

# Intelligent missing persons index system Implementation based on the OpenCV image processing and TensorFlow Deep-running Image Processing

Yeong-Tae Baek\*, Se-Hoon Lee\*\*, Ji-Seong Kim\*\*\*

## Abstract

In this paper, we present a solution to the problems caused by using only text - based information as an index element when a commercialized missing person indexing system indexes missing persons registered in the database. The existing system could not be used for the missing persons inquiry because it could not formalize the image of the missing person registered together when registering the missing person. To solve these problems, we propose a method to extract the similarity of images by using OpenCV image processing and TensorFlow deep - running image processing, and to process images of missing persons to process them into meaningful information. In order to verify the indexing method used in this paper, we constructed a Web server that operates to provide the information that is most likely to be needed to users first, using the image provided in the non - regular environment of the same subject as the search element.

▶ Keyword : Face Recognition, Area Detection, Controled Indexing

## I. Introduction

기존 상용화된 미아 등록 시스템은 텍스트 정보만을 색인요소로 사용하기 때문에 실종자의 이미지는 별도로 저장하여 출력만을 위해 사용되는 방식을 주로 사용하고 있다. 이러한 방법은 지역이나 나이와 같은 정형화 되어있는 텍스트 데이터만 실종자 검색에 사용할 수 있어 정형화 되지 않는 이미지 데이터는 사용자가 모두 확인해야 하는 방법 외에는 존재하지 않는다. 또한 실종 대상에게 정형적인 데이터수집이 어려운 영유아 또는 지적 장애 실종자 등에 대해서는 오직 이미지 데이터만이 주 색인 요소가 되는데, 이러한 방법은 실종자를 찾고자 하는 사용자에게는 많은 시간적 자원이 필요하게 되며 실종자를 찾을 수 있는 골든타임을 놓칠 가능성이 높다. 본 논문에서는 이러한 문제점을 해결하기 위해 기존 등록된 실종자의 이미지와 새로 등록되는 실종자의 이미지 사이의 유사도를 추출하여 정

형화된 데이터를 색인요소로 사용함으로써 유사도의 우선순위를 매겨 사용자가 필요할 가능성이 높은 정보를 우선적으로 제공할 수 있는 방법을 제안한다.

## II. Preliminaries

### 1. Similar systems

실종아동/장애인 검색 검색조건을 입력해 주세요.

이름입력	<input type="text"/>	성별구분	<input checked="" type="radio"/> 전체 <input type="radio"/> 남자 <input type="radio"/> 여자
대상구분	<input checked="" type="radio"/> 전체 <input type="radio"/> 실종아동 <input type="radio"/> 실종장애인	발생년도	<input type="text"/>
상종지역	<input type="text"/>		
신체특징키워드	<input type="text"/>		

Fig. 1. Existing system

• First Author: Yeong-Tae Baek, Corresponding Author: Se-Hoon Lee

\* Yeong-Tae Baek(hannae@kimpo.ac.kr), Dept. of Multimedia, Kimpo University

\*\* Se-Hoon Lee(seihoon@inhac.ac.kr), Dept. of Computer Systems & Engineering, Inha Technical College

\*\*\* Ji-Seong Kim(gogovlzk@gmail.com), Dept. of Computer Systems & Engineering, Inha Technical College

• Received: 2017. 01. 03, Revised: 2017. 01. 16, Accepted: 2017. 01. 20.

• This work was supported by KIMPO University's Research Fund.

• This paper is a revised and expanded version of a paper entitled 'Intelligent missing persons index system based on the OpenCV and TensorFlow' presented at the 2016 summer Conference of The Korea Society of Computer and Information

Fig.1은 ‘보건복지부 위탁 실종아동 전문기관 초록우산’에서 제공하는 실종자 검색 시스템이다. Application을 포함한 대부분의 실종자 색인 시스템이 위와 같이 운영되고 있으며, 실종자에 대하여 텍스트정보를 바탕으로 검색을 수행한다. 이 경우 실종자에 대한 정보가 부정확 하거나 실종자 데이터베이스에 잘못된 정보가 입력되어 있을 때 데이터베이스 안에 대상이 존재 하더라도 검색을 수행하지 않으며, 찾고자하는 대상의 정보 요소가 적을수록 사용자가 직접 구분해야할 정보가 많아지기 때문에 사용자들은 실종자 데이터베이스에 있는 정보들을 순차 열람하지 않는 이상 실종자를 찾을 수 없다는 문제가 발생한다 [1]. “미야의 발생예방 및 가족상봉 지원에 관한 법률안”[2]과 실종아동 전문기관에서는 사건이 발생했을 시점을 기준으로 48시간까지를 골든타임이라 칭하며 이를 벗어난 경우 “장기미 아사건”으로 분류한다. 이 때문에 골든타임을 지키기 위해 사건발생시점 기준으로 우선 노출시켜야 될 것이며, 골든타임이 벗어난 장시 실종자의 경우 현재는 등록 순으로 정보가 나열되어 사용자가 모두 열람해야 하는 문제가 있는데, 이러한 문제를 색인 시스템의 개선을 통한 해결이 필요하다.

2. Related works

2.1 Face recognition technology

얼굴인식 기술은 지문, 음성, 서명 등의 생체인식 기술 중 가장 활발한 연구가 진행되어 왔으며 보안 및 대상의 색인을 위해 가장 많이 사용되어지고 있다. 이 기술을 이용하여 두 개 이상의 영상 비교를 통해 신원 인증 및 확인의 문제를 해결할 수 있다.

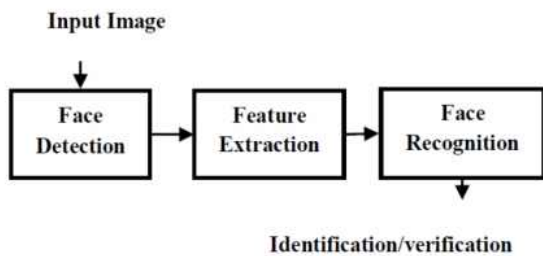


Fig. 2. Face Recognition Process[3]

얼굴인식 과정은 Fig.2와 같이 얼굴 검출, 특징 추출, 얼굴 인식의 과정으로 크게 3 단계의 프로세스로 이루어진다.

Table 1. System Environment

얼굴인식 기술	설명
기하학적 특징 기반	양안과 코의 특징 및 그 사이의 기하학적인 관계를 분석 그래프 매칭 알고리즘 등을 활용
얼굴 전체 영역 기반	얼굴 전체 영역을 분석 PCA(Principal Component Analysis) 과정이 필수
하이브리드 기반	지역적 특징과 전체적인 특징을 조합한 정보를 활용 주로 눈, 코, 입을 주요 특징으로 활용

얼굴인식 기술은 다양한 방향으로 지속적으로 발전되어 왔으며, 이는 Table 1 과 같이 크게 세 가지 방식으로 분류할 수 있다. 본 논문에서는 얼굴 전체 영역 기반 비교 방식으로 YCbCr 색공간과 히스토그램 비교연산을 통해 두 개 이상의 영상을 비교하여 유사도를 추출하는 방식을 채택하여 이미지 정보 까지 실종자의 색인 정보로서 사용할 수 있도록 하였다.

2.2 Machine Running - OpenCV Cascade



Fig. 3. System Architecture

각기 다른 피사체를 동일한 수준에서 비교하기 위해 피사체의 특징만을 추출하는 일반화 가공과정이 필요하다. 본 논문에서 피사체의 특징은 ‘얼굴’ 요소이며, 일반화 과정의 첫 번째 작업으로 이 ‘얼굴’ 요소의 영역검출이 필요한데, 이를 위해 Cascade 물체학습 알고리즘을 사용하여 인체의 ‘얼굴’ 요소를 학습시킨 뒤 요소의 영역을 검출한다. 본 논문에서는 Cascade 물체학습의 실험을 위해 Fig.3의 피사체인 ‘인형’ 물체를 학습하여 인형을 감지하는 검증을 하였다. 피사체를 판별하기 위해 Cascade 물체학습을 사용할 때 피사체가 표현된 Positive Image와 피사체가 표현되지 않은 Negative Image를 필요로 하며[4], 물체인식을 위한 최소한의 샘플로 Positive Image는 약 600 Sample, Negative Image는 Positive Image의 약 12 배의 Sample을 필요로 한다.

2.3 Machine Running - TensorFlow

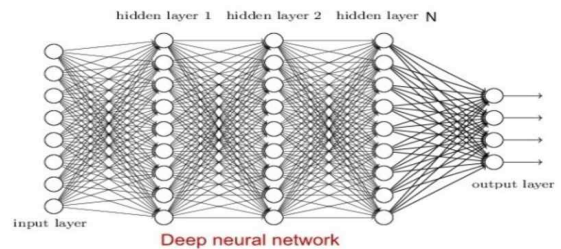


Fig. 4. Deep neural network

Fig.4는 input layer와 output layer 사이에서 다수의 hidden layer로 이루어진 인공신경망인 심층 신경망(Deep Neural Network, DNN)[5,6]을 표현하고 있다. 일반적인 인공 신경망은 비선형 관계들을 모델링 할 수 있는데, 비선형 관계의 모델링이 가능함을 복잡한 벡터 가설 식을 세워야하는 데이터 분포에서 단순 선형회귀만으로 불가능했던 특징추출이 가능하게 된다. 여기서 심층 신경망은 영상의 단순 특징 학습에서 더 나아가 영상의 기본요소들을 layer사이에서 계층적으로 구성하

기 때문에 일반적인 인공신경망과 비교하여 더 적은 수의 노드 들만으로도 복잡한 벡터 가설 식을 모델링할 수 있게 해준다. 본 논문에서는 TensorFlow를 이용해 심층신경망을 구성하였다.

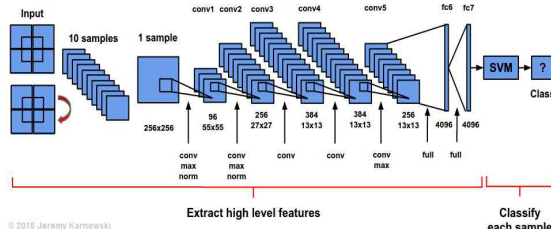


Fig. 5. Convolutional Neural Network Process

본 논문에서의 영상처리를 위해 TensorFlow deep running 도구에서 합성곱 신경망(Convolutional Neural Network, CNN) 알고리즘을 사용하였다. 이는 이미지처리에 특화된 전통적인 알고리즘으로서 성공적인 사례에 대해 문서화가 잘 되어 있기 때문에 본 실험에서 우선적으로 사용하게 되었다. 이 알고리즘은 최소한의 전처리를 사용하며, 병렬처리를 이용하여 상대적으로 빠른 처리속도의 프로세싱 능력을 가진 다 계층의 알고리즘이다. Convolution을 통해 각 노드를 추상화하고 노드에 따른 가중치 연산과 Convolution을 통해 다음 계층의 노드로 연결한다. 이 과정을 반복하여 다 계층의 구조로 생성함으로써 소량의 매개변수만으로도 층이 깊어질수록 고도화된 특징들을 검출할 수 있다는 특징을 가지고 있다.

### III. Image Recognition

#### 1. OpenCV

##### 1.1 Image Generalize

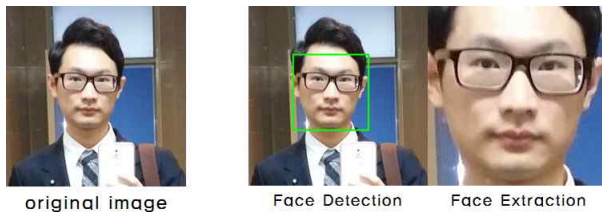


Fig. 6. Face Extraction

Fig.6은 OpenCV기반에서 Cascade 학습을 통해 사람의 얼굴을 학습시킨 후 반환된 데이터 셋을 이용해 비트맵 영상에서 피사체의 얼굴을 검출해내는 것을 표현하고 있다. 이 때 보다 높은 얼굴 인식 정밀도를 위해 검출해낸 얼굴의 영역에서 배경을 배제하여 일반화 할 필요성이 있다. 본 논문에서는 배경을 배제한 얼굴만을 추출해내기 위해 사람의 피부색을 기준으로 필요한 영역을 마스킹 하는 방법을 채택하였다. 사람의 피부색을 기준으로 마스킹하려 할 때 RGB 색 공간에서 마스킹 할 경

우 추출한 영상에서 가변 하는 광원에 따라 바뀌는 피부의 색 값에 대한 대처가 어려워진다.

$$Y = 0.299 \times R + 0.587 \times G + 0.114 \times B + 0$$

$$C_b = -0.169 \times R - 0.331 \times G + 0.499 \times B + 128$$

$$C_r = 0.499 \times R - 0.418 \times G - 0.0813 \times B + 128$$

$$R = \text{clamp}(Y + 1.402 \times (C_r - 128))$$

$$G = \text{clamp}(Y - 0.344 \times (C_b - 128) - 0.714 \times (C_r - 128))$$

$$B = \text{clamp}(Y + 1.772 \times (C_b - 128))$$

Fig. 7. RGB-YCrCb Conversion Formula[7]

따라서 위 문제는 Fig.7의 공식을 이용해 RGB 색 공간에서 YCbCr 색 공간으로 인코딩하여 문제를 해결한다.

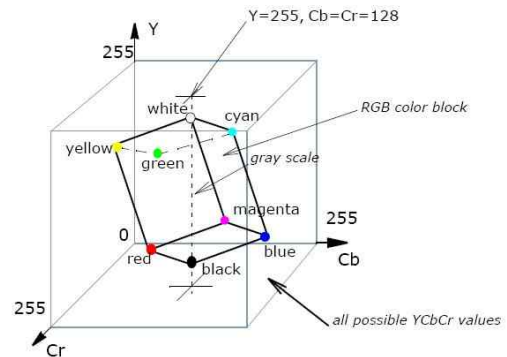


Fig. 8. Relationship between RGB and YCbCr[8]

인텔 컬러모델 문서[8]에 따르면 RGB색 공간은 빛의 3원색을 정보로서 사용하여 색에 따른 상호간 연관성이 심하며, YCbCr 색 공간은 Y는 휘도, Cb와 Cr은 상호간 색차 성분을 정보로서 사용하여 휘도 값을 배제할 수 있게 된다. 따라서 피부 색차 범위를 구한 뒤 휘도 값을 배제한다면 광원에 관계없이 피부영역을 마스킹할 수 있게 된다.

Table 2. YCbCr Skin Color Ranges

Color Unit	Maximum	Minimum
Cb	73	158
Cr	128	170

이에 따라 피부색의 범위를 도출하기 위해 RGB색 공간에서 피부가 가질 수 있는 색 범위의 최대, 최솟값을 구한 뒤 Fig.7의 공식을 이용하여 YCbCr값으로 인코딩 한다. Table 2은 본 논문에서 인코딩한 YCbCr색 범위를 나타낸다.

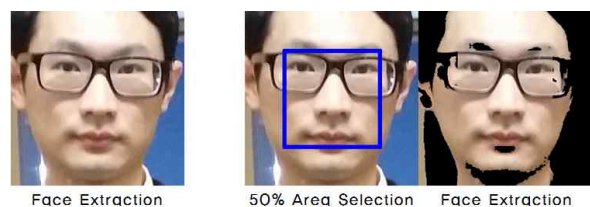


Fig. 9. Face Generalize

Fig.9는 Fig.6에서 검출된 얼굴을 피부색만을 추출하여 일반화하는 과정을 표현한다. 일반화를 위해 Fig.6에서 잘라내어진 사각형 영역의 중심으로부터 가장자리까지 50% 영역은 보존하며 이외의 영역을 Table 2의 피부 색차 값으로 마스킹 하여 일반화를 마친다.

### 1.2 Face Recognition

일반화를 마친 얼굴 영상을 비교하기 위해 본 논문에서는 히스토그램을 이용한 비교연산을 사용한다.

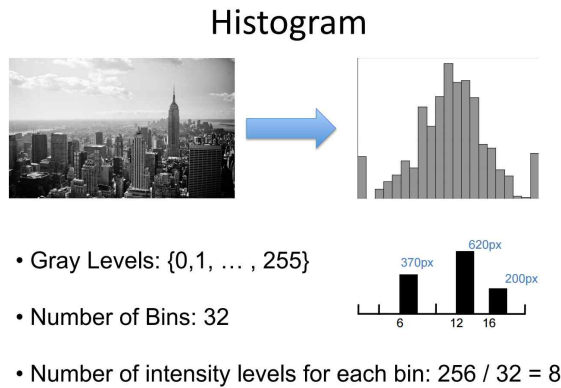


Fig. 10. Histogram Graph[9]

히스토그램 그래프란 영상의 픽셀 값의 분포를 그래프로서 표현한 것이다[10]. 그래프의 세로축은 픽셀 값의 빈도를 나타내며, 가로 축은 영상의 픽셀 값을 나타낸다. 그러므로 픽셀의 분포를 통해 영상을 단순화시켜 두 영상의 특징만을 빠르게 비교할 수 있게 된다. 따라서 일반화된 히스토그램 값을 비교하여 두 얼굴의 유사도를 판별할 수 있게 되는 것이다.

### 1.3 Recognition Test

Table 3. Histogram Matching Result

Count	Same Figure	Other Figure
1	273	865
2	230	829
3	282	1020
4	290	850
5	422	1133
6	363	823
7	438	984
8	402	956
9	301	931
10	269	852

Table 3의 결과에 따르면 동일 인물은 그래프 비교 결과가 약 500 이하의 결과를 나타내며, 서로 다른 인물은 약 800 이상의 결과를 나타낸다. 따라서 위 실험을 통해 히스토그램 그래프 비교를 이용하여 동일 인물과 다른 인물을 판별할 수 있음을 알 수 있다.

## 2. TensorFlow

### 2.1 Object Recognition

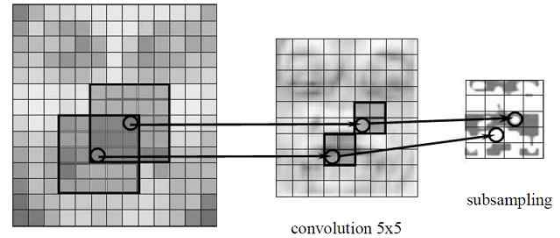
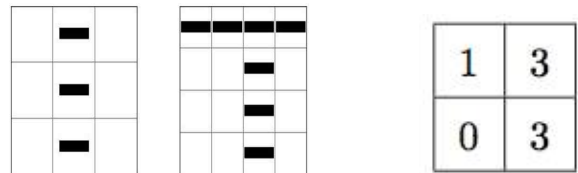


Fig. 11. pooling layer[11]

Fig.11은 이미지의 크기를 계층별로 잘라내는 과정이다. 일반적으로 CNN은 이미지에 적용되는데 이미지의 특성상 픽셀의 개수가 너무 많아 계산과 파라미터의 양을 줄기위해 표현의 공간적 크기를 줄이는 sub-sampling 과정을 pooling이라고 부른다. 이 pooling 과정의 횟수의 따라 계층 수가 나뉘게 되는데 층의 숫자가 많아질수록 분류 성능은 좋아지지만 그만큼 처리 속도가 늘어나게 되어 합리적인 계층 수를 찾아야한다.



$$s(t) = (x * w)(t) = \int x(a)w(t-a)da.$$

Fig. 12. Intuition behind convolution[12]

Fig.12는 convolution을 통해 노드에서 특징을 뽑아내는 과정을 나타낸다. 이 operation은 주어진 데이터 x에 filter w를 사용해 데이터를 처리할 때 사용되며 위의 filter는 수직의 edge를 검출하는 필터를 의미한다. 이러한 필터들을 이용하여 각 계층마다 설정된 필터로 노드에서 유의미한 특징을 검출하게 되는 것이다. pooling과 layer와 convolution layer를 반복해서 사용함으로써 각 계층 각 노드마다 유의미한 특징들을 저장함으로써 데이터 셋이 만들어지게 된다. 본 논문에서는 사람뿐만이 아닌 반려동물 또한 실종된 가족으로 간주하여 같은 시스템에서 색인 시스템을 제공하기 위해 앞서 소개한 심층신경망과 합성공 신경망을 이용하여 반려동물을 판별한다.



Fig. 13. Animal Recognition using CNN algorithm of TensorFlow

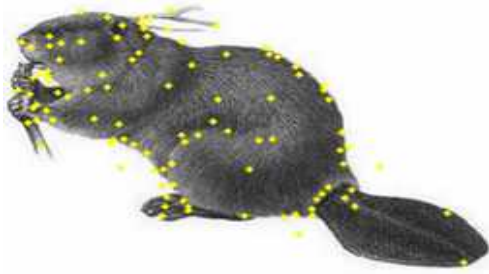


Fig. 14. Feature point detection using OpenCV

Fig.13은 Google에서 제공하는 TensorFlow의 학습된 객체 데이터 셋을 대상으로 실험한 내용이다. 추가적인 실험으로 Fig.3에서 표현된 펭귄을 형상한 봉제인형을 위 데이터 셋에서 추출한 가설 식에 대입하였을 때 ‘동물’인 펭귄이 예측된 것이 아닌, ‘봉제인형’ 과 ‘테디베어’가 결과로 예측되었다. 이는 입력된 피사체의 전체 형상만을 비교한 것이 아닌 layer 구성에서 최저 층의 요소부터 비교하여 피사체의 구성조직부터 찾아낸 것이라고 할 수 있다. 그러므로 결과적으로 반려동물의 털 상태에 따라 전혀 다른 종으로 예측 될 수 있기 때문에 Fig.14에서 표현하고 있는 OpenCV의 Feature Matching을 2차 예측 수단으로 사용하여 결과를 보정할 필요가 있다.

## IV. System Design and Implementation

### 1. System OverView

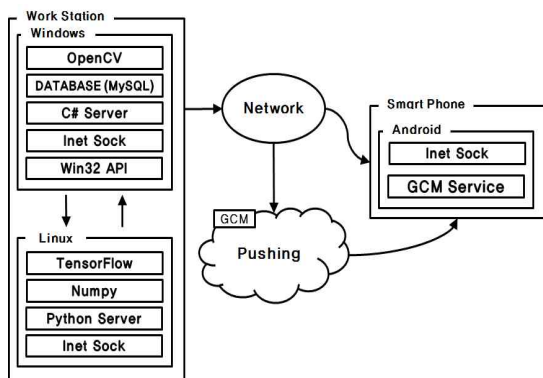


Fig. 15. System Composition

Fig.15는 본 논문에서 제안하는 전체 시스템 구성을 나타내고 있다. 워크스테이션 가상화를 통해 Windows와 Linux 두 개의 OS가 동작하며, 상호간 TCP통신을 통해 정보를 주고받는다. Windows 단에서 안드로이드와 직접적으로 TCP통신하여 페이지 요청에 따른 결과를 반환하는 방식으로 동작하며, 이벤트 발생 시 푸시 메시징 서비스를 통해 사용자에게 찾고 있을

확률이 높은 정보를 제공한다.

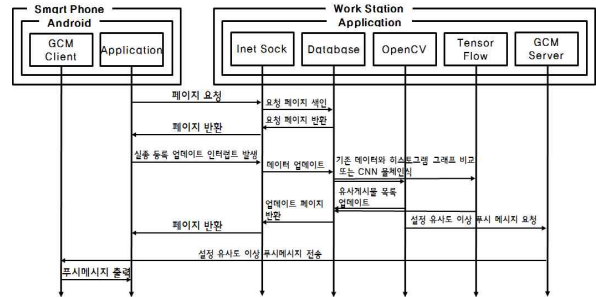


Fig. 16. Pushing Scenario Sequence Diagram

Fig.16은 사용자가 게시 글 등록 인터럽트 발생 후 푸시메시지 전송까지의 시나리오를 표현하고 있다. 첫 번째로 어플리케이션을 실행하게 되면 먼저 게시판의 종류를 출력한다. 이 때 실종자 찾기 버튼 인터럽트가 발생한다면 서버에 페이지를 요청한다. 페이지를 요청받은 서버는 DB에서 목록을 시간 순으로 색인하며 결과를 반환한다. 두 번째로 반환된 목록의 한 항목을 선택하여 인터럽트가 발생했을 때 페이지 요청과 마찬가지로 항목에 대한 상세정보를 반환받는다. 실종자 등록을 위해 실종자 등록 버튼을 누르고 상세 정보를 입력한 뒤 업로드 인터럽트를 발생시켰다면 등록할 이미지를 파일로 저장한 후 기존 저장되어 있는 데이터와 히스토그램 비교연산 또는 TensorFlow 동물인식을 통해 데이터를 정형화하며 설정한 유사도 이상의 결과 값이 반환될 경우 해당 클라이언트에게 푸시 메시지를 발송하며, 이후 최종적으로 상세정보를 데이터베이스에 기록한다.

### 2. Application Implementation



Fig. 17. Running Case1

Fig.17은 미야 발생 시 보호자가 App을 통해 실종자를 등록하는 과정이다. 보호자는 실종자에 대한 정보와 함께 실종자의 이미지를 삽입하게 된다.



Fig. 18. Running Case2

Fig.18은 보호자가 없는 미아를 찾았을 경우 App을 통해 보호자를 찾는 과정이다. Fig17과 동일한 정보를 입력하되 태그를 변경하여 실종자 등록과 별개로 등록하게 된다.



Fig. 19. Running Case3

Fig.19는 보호자를 찾고 있는 실종자에 대한 정보가 새로 입력 됐을 때 기존 실종자 리스트에서 유사한 이미지가 있는지 검색해 일정 유사도 이상의 실종자를 찾아 보호자에게 푸시메시지를 전송하는 기능을 나타낸다.

### V. Conclusions

본 논문에서 텍스트 기반의 통계색인을 바탕으로 실종자를 검색하는 방식에서 발생하는 기존 시스템의 문제점을 해결하고자 이미지 데이터를 정형화 시키는 시스템을 제안하였다. OpenCV와 TensorFlow를 이용하여 가공한 정형화된 데이터를 색인에 포함함으로써 사용자에게 요구되는 시간비용을 줄일 수 있어 장기 실종사건으로 발전될 수 있는 사건을 미연에 방지할 수 있는 기대효과를 지니고 있다. 본 논문에 사용된 얼굴 인식 기술은 얼굴 전체 영역 기반 히스토그램 비교 분석 방식

을 사용하였으나, 기하학적 특징 기반 방식을 함께 사용한 하이브리드 기반 방식으로 얼굴인식 시스템을 구성한다면 더 높은 인식률을 기대하여 좋은 결과를 보여줄 수 있을 것으로 기대된다.

### REFERENCES

- [1] Deock-Soon Son, "Establishment support system for missing children", Gyeonggi-do issue brief in the world, No. 2, April 2008.
- [2] Hee-Seon Kim, "Proposal for Prevention of Mia and Support for Family Reunion", Politics in Korea, No. 162549, August 2003.
- [3] won-seok Chae, "Face Detection Technology Trends, Electronics and Telecommunications Research Institute", 2013.02.15.
- [4] OpenCV User Guide, "Cascade Classifier Training", 2007.
- [5] Y. Bengio, A. Courville, and P. Vincent., "Representation Learning: A Review and New Perspectives," IEEE Trans. PAMI, special issue Learning Deep Architectures, 2013.
- [6] J. Schmidhuber, "Deep Learning in Neural Networks: An Overview" <http://arxiv.org/abs/1404.7828>, 2014
- [7] Logic Devices, "RGB to YCbCr Report" Logic Devices pp .2, Jan 2001.
- [8] Intel , "Color Models" Intel Developer Zone , <https://software.intel.com/en-us/node/503873>
- [9] Vietdungitb, "Online handwriting recognition using multi convolution neural networks".
- [10] Yeong-Tae Baek, Se-Hoon Lee, Ji-Seong Kim, "Gesture Recognition using binary processing and histogram graph with OpenCV", 2016.02.
- [11] Vietdungitb, "Online handwriting recognition using multi convolution neural networks".
- [12] XiangWu, Ran He, Zhenan Sun, "A Lightened CNN for Deep Face Representation", 2015

## Authors



Yeong-Tae Baek received the B.S. degree in Computer Science from Inha University, Korea, in 1989 and M.S. and Ph.D degrees in Computer Science from Inha University, Korea, in 1993 and 2002, respectively. He is currently a Professor in the Department of

Multimedia at Kimpo University. He is interested in multimedia contents, IoT platform, and mobile system.



Se-Hoon Lee received the B.S. degree in Computer Science from Inha University, Korea, in 1985 and M.S. and Ph.D degrees in Computer Science from Inha University, Korea, in 1987 and 1996, respectively.

From 1987 to 1990, he worked as a information analyst officer in the Computing Center, at Marine Corp. Dr. Lee joined the faculty of the Department of Computer Engineering at Inha Technical College, Incheon, Korea, in 1993. He is currently a Professor in the Department of Computer Systems & Engineering at Inha Technical College. He is interested in embedded software, IoT platform, and mobile system.



Ji-Sung Kim is currently studying in the Department of Computer & Engineering at Inha Technical College. He is interested in embedded software, IoT security, and NUI.