

## 다중 카메라 환경에서의 안면인식 기반의

### 영유아 활동 사진 자동 생성 시스템

이정석, 이규호, 김건희, 최창훈, 박경로, 손호준, 유홍석(교신저자)

경운대학교 항공소프트웨어공학과

e-mail: {vddv12, lkhkh0925, kkkk159258, rororo4952, hun7898, rhdqjtdkd}@naver.com, hsyoo@ikw.ac.kr

## A system for automatically generating activity photos of infants based on facial recognition in a multi-camera environment

Jung-seok Lee, Kyu-ho Lee, Kun-hee Kim, Chang-hun Choi, Kyoung-ro Park,

Ho-joun Son, Hongseok Yoo(Corresponding Author)

School of Software, Kyungwoon University

### ● 요약 ●

본 논문에서는 다중 카메라환경에서의 안면인식 기반 영유아 활동 사진 자동 생성 시스템을 개발했다. 개발한 시스템은 어린이집에서 알람장 작성을 위한 촬영하는 동안 보육에 부주의하여 안전사고가 발생하는 것을 방지 할 수 있다. 시스템은 이동식 수집기와 분류 서버로 나뉘어 작동하게 된다. 이동식 수집기는 Raspberry Pi를 이용하였고 초당 1장 내외의 사진을 촬영하여 SAMBA를 사용 공유폴더에 저장한다. 분류 서버에서는 YOLOv5를 사용해 안면을 인식해 분류한다. OpenCV와 TensorFlow-Keras를 통해 분류된 사진에서의 표정을 파악하여 부모에게 전송할 웃는사진만을 분류하여 남겨둔다. 이외의 사진은 /dev/null로 이동하여 삭제된다.

**키워드:** 율로(Yolov5), 라즈베리파이 (Raspberry Pi), 텐서플로우(Tensorflow)

## I. Introduction

현재 어린이집에서는 원생의 사진을 촬영하여 스마트 알람장을 작성한다. 이 서비스는 가정에서 어린이집으로 이어지는 보육환경의 연속성을 강화시켜 영유아의 건강한 발달을 위함이다. 제공하는 기능으로는 알람장, 공지사항, 포토앨범, 알람장 등이 있다.

아동교육 전문가 집단에서 이 서비스에 대해 많은 연구가 이뤄졌다. 이지선(2019), 류재형(2019), 유수정 외 1명(2022)<sup>1)</sup> 전문가들의 공통적인 문제점으로 영상 중심의 알람장 작성에 소요되는 보육교사의 과중한 업무 부담을 주장하였다. 보육교사들은 주장과 같은 과중한 업무로 인해 보육이라는 실질적인 목적에 집중하지 못하여 사고가 발생하기도 하였다. 2018년 물놀이장에서 현장 체험 학습을 진행한 한 어린이집에서 보육교사들이 알람장 작성을 위해 사진을 촬영하는 동안 한 아이가 물에 빠진 것을 인지 못 해 익사할 뻔한 미수사건이 있었다.

본 논문에서는 사진 촬영으로 인해 발생한 문제를 해결하기 위한

방안으로 본 시스템을 제안하였다.

먼저 카메라를 이용하여 촬영하게 된다. 촬영된 사진은 SAMBA를 이용한 공유폴더에 저장된다. 이동식 수집기는 보드에 OS를 올릴 수 있고 네트워크에 연결할 수 있어야 한다. CCTV나 바디캠을 이동식 수집기로 사용할 수 있다. 본 시스템에서는 Raspberry Pi Zero 2 W에 CSI캠을 연결하여 체스트 버클을 이용해 바디캠을 제작하였다.

다음 단계로 안면을 인식하여 사진을 분류한다. 객체를 감지하는 딥러닝 알고리즘으로는 YOLOv5를 사용하였고, YOLOv5-Weight File이라는 객체를 인식하는 특징을 저장한 파일에 인식시킬 대상의 안면 사진을 여러 장 찍어 전이 학습을 시켜두었다.

마지막으로 표정 인식단계는 학부모에게 아이들의 웃는 사진을 제공하기 위해 OpenCV라는 오픈 소스 라이브러리와 TensorFlow-Keras라는 딥 러닝의 구현을 도와주는 인터페이스를

1) 스마트 알람장을 통한 영아교사와 부모의 의사소통에 관한 연구(2019, 이지선)

어린이집 교사의 부모와의 의사소통을 위한 알람장 사용 실태 및 인식(2019, 류재형)

모바일 알람장 앱을 매개로 한 부모와의 의사소통에 대한 실험연구(2022, 유수정 외 1명) 의 내용을 저자의 용어대로 표현한것임

사용하였다.

## II. Preliminaries

### 2.1 Yolov5

YOLOv5은 실시간 객체 검출 알고리즘으로 모델을 불러와 OpenCV에서 사람을 검출하기 위해 YOLOv5의 기존에 학습된 객체 정보를 이용하였다.

### 2.2 OpenCV

OpenCV 라이브러리를 사용하여 이미지 파일을 읽어 들었다. 지정된 디렉토리에서 .jpg 확장자를 가진 모든 이미지 파일을 읽어 들였고, 읽어 들인 이미지는 YOLOv5 모델을 사용하여 이미지에서 사람을 검출하고, 나머지 경우의 이미지 파일은 /dev/null 디렉토리로 이동시켜 삭제하였다.

### 2.3 SAMBA

네트워크 파일 및 프린터 공유를 위한 소프트웨어로 서로다른 OS의 상호 작용이 가능하다. 리즈비인과 윈도우 간의 이미지 공유를 위해 사용하였다.

### 2.4 Tensorflow

TensorFlow 라이브러리를 사용하여 특정 디렉토리에 있는 이미지를 전처리 후, 기존 모델을 불러와 사람을 구분하기 위한 새로운 분류기를 추가한 뒤 전이학습을 실행한다. 해당 학습 모델을 저장한다. 사람을 분류 할 때 학습한 모델을 불러오고, 학습된 정보를 토대로 이미지를 분석하여 가장 높은 값을 가진 클래스의 인덱스를 찾아 출력한다.

## III. Design and Development

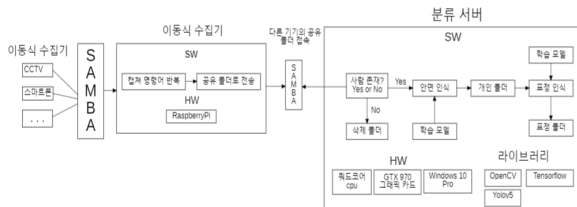


Fig. 1. Diagram Chart

Fig. 1의 이동식 수집기를 보면 자동 촬영 기능이 있다. 이것은 0.5-1FPS의 속도로 촬영 후 저장이 된다. FPS를 높게 설정한다면 중복되는 사진들로 인해 필요가 없는 많은 양의 사진이 서버에 유입되고 이는 서버와 통신과 용량에 부담을 주기 때문에 낮은 FPS를 설정하였다. HW는 Raspberry Pi Zero 2 W를 사용하였다. 무선통신과 OS를 갖춰 분류 서버와 통신을 기능을 하는 SAMBA의 사용이

가능하며 바디캠으로 사용하기 가벼운 무게를 가졌기 때문이다.

Fig. 1의 분류 서버의 SW를 보면 안면 인식과 표정 인식 2가지로 나뉜다. 안면인식은 YOLOv5라는 객체 인식 알고리즘을 사용하였는데 앞서 설명한 Weight File에 대상의 안면을 학습시켜 정보를 저장하면 YOLOv5에서는 해당 정보와 대조해 알맞은 안면을 인식하게 된다. 인식된 안면은 대상별로 디렉토리에 나뉘어 저장된다. 표정 인식은 OpenCV라는 라이브러리에서 객체 인식 라이브러리를 사용한다. TensorFlow-Keras에서는 딥 러닝 기반으로 신경망을 구성하게 되는데 여기서 표정 인식에 대한 학습과 인식이 이뤄지게 된다. 두 기능을 하는 분류 서버의 HW에는 GTX 970 그래픽 카드와 Quad core CPU를 사용하였고 OS는 Windows 10 Pro를 사용하였다. 사용된 라이브러리는 영상처리에 특화된 OpenC -V와 YOLOv5, TensorFlowKeras를 이용하였다.

사진의 흐름은 이동식 수집기에서 자동 촬영을 통해 사진을 모으고 이것을 공유폴더에 저장한다. 분류 서버에서는 공유폴더의 사진을 대상으로 안면인식을 실시해 각 디렉토리에 저장되고 분류에 실패한 사진은 /dev/null에 연결되어 삭제된다. 이어서 표정 인식을 하고 행복한 표정 이외의 사진도 마찬가지로 /dev/null로 이동하게 된다. 추후 사용자는 디렉터리별 분류된 사진을 사용하여 업무를 진행할 수 있다.

## IV. EXPERIMENT

본 시스템을 구성할 때 사용된 HW와 SW는 Fig. 1의 사양과 동일하다. 검증에는 본교의 빈 강의실을 사용하였다. 인식 대상은 4명으로 학습된 안면 2명과 학습되지 않은 안면 2명으로 설정하여 인식되지 않는 안면은 삭제가 되는지 확인하였다.



Fig. 2. Product Picture

Fig. 2의 사진(좌)는 바디캠을 착용한 모습이고 사진(우)는 바디캠이다. 약 40초 동안 바디캠을 사용하여 사진을 수집하고 스마트폰을 이용하여 개인사진을 4장 촬영하여 공유 폴더에 저장하였다.

## REFERENCES

- [1] YOLOv5 : <https://github.com/ultralytics/yolov5/wiki>
- [2] OpenCV : <https://github.com/opencv/opencv/wiki>
- [3] SAMBA : [https://wiki.samba.org/index.php/Main\\_Page](https://wiki.samba.org/index.php/Main_Page)
- [4] TensorFlow : <https://www.tensorflow.org/guide/keras?hl=en>

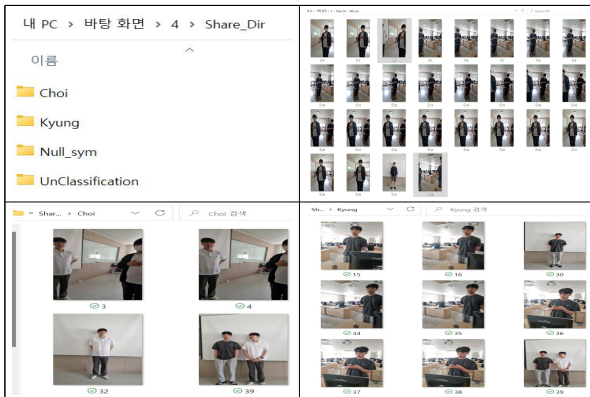


Fig. 3. Classification

Fig. 3의 첫 사진은 미리 만든 디렉터리로 찍은 사진은 UnClassification에 저장된다. 2번째 사진은 분류 전의 사진들이고 밑의 사진들은 분류 후의 사진이다. 같이 찍은 사진 학습시킨 두 안면이 동시에 인식되었을 경우 정상적으로 작동하는지 확인하기 위해 사용했으며 양쪽에 존재하는 것을 알 수 있다.

인식률은 사전 검증에선 98%였으며, 본 검증에서는 사진의 질럼으로 인해 95%까지 떨어졌다.

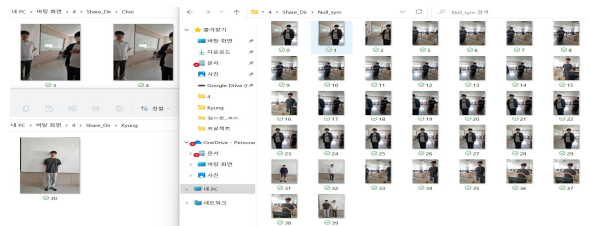


Fig. 4. Result

Fig 4는 최종적으로 분류된 사진이다. 왼쪽 위와 아래의 디렉터리인 Choi와 Kyung에 남은 사진은 표정 인식을 통해 3, 4번과 30번 사진으로 이외의 사진들은 모두 /dev/null의 링크 디렉터리인 Null\_Sym으로 이동하여 제거되었다.

## V. Conclusions

본 논문에서는 제안한 시스템은 다중 카메라 환경에서 사진을 수집하고 Weight File에 의해 분류된다. 이러한 특징을 활용하여 놀이공원을 이용하는 특정 고객들을 대상으로 고객의 사진 정보를 받아 활동사진을 CCTV를 통해 활동 사진을 수집해 제공한다. 또다른 활용 방안으로는 경찰에서 CCTV나 순찰(차량 또는 인원)을 통해 사진 및 영상을 수집하여 범죄자의 차량 번호나 안면을 식별하고 관제실 및 순찰인원에게 알람을 보내어 빠른 대처를 할 수 있다.

본 시스템은 추후에 사진이 아닌 영상으로 통신을 하고 서버에서 인식된 장면을 캡처하여 사진을 생성하는 기능으로 발전시킬 것이다.