

본 논문의 전체 시스템 구조는 <그림 1>과 같다. 메인 과정은 총 세 가지로 구분된다.

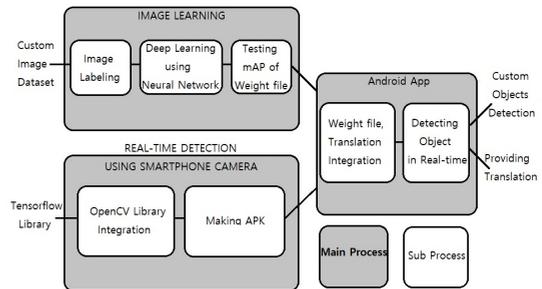


그림 1. 시스템 구조  
Figure 1. System structure

인식할 사물의 커스텀 이미지를 인공 신경망을 통해 딥러닝 시키는 이미지 러닝 과정이 첫 번째 과정이다[9][10]. 이 과정을 통해 생성된 가중치 파일을 스마트폰에서 활용할 수 있게 OpenCV 라이브러리를 안드로이드 앱에 빌드하는 과정이 두 번

째이다. 마지막으로 카메라를 사용하여 실시간으로 인식된 사물을 번역 데이터와 연동하여 사용자로 하여금 번역을 실시간으로 확인 할 수 있도록 개발하는 과정이 마지막 세 번째 단계이다.

### 3. Yolo 딥러닝 모델 및 이미지 학습

사물을 실시간으로 검출하기 위해선 먼저 대량의 사물 이미지를 클래스별로 분류하여 사물의 위치 좌표를 구한다. ‘Yolo’ 모델[11]을 사용하였기에 아래 <그림 2>처럼 그에 맞는 경계 상자 데이터를 생성한다. 각 경계 박스마다 depth 방향으로 (4 + 1 + 5) \* 3 = 30개의 depth가 쌓인다[12].

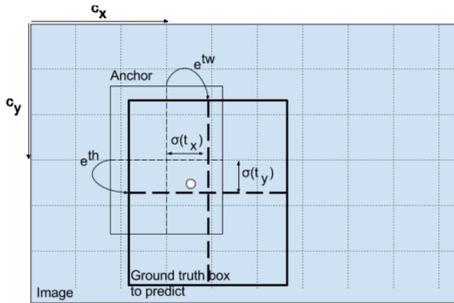


그림 2. Yolo 모델의 경계 박스  
Figure 2. The bounding box of the Yolo model

이미지별 경계 박스 크기는 아래 <그림 3>과 같이 13\*13, 26\*26, 52\*52인 3가지 종류로 나뉜다[11].

$$10,647 \text{ boxes} = ((52 \times 52) + (26 \times 26) + (13 \times 13)) \times 3$$

(anchor box count). 416x416 pixels

<그림 4>와 같이 나머지 부분은 기존에 개발된 CNN(Convolutional Neural Network), FCN(Fully Convolutional Network), ResNet, DarkNet 등을 이용한다[12].

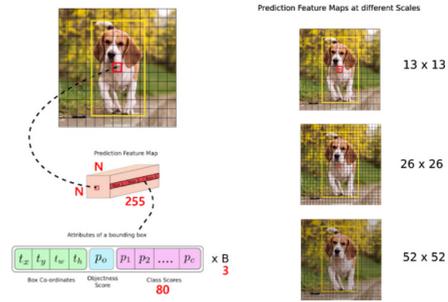


그림 3. 이미지 크기별 경계 박스  
Figure 3. Anchor size by image size

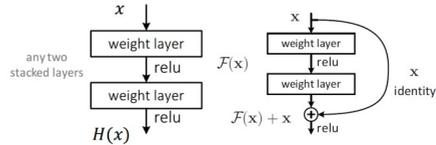


그림 4. ResNet 핵심 구조 (좌=기준, 우=ResNet)  
Figure 4. Core structure of ‘ResNet’

Up-sampling은 저해상도에서 고해상도로 이미지 스케일업할 때 사용된다. 이는 보간법과 유사한 방식으로, Transpose convolution, Deconvolution으로 불린다. 이런 입출력을 고려해 CNN 레이어를 개발하면, 다음과 같은 최종 ‘Yolo’ 딥러닝 모델이 만들어진다[13]. 이 신경망 모델의 설정값을 인식할 사물들의 특성에 맞게 변경하여 원하는 사물만 인식할 수 있도록 학습하였다.

학습에 사용한 모델은 위에 설명한 ‘Yolo’ 모델 중 안드로이드 플랫폼에서 가동할 수 있도록 제작된 ‘Yolov3-tiny’를 사용하였다[14]. ‘Yolov3-tiny’는 기본 ‘yolov3’보다 훨씬 적은 수의 convolution 레이어를 갖고 있다. 이는 많은 양의 메모리를 차지할 필요가 없으므로 하드웨어의 자원 필요성이 줄어들고 감지 속도는 크게 증가한다. 반면에 레이어의 수가 줄어들어 사물의 인식 정확도가 다소 떨어질 수 있지만[15], 스마트폰의 한정적인 시스템 자원에 맞게 구성하기 위해서는 ‘Yolov3-t

iny' 모델이 학습 모델 중 가장 최선이라 할 수 있다. 본 논문에서 개발한 앱은 외국인 근로자들이 쉽게 사용할 수 있어야 하고 언제, 어디서든 작동해야 하므로 최적 사양을 낮춰 2016년 이후에 출시한 모든 안드로이드 기반 스마트폰에서 무리 없이 작동할 수 있도록 충족해야 했고 스마트폰 시스템 사양[16]을 만족하기 위해서는 'Yolov3-tiny' 학습 모델은 최적의 선택이다.

본 논문의 시스템에서는 총 다섯 가지의 사물인 '전기파리채', '모기향', '전기밥솥', '이태리타울', '강통의자'를 인식할 수 있도록 학습하였다. 다섯 가지의 사물을 클래스별로 분류하고 <그림 5>에서처럼 다양한 각도, 형태, 색상, 크기 그리고 밝기 값이 다른 총 7,900장의 이미지 데이터를 각각 라벨링 처리하였다.

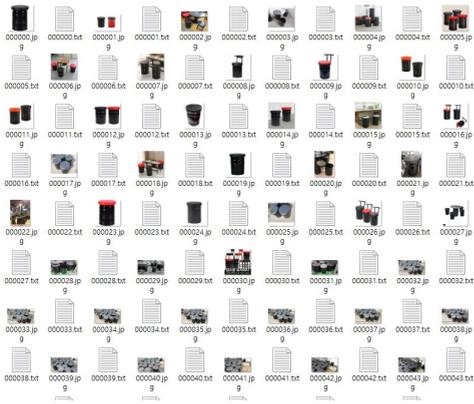


그림 5. 라벨링된 7,900장의 이미지  
Figure 5. 7,900 labeled images

라벨링 처리된 대량의 이미지들은 각기 다른 앵커 값을 갖는다. 모든 앵커값에 대해 일일이 학습하게 되면 학습 속도가 느려질 뿐만 아니라 각 사물의 인식 정확도 또한 현저하게 떨어진다. 이를 해결하기 위해 IoU 값을 계산한다. IoU란, 두 영역의 교차 영역의 넓이를 두 영역의 합으로 나눈 값

이다. 사물 검출에서 예측된 앵커 박스의 정확도를 평가하는 지표 중 하나로, 예측된 앵커 박스와 실제 앵커 박스의 IoU를 해당 앵커 박스의 정확도로 간주한다. 사물 인식 모델의 정확도는 테스트용 데이터셋을 실행해본 다음에 precision-recall curve와 average precision을 구해 수치화 할 수 있는데, 이때 IoU값을 사용하게 된다. 정확하고 빠른 반복 학습을 위해 <그림 6>과 같이 IoU 계산식을 이용하여 라벨링된 앵커의 위치 평균값을 계산하고 이 계산 값을 cfg 파일에 설정하여 <그림 7>의 IoU를 81.17%까지 충족시켰다.

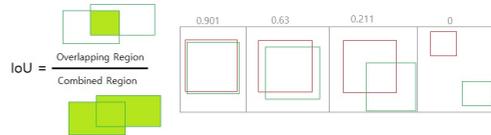


그림 6. IoU 계산식  
Figure 6. IoU calculation formula

```

C:\darknet-master\build\darknet%64>
C:\darknet-master\build\darknet%64>darknet.exe detector calc_anchors data\obj\data -num_of_clusters 6 -w 416 -h 416
CPU version: 10010 (11010), cuDNN: 7.6.5, CUDNN_HALF=1, GPU count: 1
CUDNN_HALF=1
OpenCV version: 4.0.1

num_of_clusters = 6, width = 416, height = 416
read labels from 555 images
loaded image 554 box: 1168
all loaded.
calculating k-means+ ...
iterations = 20

counters_per_class = 1168

avg IoU = 81.17 %
Saving anchors to the file anchors.txt
anchors = 112,151, 227,148, 151,250, 377,225, 225,365, 381,381
    
```

그림 7. 평균 앵커, IoU  
Figure 7. Anchors average, IoU

이미지 학습에서 중요하게 봐야 할 요소는 반복 학습 횟수인 max\_batches 값과 반복 학습을 하는 동안 이미지 손실률인 avg\_loss이다. batch 값을 최대한 올리면 반복 학습 횟수가 증가하고 반복 학습을 하면 할수록 인식률은 높아진다. 손실률 또한 0에 가까운 값으로 수렴한다. 그러나 인식 정확도가 높아진다는 말은 학습에 사용한 이미지만을 정

확하게 인식한다는 뜻이기 때문에 같은 사물이라도 방향, 색상 또는 크기가 학습된 이미지와 조금이라도 다르게 되면 인식하지 못하게 되어 일반화에 실패하고 만다. 사물 인식률은 90% 이상 유지하고 손실률은 1 이하로 유지할 수 있도록 max\_batches 값은 <그림 8>과같이 iteration을 20,000회로 설정하여 각 클래스당 반복 학습 횟수를 4,000번으로 조절하였고 avg loss는 약 0.2까지 안정화하였다. 라벨링된 이미지는 ‘Darknet’ 신경망을 통해 ‘Yolov3-tiny’ 학습 모델로 이미지를 학습시켰다.

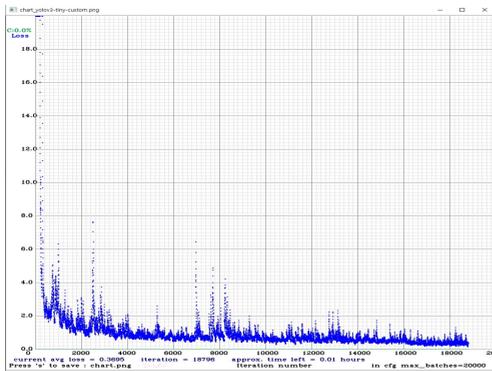


그림 8. 반복 학습 횟수별 avg loss  
Figure 8. avg loss per training iteration

#### 4. 실험 및 구현

반복 학습을 1,000회 할 때마다 가중치 파일을 생성하였고 총 max\_batches 값인 20,000회의 반복 학습 중 인식률이 가장 좋은 19,000회 제의 학습 결과는 <그림 9>의 내용과 같았다. 최종적으로 91.3% 인식 예측률을 구현한 weights 파일을 생성할 수 있었다. 동영상 테스트를 통해 움직이는 영상에서의 사물도 정확하게 인식하고 여러 가지 사물이 겹쳐 있어도 충분히 인식이 가능할 정도로 mAP가 높음을 확인하였다.

```
<19000>
class_id = 0, name = Canchair, ap = 80.07% (TP = 1034, FP = 138)
class_id = 1, name = Mogihyang, ap = 89.98% (TP = 485, FP = 65)
class_id = 2, name = FlyKiller, ap = 97.68% (TP = 1069, FP = 20)
class_id = 3, name = RiceCooker, ap = 98.91% (TP = 3116, FP = 111)
class_id = 4, name = ItalyTowel, ap = 90.27% (TP = 647, FP = 14)

for conf_thresh = 0.25, precision = 0.95, recall = 0.86, F1-score = 0.90
for conf_thresh = 0.25, TP = 6351, FP = 348, FN = 1019, average IoU = 76.45 %

IoU threshold = 50 %, used Area-Under-Curve for each unique Recall
mean average precision (mAP@0.50) = 0.913845, or 91.38 %
```

그림 9. 학습이 완료된 가중치의 인식률  
Figure 9. Recognition rate of weights that have been trained

위의 과정을 거쳐 높은 정확도로 사물 인식이 가능한 가중치 파일이 생성되면 이를 다양한 언어로 번역할 수 있는 번역 데이터와 연동한다. 법무부 통계에 따르면 국내 거주 중인 외국인인은 약 220만 명이고 미국과 중국을 제외한 가장 많은 외국인 수를 차지하는 베트남, 태국, 우즈벡, 인도네시아 4개국을 선정하여 이들의 언어로 사물의 이름과 설명을 번역하였다.

번역 데이터는 번역 API를 사용하지 않고 외국인 근로자들의 스마트폰 사용 환경을 고려하여 <그림 10>과 같이 직접 구글 번역기를 사용하여 안드로이드 asset folder에 string 자료형으로 입력하였다.

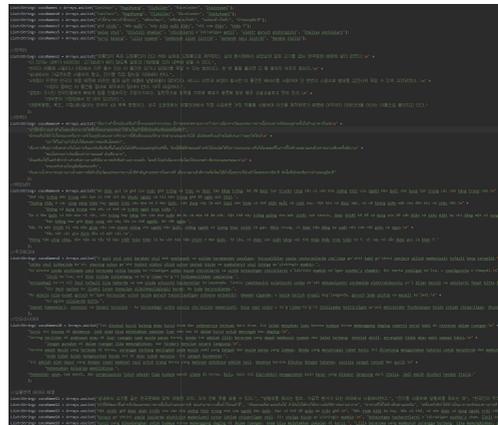


그림 10. Asset folder의 번역 데이터  
Figure 10. Translation data in the asset folder





그림 14. 인식 화면에 설명 바로 출력  
Figure 14. printing out the object translation direct on the detection window

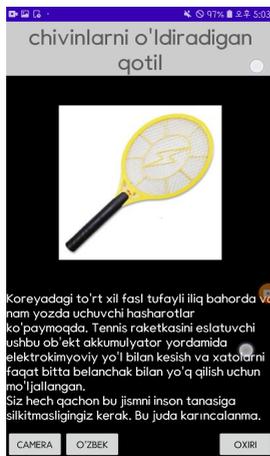


그림 15. 사물을 인식 및 번역한 앱 화면  
Figure 15. App screen that recognizes and translates objects

카메라 화면에 인식할 사물이 없는 경우에는 인식 오류를 방지하기 위해서 인식할 사물이 없음을 <그림 16>과 같이 팝업으로 알려준다.

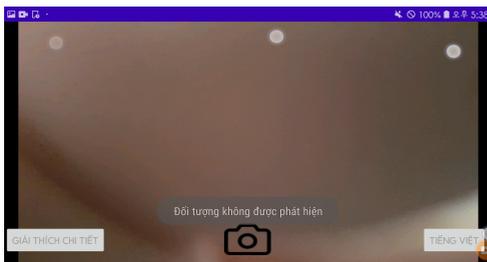


그림 16. 인식할 사물이 없는 경우에 대한 처리 방법  
Figure 16. What to do when there is no object to recognize

화단의 상세페이지 버튼과 언어 선택 버튼을 비활성 처리하여 사물을 다시 인식하도록 하였고 선택된 언어에 맞게 팝업창의 설명도 번역되도록 구현하였다.

## 5. 결론

외국인 근로자들이 한국 생활 적응에 크나큰 어려움을 겪고 있으며, 이를 해결하고자 정부의 다양한 노력에 불구하고 문화적 소외감으로 생긴 스트레스는 업무에 지장을 주고, 한국 생활에 부적응 현상을 보여주고 있다. 본 논문에서는 외국인 근로자들은 문화적, 언어적 소외감을 줄이기 위해서 한국문화에 독특한 사물을 인공지능 기법을 이용하여 번역해주는 시스템을 개발하였다. 본 연구를 통해 외국인 근로자들이 한국 생활 정착에 도움이 되기를 희망한다.

## References

- [1] J-S. Han, *Koreans and korean culture viewed by foreign workers from cambodia*, Major in Multi-Cultural Education Seoul National University of Education Graduate School of Education master's thesis, 2018.
- [2] J-D. Park, *A study on the actual conditions of foreign workers in livestock industry*, Department of Food Marketing & Economics Graduate School of Agriculture & Animal Science Kon kuk University master's thesis, 2013.
- [3] Y-H. Lee, *Effect of foreign workers' domestic travel motivation, attitude, and sociocultural adaptation on behavioral intention : applying leisure restriction negotiation theory*, Department of Tourism

- Graduate School of Kyung Hee University  
Doctoral dissertation, 2019.
- [4] J-Y. Ko, *A study on the life and adaptation of foreign workers-focusing on foreign workers engaged in fishery in Jeju island-*, Major in Multi-Cultural Education Graduate School of Education, Gyeongin National University of Education master's thesis, 2020.
- [5] B-J. Kim, and J-K. Kim *The impact of acculturation stress on subjective quality of life among migrant workers: focusing on the moderating effect of social support*, Korea Academy of Care Management, No. 28, pp. 57-90, 2018.
- [6] H-S. Jung, *A qualitative exploratory study on intercultural communication conflict : a case analysis of lived experiences of roreign factory workers in Korea*, Korean Journal of Communication Studies, Vol. 12, No. 12, pp. 27-45, 2004.
- [7] S-H. Lee, Y-J. Lee, S-Y. Kim, and S-J. Kim, *Social support and acculturative stress in migrant workers*, Journal of Korean Academy of Nursing, Vol. 39, No. 6, pp. 899-910, 2009.
- [8] National Folk Museum of Korea, youtube channel, [https://www.youtube.com/channel/UCUN2CadW\\_WjICFZfyNfqGXg/videos](https://www.youtube.com/channel/UCUN2CadW_WjICFZfyNfqGXg/videos) Oct. 2020.
- [9] D-H. lee, and B-Y. lee, *A study on the analysis of Jeju island precipitation patterns using the convolution neural network*, Korea software sssessment and valuation society, Dec. 2019.  
<http://dx.doi.org/10.29056/jsav>. Dec. 2019.
- [10] H-S. Kang, *A study on the establishment of the IDS using machine learning*, Korea Software Assessment and Valuation Society, Dec. 2019.  
<http://dx.doi.org/10.29056/jsav>. Dec. 2019.
- [11] J. Redmon, and A. Farhadi, *Yolov3: An incremental improvement*, p. 1, 2018.
- [12] Daddy Makers, Data labelling based on yolov3 deep learning, <http://daddynkidsmakers.blogspot.com/2020/05/yolo.html>, Sept. 2020.
- [13] J. Redmon, S. Divvala, R. Girshick, and Ali Farhadi, *You only look once: unified, real-time object detection*, IEEE Conference on Computer Vision and Pattern Recognition(CVPR), pp. 779-788, 2016.
- [14] X. Zhang, X. Dong, Q. Wei, and K. Zhou, *Real-time object detection algorithm based on improved YOLOv3*, Journal of Electronic Imaging, Vol. 28, No. 5, pp. 53022, 2019.
- [15] Y. Zhang, Yongliang, and Jun, Zhang, *An imporved tiny-yolov3 pedestrian detection algorithm*, Optik, Vol. 183, No. 3 pp. 17-23, 2019.
- [16] Galaxy S7 Smartphone(SM-G930S), SAMSUNG.  
<https://www.samsung.com/sec/support/model/S-M-G930SZWASKO> Oct. 2020.
- [17] Kari Pulli, Anatoly Baksheev, Kirill Korniyakov, and Victor Eruhimov, *An imporved Real-time computer vision with OpenCV*, Communications of the ACM, Vol. 55, No. 6 p. 61, 2012.
- [18] Ivangrov, Anroid-Deep\_Learning-with-OpenCV, <https://github.com/ivangrov/Android-Deep-Learning-with-OpenCV> Sept. 2020.

## 인공지능을 이용한 외국인 근로자를 위한 실시간 사물 인식 및 번역 앱

김범수<sup>1</sup>, 김영현<sup>2</sup>, 전병태<sup>3</sup>

<sup>1,2</sup>한경대학교 컴퓨터응용수학부 학부생

<sup>3</sup>한경대학교 컴퓨터응용수학부 교수

### 요 약

한국에 입국하는 외국인 근로자의 수는 매년 증가하는 추세이다. 이에 따라 정부 부처의 공공직업 훈련 기관에서 한국어 교육과 한국문화의 이해에 대한 교육을 진행한다. 이러한 노력에도 불구하고 외국인 근로자가 한국문화를 처음 접했을 때 많은 어려움을 겪고 있다고 볼 수 있다. 본 논문에서는 외국인 근로자들의 문화적 어려움을 해결하기 위해 개발한 공익적 앱을 소개한다. 한국에서 대중적인 사물 중에서 외국인 근로자들이 다소 생소하게 느끼는 사물에 대하여 연구를 진행했다. 스마트폰에서 실시간으로 사물을 인식하고 그 사용 용도를 여러 언어로 번역하여 제공하는 앱을 개발하였다. 언어 선정은 한국 입국 비중이 큰 다섯 국가의 언어를 대상으로 했다. 학습 이미지로 7,900장 수집하여 라벨링하였다. AI 신경망인 Darknet을 이용하여 YOLOv3-tiny 학습 모델로 총 20,000회의 반복 학습을 시켰다. 개발된 앱은 평균 91.3%의 정확도로 스마트폰에서 실시간 사물을 감지할 수 있다. 감지된 사물의 설명을 버튼 하나로 5개국 언어로 번역하여 제공하는 시스템이다. 이 앱을 이용하면 외국인 근로자들의 일상생활의 문화적, 언어적 차이를 해소할 수 있을 것으로 본다.

### 감사의 글

본 논문은 2020년도 한국전자통신연구원의 연구비를 지원 받음.



**Bum-Soo Kim** is Students in School of Computer Engineering & Applied Mathematics at Hankyong National University.

*E-mail address:* iimkbsu@hknu.ac.kr



**Young-Hyen Kim** is Students in School of Computer Engineering & Applied Mathematics at Hankyong National University.

*E-mail address:* credita@hknu.ac.kr



**Byung-Tae Chun** Master's degree from Graduate School of Soongsil University and doctor's degree from Graduate School of Korea University

Worked as researcher for KIST (Korea Institute of Science and Technology)/System Engineering Research Institute (SERI) from 1999 to 1996 and worked as executive researcher for ETRI(Electronics and Telecommunications Research Institute) from 1996 to 2004. From 2004 to current: Professor of computer engineering Dept. of Hankyong National University. From 2013 to current: Vice president of TTA/TC4/PG413. From 2010 to 2018 : Chairman of WG SC6 subcommittee of Intelligent Robot Standard Forum (KOROS) Committee

*E-mail address:* chunbt@hknu.ac.kr