

영상 데이터 개인정보 영역 객체 검출 최적화 기법 연구

송인준, 김차중*

한밭대학교 정보통신전문대학원, *한밭대학교 컴퓨터공학과

ijsong@watosys.net, cjkim@hanbat.ac.kr

Object Detection Optimization Technique for Personal Information Area of Video Data

Injun Song, Chajong Kim*

Department of Computer Engineering, Graduate School of Information & Communications, Hanbat National University

*Department of Computer Engineering, Hanbat National University

요 약

본 연구는 영상 데이터에 대해 개인정보를 비식별화하기 위한 임베디드 시스템에서의 인공지능 기반 개인정보 영역 객체 검출 최적화 방법을 제안한다. 객체 검출 최적화 방법으로는 먼저, 객체 검출 시 개인정보 영역에 대한 검출율을 높이기 위하여 자이로센서를 이용하여 영상 획득 시 영상 데이터에 대한 촬영 각도를 수집하고 수집된 촬영 각도를 통해 영상 데이터를 수평 영상으로 변환하였고, 이를 기반으로 학습 데이터의 영상 해상도의 크기 변화와 학습 엔진의 학습 방법에 변화에 따른 각각의 학습모델을 생성하여 실험적 방법으로 최적의 학습 모델을 선정하여 평가한 결과 유효성을 확인할 수 있었다.

I. 서 론

최근 인공지능의 한 분야인 딥러닝 기술의 발달에 따라 자율주행과 로봇, 스마트폰 응용 등의 분야에 많이 활용되고 있다. 이러한 인공지능 기술이 우리 생활에 더욱 밀접하게 확장되면서, 인공지능이 자동으로 데이터를 수집하고 학습하는 경우가 증가하고 있다[1]. 특히, 자율주행차나 로봇 분야에서는 실제 환경의 인공지능 학습데이터를 위해 많은 양의 영상 데이터를 수집하고 있으며, 개인정보(얼굴, 차량 번호판 등)가 포함되어 있어서 법적으로 많은 논란이 되고 있다[2]. 개인정보가 포함된 영상 데이터는 법적 요구에 따라 개인정보 영역에 대해 마스킹 처리 등과 같은 비식별화를 거쳐 저장하고 관리, 이용, 전송되어야 한다. 자율주행차, CCTV 등에서 촬영한 영상 데이터의 비식별화는 법적으로 요구되는 매우 중요한 과정이다[3]. 개인정보 영역을 검출하는 기존 연구에서는 전처리 과정과 개인정보 영역 검출 과정에서 딥러닝이나 영상 처리 방법을 활용하고 있어, 임베디드 시스템과 같이 컴퓨팅 자원이 부족한 환경에서는 이러한 복잡한 과정을 적용하기 매우 어렵다[4-6]. 본 논문에서는 카메라로부터 취득한 영상 데이터에서 개인정보 영역 검출하기 전 단계인 영상 전처리 과정에서 수평 각도의 영상 데이터로 변화하여 검출 성능을 높이기 위한 전처리 기법을 제안하고, 임베디드 시스템을 통해 개인정보 영역을 검출하기 위한 최적화된 입력 영상 해상도를 설정하고 학습모델을 찾는 방법을 제안한다.

II. 영상 데이터 개인정보 영역 검출 최적화

1. 영상 데이터 개인정보 영역 검출을 위한 전처리

본 논문에서는 대부분의 딥러닝 알고리즘은 비정형의 변형에 강인한 특성이 있으므로 카메라로부터 수집되는 영상은 카메라 자체 오류로 인한 비정형 변형을 제외하고 회전과 평행이동만 있는 것으로 가정한다. 자체 오류의 비정형으로 인해 수학적 보정이 어렵다.

영상 데이터의 3차원 공간의 점 (X, Y, Z)를 X축, Y축, Z축을 중심으로 θ 라디안(radian) 회전시키는 행렬을 각각 $R_x(\theta)$, $R_y(\theta)$, $R_z(\theta)$ 라 하고 수식 (1)로 표현되며, 카메라에 부착된 기울기 센서로부터 구한 X, Y, Z축의 기울기 값을 적용한다.

$$R_x(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}, \dots\dots\dots(1)$$

$$R_y(\theta) = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix},$$

$$R_z(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

카메라의 경우 중력의 영향으로 X축과 Y축의 변형이 주로 생기며, Z축은 바람과 같은 환경적인 요인으로 발생한다. 기울기 센서의 값에 대한 데이터와 환경적인 요인을 분석하면, 영상 데이터의 상황에 따른 변형 정도를 예측할 수 있다.

2. 영상 데이터 개인정보 영역 검출을 위한 학습모델

영상 개인정보 비식별화를 위해서는 먼저 영상 이미지에서 개인정보 영역을 추출해야 한다. 많은 연산량을 요구하는 인공지능 딥러닝을 임베디드 시스템에 적용하기 위해서는 서버나 일반 PC에서 사용되는 인공지능 모델이 아닌 경량의 모델을 사용해야 한다. 이로 인해 대부분의 임베디드 시스템에서는 검출률과 처리 속도에 많은 제약이 따를 수밖에 없다. 현재의 AI 고속연산이 가능한 칩셋을 사용하는 범용적인 사양의 임베디드 시스템 환경에서 빠른 처리 속도와 높은 검출률을 얻을 수 있는 모델을 얻기 위해서는 학습 영상 데이터의 해상도 조절을 통해 최적의 학습 모델을 선정하는 방법이 필요하다. 이를 위해 본 논문은 학습데이터 이미지의 입력 영상 해상도 변화에 따른 생성된 학습 모델을 실험적인 방법을 통해 주어진 환경의 임베디드 시스템에서 실험을 통해 최적의 학습 모델을 찾

* Corresponding Authors: Chajong kim(cjkim@hanbat.ac.kr)

고자 하였다.

III. 실험

1. 실험 준비

최적의 영상 데이터 해상도를 얻기 위해서는 다양한 실험이 필요하다. 따라서 본 논문에서는 2K(1920 x 1080) 해상도 이미지의 입력 영상 데이터에 대해 실험적인 방법을 통해 학습 영상 데이터의 해상도를 최적화하고자 한다. 본 논문에서는 표 1과 같이 학습용 원본 영상 데이터는 총 2천 장이며, 이 중에서 1,800장은 학습용으로 사용하고, 200장은 검증용으로 사용한다.

표 1 학습 영상 데이터 원본의 구성

종류	이미지 크기	이미지 개수	차량 번호판 객체 수	얼굴 객체 수
학습용 (train)	1920 x 1080	1800	3543	3764
검증용 (valid)	1920 x 1080	200	411	106
성능실험용 (test)	1920 x 1080	200	412	208

그림 1은 학습 데이터와 검증용 데이터의 일부이며, 임베디드 장치에 연결된 2K 영상획득장치에서 촬영된 도로주행 영상 데이터와 스마트폰의 2K 해상도로 촬영된 일반 영상 데이터를 대상으로 실험하였다.

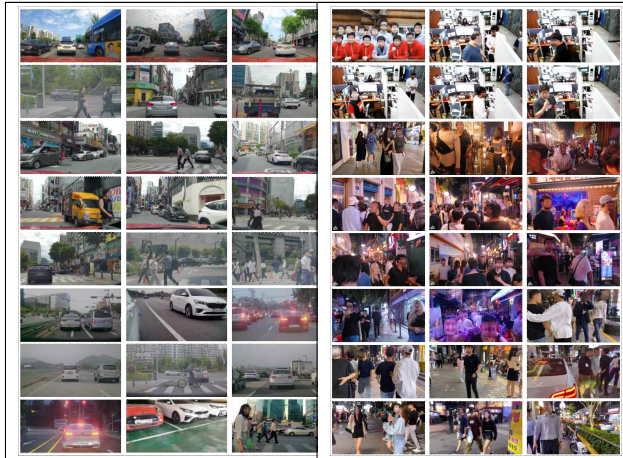


그림 1 학습 데이터

본 논문에서 목표로 하는 검출률과 처리 속도는 표 2에서와 같이 일반적인 PC 환경에서 가장 좋은 성능을 보인 경우를 기준으로 하며, 다양한 조건으로 해상도로 변환하여 학습과 추론을 하였을 때 기준에 가장 근접한 성능을 보이는 경우를 찾고자 한다.

표 2 PC 환경에서의 개인정보 영역 검출 성능 결과

대상 개체수	찾은 개체	매칭 개체수	못찾은 개체	에러	불명확	매칭률 (%)	속도 (fps)
620	662	598	13	55	9	96.45	31.25

2. 실험

본 논문에서 설명한 입력 영상에 대한 기울기센서값에 의한 전처리 실험

결과 결과는 그림 2와 같다. 그림 2에 보인 것과 같이 기울어진 영상에 대한 전처리 후 객체 검출률이 최대 13% 이상 높아지는 것을 확인할 수 있다.



그림 2. 전처리 실험 결과

또한, 검출률과 처리 속도 모두를 만족할 수 있도록 최적의 해상도를 찾기 위해 표 3, 표 4와 같이 2가지 실험 방법으로 학습 모델을 생성하여 그중 표 2의 성능에 가장 근접한 성능의 학습 모델을 선정하였다.

표 3 <실험 방법 1>의 입력 영상과 학습 영상 크기

순번	입력 영상 크기 (가로 크기 기준)	학습 영상 크기
1	320	320
2	416	416
3	480	480
4	640	640
5	800	800
6	960	960

표 4 <실험 방법 2>의 입력 영상과 학습 영상 크기

순번	입력 영상 크기 (가로 크기 기준)	학습 영상 크기
1	1920	320
2	1920	416
3	1920	480
4	1920	640
5	1920	800
6	1920	960

<실험 방법 1>과 <실험 방법 2>를 통해 학습된 딥러닝 학습모델에 대한 성능 검증은 검출 대상 객체를 찾은 영역(True Positive)과 검출하지 못한 객체 영역(True Negative), 객체 영역이 아닌데 객체로 검출한 영역(False Positive)을 수치로 분석하여 측정하였다. 일반적으로 딥러닝 학습 모델의 성능은 정밀도(Precision)와 재현율(Recall)로 표현하는데, 다음의 수식으로 정의된다. 정밀도는 검색된 결과 중에서 관련 있는 것으로 분류된 결과물의 비율이고, 재현율은 관련 있는 것으로 분류된 항목 중에서 실제 검색된 항목들의 비율이다.

$$Precision = \frac{TP}{TP + FP} \dots\dots\dots(2)$$

수식 (2)에서 정밀도는 TP(True Positive)는 True로 분류한 것 중에서 실제로 True인 것의 비율이다.

$$Recall = \frac{TP}{TP + FN} \dots\dots\dots(3)$$

(3)에서 재현율은 실제 True인 것 중에서 모델이 True라고 한 것의 비율이다. 본 연구에서는 <실험 방법 1>과 <실험 방법 2>를 통한 학습의 경우, 표 5와 같이 mAP가 더는 증가하지 않는 학습 횟수를 최적의 학습 횟수로 선택하고, 학습 모델을 생성하였다.

표 5 실험 방법 2의 416 크기 학습 결과 예시

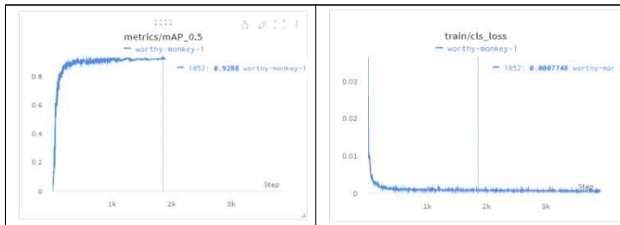


표 6은 실험 방법 1, 2와 전처리 방법을 적용하여 임베디드 시스템에서 최종적으로 실험과 성능 결과이며, 그림 3은 학습 모델을 통해 검출된 개인정보 영역에 대한 비식별화 처리 결과이다.

표 6 학습모델의 성능평가 결과

순번	학습모델	GT	DA	TP	TF	정밀도 (%)	재현율 (%)	속도 (FPS)
1	320크기 6,268번 학습	620	186	154	440	82.80	24.84	45.79
2	416크기 3,312번 학습	620	375	348	282	92.80	56.13	31.02
3	480크기 4,096번 학습	620	445	407	182	91.46	65.65	24.39
4	640크기 2,238번 학습	620	530	470	101	88.68	75.81	13.03
5	800크기 1,598번 학습	620	557	519	78	93.18	83.71	8.05
6	960크기 1,154번 학습	620	578	527	61	91.18	85.00	6.06
7	320크기 2,052번 학습	620	547	501	95	91.59	80.81	47.60
8	416크기 1,852번 학습	620	587	525	74	89.44	84.68	30.02
9	480크기 2,238번 학습	620	515	474	118	92.04	76.45	24.16
10	640크기 1,776번 학습	620	563	509	76	90.41	82.10	12.96
11	800크기 1,628번 학습	620	572	505	76	88.29	81.45	8.25
12	960크기 1,158번 학습	620	598	538	51	89.97	86.77	5.91

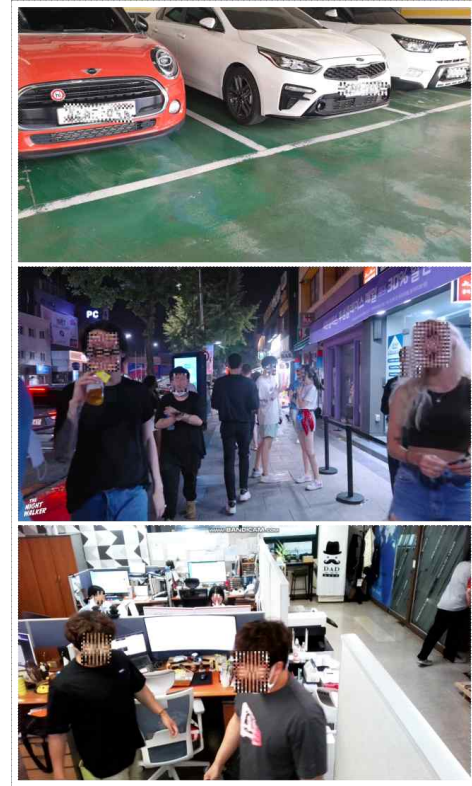


그림 3. 영상 데이터 개인정보 영역 객체 검출 결과

III. 결론

영상 데이터의 전처리 과정은 기존 연구의 단점인 처리 시간을 단축하기 위한 해결 방안으로 6축의 기울기 센서를 활용할 것을 제안하였으며, 임베디드 시스템에 적용할 수 있는 재현율과 실시간 처리 속도를 얻을 수 있었던 <실험 방법 2>에서 학습 데이터의 이미지 해상도 1920 사이즈를 학습 시 416 사이즈로 변환하여 생성된 학습 모델을 사용하였을 때 정밀도 89.44%와 재현율 84.68%, 처리 속도 30.02 FPS의 성능을 얻을 수 있었다. 본 연구에서 실험적으로 사용한 영상 데이터의 다양성을 높이고 실험적 방법으로부터 정형화된 방법으로의 개선을 위한 연구가 요구되며 개인정보 영역에 대한 오검출이나 미검출을 최소화할 수 있는 방안을 향후 연구함으로써 산업적인 활용 가치를 좀 더 높일 수 있을 것으로 기대된다.

참 고 문 헌

- [1] 김윤정, 윤혜선, “인공지능 기술의 활용과 발전을 위한 제도 및 정책 이슈”, 한국과학기술기획평가원, ISSUE PAPER 2016-07, pp.3-30, 2016.
- [2] 민경욱, 최정단, “자율주행 인공지능 기술 개발을 위한 데이터 수집 및 학습 플랫폼”, 2020년도 한국자동차공학회 춘계학술대회, 2020.
- [3] 원병철, “어린이집 CCTV 영상 반출하려면 모자이크 필수... 비용 문제는 해결 못해”, 보안뉴스, 2021, 04.23., <https://www.boanews.com/media/view.asp?id=x=96831>
- [4] 홍은빈, 전준호, 조성현, 이승용, “딥러닝을 이용한 영상수평 보정”, 한국컴퓨터그래픽스학회, v.23 no.3, pp.95 - 103, 2017.
- [5] M. Osadchy, Y.L. Cun and M.L. Miller, “Synergistic face detection and pose estimation with energy-based models”, Journal of Machine Learning Research, pp. 1197-1215, 2007.
- [6] Wang Yongmei Michelle, Zhang Hongjiang, “Detecting image orientation based on low-level visual content”, Computer Vision and Image Understanding, 93(3), pp.328-346, 2004.