

범죄예방을 위한 CCTV 영상 기반의 실시간 안면인식 시스템

김현빈*¹, 최낙훈*², 강지수*³, 임소현*⁴, 김희열**

Real-time Face Recognition System in CCTV Video for Crime Prevention

Hyun-Bin Kim*¹, Nakhoon Choi*², Ji-Soo Kang*³, So-Hyun Lim*⁴, and Heeyoul Kim**

본 연구는 경기도의 경기도 지역협력연구센터 사업의 일환으로 수행하였음.
[GRRC경기2020-B04, 영상 및 네트워크 기반 지능정보 제조 서비스 연구]

요 약

현대 사회에서 성범죄율은 지속적으로 증가하고 있다. 특히 성범죄자의 경우 재범률이 높기 때문에 재발을 막기 위한 제도적 혹은 사회적 필요성이 점차 증가하고 있다. 이를 위해 본 논문에서 범죄예방을 위한 CCTV 영상 기반의 실시간 범죄자 안면 인식 시스템을 제안한다. 이는 인공지능 기반의 안면인식을 통해 성범죄 예방 및 성범죄자의 재범 방지를 위한 일련의 프로세스다. 하지만, 안면 인식 기술을 실질적으로 활용하기에는 문제점이 있다. 안면 인식 시스템은 영상에서 안면을 찾는 탐지 단계와 탐지된 안면과의 일치도를 계산하는 매칭 단계로 이루어지며 단일 작업에 비해 큰 연산 비용이 요구된다. 때문에 본 연구에서는 안면을 탐지할 때 벡터화 작업을 통해 처리 시간의 소모를 줄여 실시간 작업 시에 연산 비용을 줄이고, 이후 탐지된 안면 영상의 크기를 늘리는 AB-box 방식을 사용해 더 높은 인식 정확도를 보인다.

Abstract

In modern society, the rate of sexual offenses continues to increase. In particular, since the recidivism rate of sex offenders is high, the institutional or social need to prevent recurrence is gradually increasing. In this paper, we propose a real-time criminal face recognition system in CCTV video for crime prevention. This is a series of processes to prevent sexual offenses and re-offenses of sex offenders through artificial intelligence-based facial recognition. However, there is a problem in practically utilizing facial recognition technology. The facial recognition system consists of a detection step for finding a face in an image and a matching step for calculating the degree of matching with the detected face. It requires a high computational cost compared to a single operation. Therefore, this study shows higher recognition accuracy by using the AB-box method, which reduces the processing time through vectorization when detecting a face. Also, we reduce computational cost during real-time operation and increase the size of the detected facial image afterward.

Keywords

crime prevention, face recognition, face matching, real-time, CCTV

* 경기대학교 컴퓨터과학과 석사과정

- ORCID¹: <https://orcid.org/0000-0001-5295-7507>

- ORCID²: <https://orcid.org/0000-0001-6063-5030>

- ORCID³: <https://orcid.org/0000-0003-3433-0094>

- ORCID⁴: <https://orcid.org/0000-0001-7172-0871>

** 경기대학교 컴퓨터과학과 교수(교신저자)

- ORCID: <https://orcid.org/0000-0001-6341-580X>

• Received: Feb. 09, 2021, Revised: Aug. 06, 2021, Accepted: Aug. 09, 2021

• Corresponding Author: Heeyoul Kim

Dept. of AI Computer science, kyonggi University, 154-42, Gwanggyosan-ro, Yeongtong-gu, Suwon, 16227, Korea

Tel.: +82-31-249-9607, Email: heeyoul.kim@kgu.ac.kr

1. 서 론

현대 사회에서 성범죄에 대한 관심과 처벌이 강화되는 것과 다르게 범죄율은 지속해서 증가하고 있다[1]. 성범죄 형법 범죄율은 2009년 인구 10만 명당 32.8건에서 2019년 61.9건으로 약 2배가량 상승하였으며 범행의 지속성이 강하게 나타난다[2]. 특히 아동(만 13세 미만) 대상 성폭력 범죄는 2009년 1,180 건에서 2019년 1,376건으로 지난 10년 동안 16.6% 증가하였다. 아동 성범죄는 죄질이 나쁘고 재범률이 21.2%로 굉장히 높기 때문에 이를 사전에 차단하고 범죄자들을 추적할 필요성이 있다[3]. 성범죄를 예방하기 위한 제도적 장치로써 전자감독제도, 신상정보공개제도 등이 시행되고 있으나 성범죄자들을 완벽하게 감시하고 교정하는 것에 한계가 있다. 전자감독 제도는 재범위험성이 높은 특정 범죄자의 신체에 전자장치를 부착하여 이동경로를 파악하여 재범을 방지하는 제도이다. 하지만 전자감독관 전담인력이 충분하지 않은 상황에서 지속적인 감독이 어렵고 준수사항 위반자에게 신속한 제재를 가하기 어렵다는 한계가 있다[4]. 성범죄자 신상공개제도의 경우 성범죄자의 얼굴 사진, 신체정보 등의 신상정보를 인터넷과 모바일 앱을 통해 공개한다[5]. 이는 범죄자의 얼굴을 직접 기억해야 하기 때문에 범죄 예방에 실질적인 도움을 주기 어렵다. 따라서 실시간으로 범죄자를 추적함으로써 실질적으로 범죄 예방이 가능한 방법이 필요하다.

본 논문에서는 범죄예방을 위한 CCTV 영상 기반의 실시간 범죄자 안면인식 시스템을 제안한다. 본 연구의 목적은 CCTV 영상에서 실시간으로 범죄자를 탐지하고 즉시 보고 및 추적하여 범죄를 예방하는 것이다. 따라서 실시간적인 범죄 예방을 위하여 CCTV 영상에서 범죄자를 정확하게 인식하는 일련의 시스템이 필요하다. 제안하는 방법은 범죄예방을 위한 비디오 조사를 자동화하는 시스템이다. 우리는 영상에서 사람의 얼굴을 탐지하고 인식하기 위해 머신러닝 및 딥러닝 기술을 적용한다. 또한 데이터베이스에 저장된 범죄자 안면을 CCTV 영상 내에서 실시간으로 찾아내기 위한 방법을 제안한다.

본 연구에서 제안하는 실시간 범죄자 안면 인식 시스템은 범죄예방뿐 아니라, 사건 발생 이후

CCTV를 분석하는 데에도 유용하게 활용될 수 있다. 어린이집 CCTV 영상에서 성범죄자가 어린이집 근처에 있다는 것을 탐지한다면 어린이들을 성범죄자로부터 보호하는 것이 가능하다. 결과적으로 본 연구를 통해 어린이집 등의 교육기관에 설치된 CCTV에서 범죄자를 인식하고 관련 기관에 통보함으로써 범죄 수사 또는 예방에 기여하는 것을 최종 목표로 한다.

본 논문의 구성은 다음과 같다. 2장에서는 일반적인 안면 인식과 CCTV영상에서의 안면인식 연구 동향에 대해 기술한다. 3장에서는 제안하는 범죄예방을 위한 CCTV 영상 기반의 실시간 범죄자 안면인식 시스템에 대해 기술한다. 4장에서 제안하는 모델과 성능평가 결과에 대해 다루고 5장에서 결론을 맺는다.

II. 관련 연구

2.1 범죄자 인식을 위한 안면 인식 시스템

컴퓨터비전 분야에서 영상으로부터 유의미한 정보를 추출하는 머신러닝 및 딥러닝 기술이 발전함에 따라 영상에서 사람을 탐지하고 인식하는 다양한 연구가 진행되고 있다[6]. 비전 연구에서 안면인식은 얼굴을 탐지하고, 매칭하는 일련의 과정이다[7]. 안면인식은 영상이나 이미지 데이터에서 인식할 얼굴을 탐지하는 안면 탐지(Face detection)와 탐지된 얼굴을 데이터베이스의 얼굴과 비교하는 안면 매칭(Face matching)으로 구성된다. 안면 탐지의 경우 Haar, HOG(Histograms of Oriented Gradients)등과 같은 전통적인 컴퓨터 비전 알고리즘을 활용하는 방법과 딥러닝 네트워크를 이용한 방법이 있다[6]-[8]. 기존의 전통적인 컴퓨터 비전 알고리즘 기법은 연산 속도가 빠르다는 장점이 있으나 사진 속 얼굴의 위치와 방향 및 카메라의 각도에 따라 얼굴 영상을 얻는 것에 어려움이 있어 정확도가 떨어지는 한계가 있다[8]. 딥러닝 네트워크를 활용한 방법은 활발하게 연구가 진행되고 있으며 대규모 데이터와 대회도 구축되어 있다. 딥러닝 기반 안면 인식은 일반적인 객체 탐지와 원리가 유사하며 성능이 우수하고 기존의 전통적인 알고리즘에 비해 더 높

은 정확도를 보여주는 장점이 있다. 하지만 연산량이 많고 높은 컴퓨팅 파워가 필요한 단점이 있다[7].

안면 매칭은 영상에서 안면 탐지를 통해 탐지한 얼굴과 같다고 판단되는 얼굴이 데이터베이스에 있을 때, 해당 라벨을 출력하여 누구인지 판단한다. 작동 원리는 탐지한 안면이 어떤 인물인지 다대일 관계에서 연결시키는 식별과 일대일 대조하여 참 또는 거짓의 결과로 확인하는 검증으로 이루어져 있다. 식별은 탐지된 안면의 특징점(Facial landmark)을 추출하고 데이터베이스에서 유사한 특징점을 가진 데이터를 찾는 것이다. 이후 유사한 특징으로 추출된 데이터를 KNN(K-Nearest Neighbor) 또는 SVM(Support Vector Machine)등의 분류기를 사용해 특징들의 일치 여부를 확인함으로써 검증은 진행된다.

2.2 CCTV 안면인식 연구동향

안면 인식은 컴퓨터 비전 연구에서 가장 중요하고 오래된 주제이다. 오랫동안 연구된 주제인 만큼 그 성능이 매우 뛰어나며 이를 활용해 보안, 헬스케어, 마케팅과 같은 실생활의 다양한 분야에서 활용되고 있다[9]. 하지만 이러한 연구에도 불구하고, 여전히 CCTV 영상에 있어서는 시스템의 효율성, 또는 성능에서 몇 가지 문제점을 보인다.

CCTV 영상에서 범죄자의 안면 인식에는 몇 가지 도전과제가 있다. 첫 번째는 다양한 환경에서 촬영된 CCTV 영상에서 작은 객체를 탐지하는 것이다. 나이, 빛과 같은 다양한 조건 아래에서 촬영되는 CCTV 영상의 특성 때문에 탐지되는 얼굴 객체가 아주 작고 저화질이다. 이는 탐지된 얼굴로부터 얼굴의 특징인 안면 특징점을 추출하기 어렵도록 만드는 원인이다. 두 번째는 모자나 마스크, 선글라스 때문에 얼굴의 일부분이 가려지는 폐색(Occlusion)이다. 안면의 일부분이 가려지더라도 안면 탐지는 가능하다. 하지만 얼굴의 특징을 적절히 추출하지 못하여 안면 매칭이 어려워지는 한계가 있다.

이러한 CCTV 영상에서의 문제점을 해결하기 위해 폐색 영상에 대해 우수한 성능을 보이는 부분적 접근방법, 얼굴 영상을 픽셀 매트릭스로 표현하고

이를 특징 벡터로 변환하는 전체적 접근방법, 두 가지의 방법을 혼합해 사용하는 하이브리드 접근방법이 연구되어 왔다[9]. 각각의 방법 모두 기존의 전통적인 방식에서 딥러닝 모델로 발전되어 왔으며 본 논문에서는 딥러닝을 기반으로 CCTV 영상에서 얼굴을 탐지한 후 실시간으로 범죄자와의 일치 여부를 확인하는 하나의 시스템을 제안한다.

III. 범죄예방을 위한 CCTV 영상에서의 실시간 안면인식 시스템

3.1 범죄예방을 위한 안면인식 시스템

본 연구는 CCTV 영상에 대한 실시간 안면 인식 시스템을 구현한다. 이를 위해 인공신경망이 안면 데이터로부터 정확히 얼굴의 특징점을 추출하도록 학습시키고, CCTV 영상에 적용한다. 또한 CCTV의 영상에서 추출한 사람 안면의 특징이 성범죄자 데이터베이스에 존재하는 특정 인물임을 추정 가능해야 한다. 대중적인 인공신경망 기반 안면인식 시스템인 OpenFace에서 제시된 방법은 FaceNet의 Triplet loss를 활용하여 안면 인식 성능 향상을 도모한다 [10][11]. 안면 인식 시스템에서는 안면 탐지 과정이 선행되어 입력 영상 전체로부터 안면 부분을 추출하는 구조를 채택한다. OpenFace에서는 HOG 알고리즘을 기반으로 하는 dlib의 사전 훈련된 안면 탐지기(Pre-trained face detector)를 활용한다. 이는 최근 가장 높은 성능을 보이는 안면 탐지기에 비해 매우

낮은 탐지 성능을 보이기 때문에 본 연구에서 제안하는 시스템은 FaceNet의 안면 인식 시스템 구조를 유지하되 RetinaFace 안면 탐지기를 채택한다 [12]. 안면 인식의 경우 dlib의 ResNet 기반 안면 인식 모델을 활용한다. 이 시스템은 실시간 스트리밍 데이터가 입력되기에 앞서 대조군(범죄자) 안면 데이터 전체를 인코딩하는 과정이 선행된다.

그림 1은 본 논문에서 제안하는 시스템의 전체 구조를 나타낸다. 시스템은 RetinaFace 안면 탐지기를 통해 탐지된 얼굴 부분을 인코딩한 후 범죄자 얼굴 데이터베이스와의 대조를 통해 관련 기관에 통보하는 일련의 프로세스를 가진다.

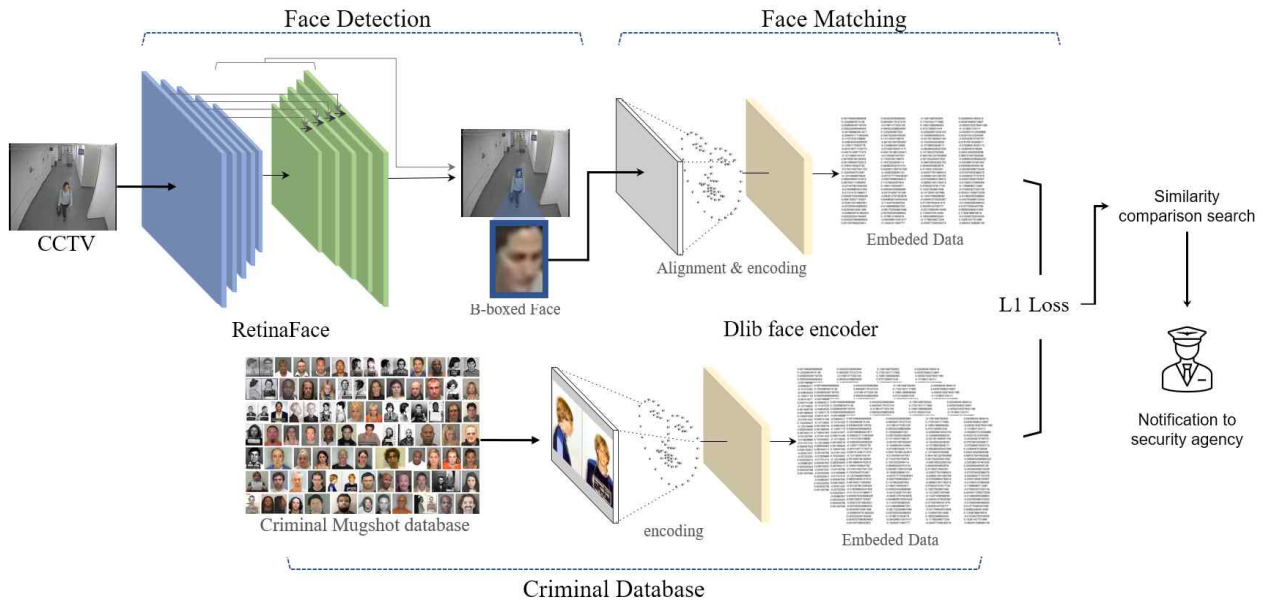


그림 1. 범죄예방을 위한 CCTV 영상에서의 실시간 범죄자 안면 인식 시스템 전체 프로세스
 Fig. 1. Process of real-time criminal face recognition system in CCTV video for crime prevention

3.2 실시간 안면 탐지를 위한 위치추정 방식

범죄자의 얼굴을 탐지하기 위한 모델은 RetinaFace의 안면 탐지기를 활용한다. RetinaFace가 제시하는 안면 특징 추출기로는 152층과 50층의 깊이를 가지는 ResNet 모델들과 Mobilenet0.25 모델이 있다[12]. 실시간 처리가 필요한 스트리밍 안면 인식 시스템의 특성상 ResNet-152 모델은 매우 높은 연산비용으로 인해 적합하지 않다. 따라서 이를 제외한 ResNet-50과 Mobilenet0.25 모델을 채택하여 사용한다. 또한, 더 빠른 안면 탐지를 위해 비교적 많은 반복 연산을 요구하는 Box Voting 연산[13]을 벡터화하여 추론 시간을 감소시켰다.

표 1은 CCTV 영상에서 위치추정방식에 따른 연산 시간 비교를 나타낸다. VGA, HD, FHD 해상도를 가지는 동영상 데이터에 대해 동일한 특징추출기를 활용하는 조건으로 위치추정방식에 따른 연산 시간을 측정하였다. 그 결과 위치추정연산 시간이 최소 26% 감소하는 것을 확인할 수 있다. 입력 영상의 크기가 커질수록 연산 방식 변화에 따른 연산 시간 감소율이 증가한다.

3.3 AB-box를 통한 안면 매칭

CCTV 동영상에서 안면을 효과적으로 인식하기 위해 영상을 프레임 별로 나눠 탐지를 진행한다. 하지만 이렇게 탐지된 안면 영상의 경우 크기가 너무 작아 안면을 인식하는데 어려움이 있다. 이를 해결하기 위해 본 연구에서는 패딩(Padding)을 추가한 더 큰 영역의 안면 영상인 AB-box(Additional Bounding Box)를 제안한다. 그림 2는 제안하는 AB-box의 개념도를 나타낸다. AB-box를 사용해 안면 탐지에서 추출된 영상보다 더 넓은 범위를 인코딩함으로써 안면 인식의 성능을 높이는데 기여한다.

안면 탐지 결과에 대해 dlib의 안면 인식 모델을 통해 다수의 좌표 정보로 구성된 데이터의 형태로 인코딩하고, 해당 영상과의 식별을 위해 앞서 언급한 대조군 안면 데이터 전체를 인코딩을 진행해 다수의 신원을 확보한다. 이렇게 만들어진 다수의 대조군 안면 인코딩 데이터들과 각 프레임에서 탐지된 안면의 인코딩 데이터를 대조해 데이터별 검증을 진행한다. 대조에는 인코딩 데이터 쌍의 벡터 차에 대해 L1 오차를 적용하여 유사도를 측정하고, 이렇게 대조한 얼굴 중에서 유사도가 가장 높은 것을 선택하여 탐지된 안면의 신원을 식별한다.

표 1. CCTV 영상에서 위치추정방식에 따른 연산 시간 비교

Table 1. Comparison of operation time according to position estimation method in CCTV video

Method			Time			
Resolution	Backbone	Post processing	Total	Inference	Post	matching
VGA(640 x 360)	ResNet50	Original	0.0313 s (32.0 fps)	0.0074 s	0.0138 s	0.0003 s
VGA(640 x 360)	ResNet50	Vectorized	0.0281 s (35.6 fps)	0.0070 s	0.0102 s	0.0002 s
HD (1280 x 720)	ResNet50	Original	0.0797 s (12.5 fps)	0.0089 s	0.0507 s	0.0008 s
HD (1280 x 720)	ResNet50	Vectorized	0.0626 s (16.0 fps)	0.0091 s	0.0315 s	0.0008 s
FHD (1920 x 1080)	ResNet50	Original	0.1690 s (5.9 fps)	0.0094 s	0.1177 s	0.0015 s
FHD (1920 x 1080)	ResNet50	Vectorized	0.1236 s (8.1 fps)	0.0089 s	0.0709 s	0.0015 s

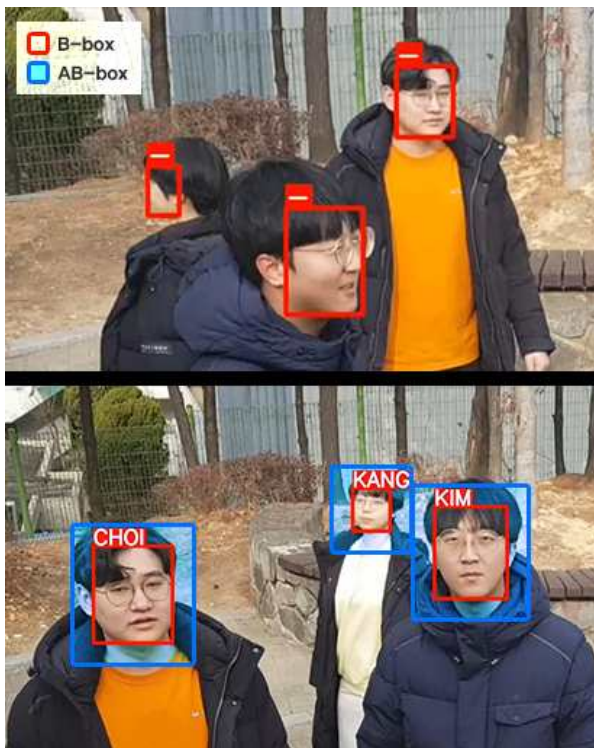


그림 2. AB-box의 개념도
Fig. 2. Concept image of AB-box

따라서 최종 안면 탐지 결과 영상의 인코딩 데이터와 범죄자 데이터베이스에 매칭되는 인코딩 데이터가 있을 경우, 해당 라벨을 출력함으로써 CCTV 영상에서 실시간으로 범죄자를 찾고 범죄를 예방하는 것이 가능하다.

IV. 실험 및 성능평가

제안하는 모델의 성능을 평가하기 위해 탐지 범위 추가에 따른 인식 정확도와 탐지 객체와 CCTV의 거리에 따른 인식 정확도 측정을 수행한다. 본

실험에서 사용된 하드웨어는 Ryzen 3700X, 32GB RAM, Geforce 1080Ti를 활용했으며, 소프트웨어는 CUDA 10.1, Pytorch 1.7.0, dlib 19.21.0을 사용하였다.

4.1 실험데이터 구성 및 구축

실험을 위해 정답이 될 범죄자의 얼굴 정보와 안면 인식을 수행할 얼굴을 포함한 CCTV 영상 데이터가 필요하다. 실험 영상 데이터는 1~4명의 사람을 대상으로 3/4/5m 거리에서 촬영되었다. 정답 라벨로 사용되는 얼굴 데이터는 저자 4명의 사진 데이터와 Helen 데이터[14]에서 추출한 396명의 사진 데이터를 사용하여 총 400명의 얼굴 데이터로 구성된다. 이는 대조할 범죄자의 데이터가 매칭을 수행할 객체의 수보다 많을 경우의 인식정확도를 고려하기 위함이다.

제안하는 안면 인식 시스템의 주요 성능으로는 처리속도, 탐지 성능, 인식 성능이 있다. 처리속도는 스트리밍 영상 데이터가 입력될 때 프레임당 처리비용을 계산한다. 프레임당 처리비용은 안면 탐지, 인식 및 시각화까지의 모든 과정에 대한 비용을 포함한다. 처리시간은 입력 영상의 해상도에 따라 다르고, 비선형적으로 증가하므로 직접 촬영한 FHD, HD 해상도의 영상을 활용해 각각 측정한다. 그림 3은 CCTV 실험 영상 데이터의 예시를 나타낸다.

4.2 패딩 추가에 따른 인식 정확도 성능 평가

본 실험에서는 제안하는 모델의 탐지 범위 추가에 따른 인식 정확도를 측정한다.



그림 3. CCTV 실험 영상 데이터의 예시
Fig. 3. Example of CCTV experimental video data

안면 탐지 결과는 B-box의 형태로 출력되며 출력된 B-box를 통해 각 안면에 대해 검증한다. 지나치게 작은 B-box 범위는 인식 성능에 악영향을 미치므로 본 연구에서는 최적의 여백을 가질 수 있도록 탐지 범위를 확장하는 AB-box를 사용하여 성능을 향상시킨다.

본 실험에서는 4명의 탐지 객체를 포함하는 CCTV의 영상 데이터에서 AB-box의 크기를 0부터 20까지 변경하며 수행되었다. 그림 4는 탐지 범위 추가에 따른 인식 정확도 성능평가를 나타낸다. 성능평가 지표는 분류 정확도(Accuracy)를 사용하며 세로축에 나타난다. 가로축은 AB-box의 크기를 의미하며 이는 B-box에 해당 패딩을 더한 값을 나타낸다. 패딩이 0인 경우, 제안하는 AB-box가 아닌 탐지 결과를 통해 도출된 B-box 그대로 사용함을 나타낸다. 각 그래프의 지표는 화질과 대조군 안면 인코딩 데이터의 수에 따라 각각 FHD/HD, 4/400으로 표기한다.

그림 4는 AB-box 크기에 따른 실험결과를 나타낸다. 실험 결과에 따르면 안면 탐지 결과에 AB-box 크기를 추가함에 따라 성능이 점진적으로 증가한다. 특히, 16~17픽셀을 추가했을 때 대조군이 4개인 경우에는 최대 95.5%, 400개인 경우에는 최대 86.6%로 높은 인식 정확도 성능을 보인다. 따라서 제안하는 모델이 단순히 탐지 결과를 통해 도출된 B-box 그대로 사용하여 매칭을 수행하는 기존의 모델보다 인식 성능을 30%p가량 상승시킨다. 또한 FHD와 HD를 비교했을 때, 화질이 더 좋은 FHD의 성능이 더 높다.

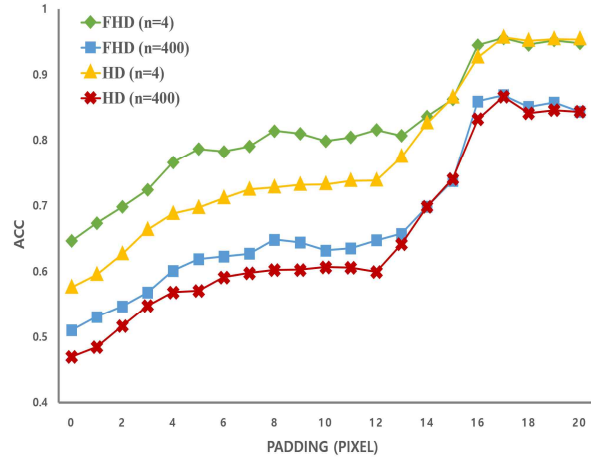


그림 4. AB-box 크기에 따른 인식 정확도 성능평가
Fig. 4. Recognition accuracy performance evaluation according to AB-box size

4.3 CCTV와의 거리에 따른 인식 정확도 성능평가

본 실험에서는 제안하는 모델의 탐지 객체와 CCTV의 거리에 따른 인식 정확도 측정한다. 이는 탐지 객체가 CCTV와 떨어져 인식하고자 하는 객체가 작을 경우에도 인식 정확도가 유지되는지를 평가하기 위함이다. 표 2는 객체와 CCTV의 거리에 따른 인식 정확도를 나타낸다. 성능평가 지표는 분류 정확도(Accuracy)를 사용한다.

실험 결과, 탐지 객체와 CCTV의 거리가 멀어지더라도 인식 정확도는 유지된다. 표 2를 보면 B-box를 사용한 경우보다 AB-box를 사용한 경우에 모든 상황에서 인식 정확도 성능이 높은 것을 확인할 수 있다. 특히 인물과 CCTV간의 거리가 5m인 경우, 대조군이 4장인 경우와 400장인 경우 모두 18% 이상의 더 높은 성능 정확도를 보인다. 즉, 제안하는 모델이 탐지 객체가 CCTV와 떨어져 작게 탐지되는 small object problem에 강경한 모델임을 의미한다.

표 2. 객체와 CCTV의 거리에 따른 인식 정확도
Table 2. Recognition accuracy according to distance between object and CCTV

Box	Embeddings	3m	4m	5m
B-box	4	0.8463	0.8749	0.5397
	400	0.8045	0.8396	0.2827
AB-box	4	0.9989	0.9658	0.7453
	400	0.9672	0.9401	0.4648

결과적으로, 성능평가 결과 제안하는 모델은 AB-box를 사용하여 3m일 경우에 15%, 4m일 경우에 9%, 5m일 경우에 18%로 인식 정확률이 더 향상되고 CCTV에 범죄자가 작은 모습으로 발견되더라도 꾸준한 탐지가 가능하다. 따라서 CCTV의 실제 범죄자에 대해 강경한 탐지와 인식을 통해 범죄 예방에 기여한다.

V. 결 론

본 논문에서는 어린이집 또는 유치원에 설치된 CCTV로 영상을 촬영하여, 영상 속 인물이 전과자 인지 판단하여 미리 경고하는 시스템을 제안한다. 감시 카메라가 어린이집과 같은 특정 위치에 성범죄자 출몰 등의 비정상적인 상황을 포착하면 공공 안전 시스템이 상황을 인식하고 즉시 기관에 보고함으로써 범죄 예방을 도모한다.

범죄자 얼굴 인식을 위한 얼굴 탐지기로는 RetinaFace 안면 탐지기를 사용했다. 안면 탐지기를 실시간으로 실행할 때 필요한 연산을 줄이고자 벡터화 방식을 도입했으며 AB-box를 사용하여 인식 성능의 상승을 도모하였다. 또한, 탐지 객체와 CCTV 간의 거리에 따른 인식 정확도 실험을 통해 작은 객체에 대해서도 강건한 모델임을 입증하였다. 따라서 본 연구를 통해 어린이집 CCTV에서 범죄자를 인식함으로써 범죄 수사에 도움을 주고 범죄 예방에 기여한다.

References

[1] KOSIS, KOREAN Statistical Information Service, <http://www.kosis.go.kr>. [accessed: Feb. 05, 2021]

[2] Kindicator, <https://www.index.go.kr>. [accessed: Feb. 05, 2021]

[3] J. Kim, J. Hwang, S. Choi, and H. Kim, "Trends of Sex Offenses against Minors: Focus on Sex Offenders Subject to Personal Information Registration Due to Offenses Committed in 2017", MOFEF, Dec. 2018.

[4] S. Jeong, "Preventive effect of releasing identities

of sexual offenders", Korean Law Association, Vol. 39, No. 39, pp. 269-290, Oct. 2010.

[5] Korea Ministry of Government Legislation, <http://www.law.go.kr>. [accessed: Feb. 05, 2021]

[6] S. Zhang, Y. Wu, and J. Chang, "Survey of Image Recognition Algorithms", In 2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), IEEE, pp. 542-548, Jun. 2020, <https://doi.org/10.1109/itnec48623.2020.9084972>.

[7] I. Masi, Y. Wu, T. Hassner and P. Natarajan, "Deep Face Recognition: A Survey", In 2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI). IEEE. pp. 471-478, Nov. 2018, <https://doi.org/10.1109/sibgrapi.2018.00067>.

[8] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection", 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), Vol. 1, pp. 886-893, Jun. 2005, <https://doi.org/10.1109/CVPR.2005.177>.

[9] Y. Kortli, M. Jridi, A. A. Falou, and M. Atri, "Face Recognition Systems: A Survey", Sensors, Vol. 20, No. 2, pp. 342. Jan. 2020. <https://doi.org/10.3390/s20020342>.

[10] B. Amos, B. Ludwiczuk, and M. Satyanarayanan, "OpenFace: A general-purpose face recognition library with mobile applications", CMU School of Computer Science, Vol. 6, No. 2. Jun. 2016.

[11] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering", In 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, pp. 815-823, Jun. 2015, <https://doi.org/10.1109/cvpr.2015.7298682>.

[12] J. Deng, J. Guo, E. Ververas, I. Kotsia, and S. Zafeiriou, "RetinaFace: Single-Shot Multi-Level Face Localisation in the Wild", IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 5203-5212, Jun. 2020,

<https://doi.org/10.1109/cvpr42600.2020.00525>.

[13] S. Gidaris and N. Komodakis, "Object Detection via a Multi-region and Semantic Segmentation-Aware CNN Model", IEEE International Conference on Computer Vision (ICCV), pp. 1134-1142, Dec. 2015, <https://doi.org/10.1109/iccv.2015.135>.

[14] V. Le, J. Brandt, Z. Lin, L. Bourdev, and T. S. Huang, "Interactive Facial Feature Localization", Interactive Facial Feature Localization. In Computer Vision - ECCV, Vol. 7574, pp. 679-692, 2012, https://doi.org/10.1007/978-3-642-33712-3_49.

저자소개

김 현 빈 (Hyun-Bin Kim)



2021년 2월 : 경기대학교
컴퓨터공학부 (공학사)
2021년 2월 ~ 현재 : 경기대학교
컴퓨터과학과 (석사과정)
관심분야 : 컴퓨터 비전, 객체탐지

최 낙 훈 (Nakhoon Choi)



· 2020년 2월 : 경기대학교
컴퓨터공학부 (공학사)
· 2020년 2월 ~ 현재 : 경기대학교
컴퓨터과학과 (석사과정)
· 관심분야 : 블록체인, 인공지능

강 지 수 (Ji-Soo Kang)



2020년 2월 : 경기대학교
컴퓨터공학부 (공학사)
2020년 2월 ~ 현재 : 경기대학교
컴퓨터과학과 (석사과정)
관심분야 : 데이터 마이닝,
헬스케어, 빅데이터 분석, 추천
시스템, 인공지능

임 소 현 (So-Hyun Lim)



2016년 2월 ~ 2020년 2월 :
경기대학교 컴퓨터공학부
(공학사)
2020년 2월 ~ 현재 : 경기대학교
컴퓨터과학과 (석사과정)
관심분야 : 영상 처리, 컴퓨터
비전, 인공지능

김 희 열 (Heeyoul Kim)



2000년 : 한국과학기술원 전산학과
(공학사)
2002년 : 한국과학기술원 전산학과
(공학석사)
2007년 : 한국과학기술원 전산학과
(공학박사)
2009년 ~ 현재 : 경기대학교

컴퓨터공학부 교수

관심분야 : 정보보호, 블록체인