

감시 영상을 활용한 OpenPose 기반 아동 학대 판단시스템

유혜림¹ · 이봉환^{2*}

An OpenPose-based Child Abuse Decision System using Surveillance Video

Hye-Rim Yoo¹ · Bong-Hwan Lee^{2*}

¹Lecturer, Department of Electronics, Information and Communications Engineering, Daejeon University, Daejeon, 34520 Korea

^{2*}Professor, Department of Electronics, Information and Communications Engineering, Daejeon University, Daejeon, 34520 Korea

요 약

최근 어린이집이나 유치원 등 교육기관에서 아동학대가 빈번히 발생하고 있다. 정부는 CCTV 설치를 의무화하였지만 CCTV 영상을 열람하는 것이 쉽지 않다. 본 논문에서는 CCTV 영상을 이용하여 아동학대를 판단하는 모델을 제안한다. 먼저 아동학대란 성인이 물리적으로 아동에게 가해를 하는 것이므로 성인과 아동을 분류하는 모델이 필요하다. 기존의 Haar기법을 사용하여 성인과 아동을 분류하려면 정면 영상이 필요하지만 OpenPose를 사용하면 정면과 측면에 구애받지 않고 성인과 아동을 분류할 수 있다. 본 연구에서는 아동이 학대를 당할 때 성인과 아동의 자세의 특성을 적용하여 아동 학대 판단 모델을 설계 및 구현하였다. 구현한 시스템은 현재 설치되어있는 CCTV를 활용하므로 추가적인 설치가 필요하지 않고 실시간으로 아동학대가 발생하고 있는지 모니터링 할 수 있으므로 이에 따른 빠른 대처가 가능할 것으로 사료된다.

ABSTRACT

Recently child abuse has occurred frequently in educational institutions such as daycare center and kindergarten. Therefore, government made it mandatory to install CCTVs, but it is not easy to inspect the CCTV images. In this paper, we propose a model for judging child abuse using CCTV images. First of all, child abuse is a physical abuse of children by adults, thus a model for classifying adults and children is needed. The existing Haar scheme uses the frontal image to classify adults and children. However, the OpenPose allows to classify adults and children regardless of frontal and side image. In this research, a child abuse judgment model was designed and implemented by applying characteristics of adult and child posture when a child was abused. Since the implemented system utilizes the currently installed CCTV image, it is possible to monitor the child abuse in real time without any additional installation, which enables us to cope with the abuse promptly.

키워드 : 머신러닝, OpenPose, 의사결정트리, 아동학대, CCTV

Key word : Machine Learning, OpenPose, Decision Tree, Child Abuse, CCTV

Received 15 December 2018, Revised 31 December 2018, Accepted 22 February 2019

* Corresponding Author Bong Hwan Lee(E-mail:blee@dju.kr, Tel:+82-42-280-2553)

Professor, Department of Electronics, Information and Communications Engineering, Daejeon University, Daejeon, 34520 Korea

Open Access <http://doi.org/10.6109/jkiice.2019.23.3.282>

print ISSN: 2234-4772 online ISSN: 2288-4165

©This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
Copyright © The Korea Institute of Information and Communication Engineering.

1. 서 론

아동에 대한 학대는 매우 오래전부터 여러 가지 형태로 자행되어왔다. 우리나라는 2000년도에 들어서면서 아동복지법 개정을 통하여 사회가 아동학대에 책임을 져야 한다는 법적, 제도적 토대를 마련하였다. 심리적으로나 육체적으로 발달상태에 있는 미완의 아동이 학대에 의해 신체적 건강과 정신적 건강이 침해되면 향후 성장과 발달에 치명적인 영향을 받을 수밖에 없다. 특히, 아동을 보호해야 하는 기관에서 빈번히 아동학대가 발생한다는 것은 우리 사회가 얼마나 아동학대 문제에 대해 관심을 기울여야 하는지를 보여준다[1].

정부는 2015년 1월 인천 송도 어린이집 교사의 4세 여아 폭행 사건을 계기로 어린이집 CCTV 설치를 의무화하도록 하였다. CCTV 설치는 현재 대부분 완료되었으나 CCTV의 영상을 부모가 보고자 하여도 어린이집에서 내건 조건에 맞아야 확인할 수 있다. 문제는 개정안은 ‘아동학대나 안전사고 피해가 의심될 경우 보호자가 어린이집 원장에게 CCTV 열람을 요청할 수 있고 원장은 10일 이내 학부모에게 승인 여부를 통지’ 하도록 했다. 열람 요구의 정당성 여부를 판단하는 1차적 역할을 조사대상일 수도 있는 어린이집 원장에게 맡긴 것이다. 학대 의심 사유를 적어내도 CCTV 설치 및 운영 자체를 어린이집 원장이 주도해 열람 조건에 적합하지 않다고 결론을 내리면 의사 소견서를 제출하거나 경찰에 고소하는 등 힘든 과정을 거쳐야 한다. 또한, 어린이집 측이 CCTV 영상을 고의로 삭제하거나 편집할 가능성도 배제할 수 없다. 아동학대 예방과 증거 확보 측면에서 CCTV의 효용성은 입증되었지만 증거를 확보하여 열람하기가 쉽지 않다.

따라서 본 논문에서는 CCTV의 실시간 영상을 통해 아동학대가 의심되는 상황을 판단하는 시스템을 제안하고자 한다. 2장에서는 본 논문에서 사용한 OpenPose Library와 기계학습의 특징을 설명하고 3장에서는 제안하는 아동학대 판단시스템의 구상도에 대해 설명한다. 4장에서 성능평가를 통한 결과를 도출한다.

II. 이론적 배경

2.1. OpenPose Library

OpenPose는 Carnegie Mellon University에서 개발한 라이브러리로 사진 또는 영상에서 실시간으로 여러 사람들의 동작에서 몸, 손, 그리고 얼굴 형태의 포인트를 추출할 수 있게 한다. 이 소프트웨어는 컴퓨터 비전과 머신러닝을 사용해 비디오 프레임을 처리하면서 개인과 연관시키는 것을 배우기 때문에 가령 특정인의 손이 자신의 팔꿈치에 가까워지는 것 등을 파악할 수 있게 한다. 이렇게 하면 한 번에 여러 사람을 추적할 수 있다. 기존의 Haar기법[2]을 사용하여 성인과 아동을 분류하려면 정면 영상이 필요하지만 OpenPose를 사용하여 정면과 측면에 구애받지 않고 성인과 아동을 분류할 수 있다.

Open Pose는 일반 카메라로도 사람의 스켈레톤 데이터를 따올 수 있게 만든 딥러닝 네트워크이다. 그림 1은 OpenPose의 전체 구조를 나타낸 것이다[3][4][5].

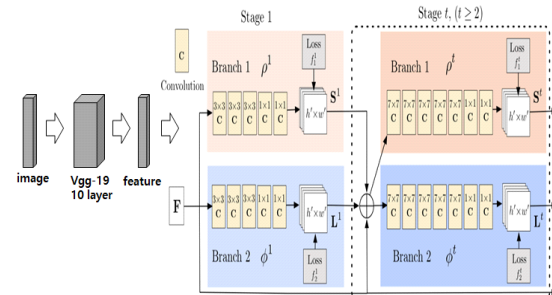


Fig. 1 Configuration of the OpenPose System

우선 vgg-19 network의 앞에서부터 10개 계층(layer)을 통과한 이미지 데이터는 특징(feature)이 강조된 형태로 출력(output)된다. 그리고 해당 출력은 1~6 스테이지(stage)의 입력으로 활용되어 affinity field와 confidence map을 구성하는데 사용된다. 각 단계에서 ground truth라는 레이블(label)과 비교를 통해 손실(loss)을 구하고 이를 최적화하는 형태로 네트워크(network)를 학습시킨다. 이것이 네트워크의 전체 구조이다. 그림 2는 Open Pose에서 관절 추출 순서를 나타낸 것이다.



Fig. 2 Joint extraction procedure in Open Pose

2.1.1. 몸체 탐지(body detecting)를 위한 이미지 입력 배치(batch) 형태의 이미지 데이터가 vgg-19 network의 입력으로 들어간다. 이미지는 네트워크를 통과하고 나면 특징점이 강조된 데이터(data)로 나오게 된다. 이제 이 데이터를 스테이지의 입력으로 넣는다.

2.1.2. 입력 이미지에서 각 몸체 부분 추출

스테이지의 입력으로 넣게 되면 2개의 가치를 쳐서 각각 confidence map과 affinity field를 구하는데 사용된다. Confidence map은 이미지 속 사람의 관절 위치를 파악하는데 사용하고 affinity field는 이미지에서 추출된 관절의 주인이 누구인가를 파악하는데 사용된다.

2.1.3. Confidence map을 통과한 사람 관절 파악

학습 초기 단계에서 입력 이미지에 따른 별 의미 없는 특징을 추출하게 된다. 하지만 이 특징을 human pose data와 비교하고 그 차이를 줄여나가는 방향으로 최적화를 진행하게 되면 특징점(feature)들은 점점 사람의 관절 위치를 가리키는 방향으로 나아간다.

2.1.4. Affinity field 추출

Affinity field는 각 관절의 주인을 파악하는 것이다. 주인을 파악하려면 각 부분(part)의 운동 방향에 대해 알아야 한다. Affinity field 자체가 벡터(vector field)이기 때문이다. Confidence map과 affinity field를 조합하여 완성된 인체 골격(human skeleton)을 조합할 때에는 greedy relaxation을 통해 각 부분을 조합하게 된다.

Confidence map을 통해 나온 각 부분을 조합할 때 이미지 속 사람이 여러 명이라면 greedy relaxation을 통해 각 부분의 주인이 누구인가에 대한 판단을 한다. 이러한 greedy relaxation에는 헝가리 해법(hungarian method)가 바탕이 된다. 위와 같은 과정을 반복하여 학습하면 결과적으로 Open Pose를 완성할 수 있게 된다.

2.2. 머신러닝(Machine Learning)

머신러닝(Machine Learning)은 인공지능의 한 분야로, 컴퓨터가 학습할 수 있도록 알고리즘 기술을 개발하는 분야를 말한다. 분류나 예측, 군집과 같은 기술, 모델, 알고리즘을 이용해 문제를 해결하는 것을 컴퓨터의 관점에서 머신러닝이라 한다. 머신러닝은 지도학습(Supervised Learning)과 비지도 학습(Unsupervised Learning)으로 분류된다. 지도학습에는 결과에 대한 사전지식이 필요

하다. 예를 들어 과거의 데이터를 다루는 경우 데이터를 판단하는 판단 값에 따라 각 데이터의 가중치를 두어 데이터의 신뢰도를 확인할 수 있다. 즉, 지도학습은 예측과 분류를 통해 이루어진다. 비지도 학습은 구체적 결과에 대한 사전지식이 없지만 데이터를 통해 유의미한 결론을 도출하고자 하는 경우 사용된다[6]. 사용자는 어떠한 결과를 도출하기 위해서 가설을 세우게 되고 컴퓨터 (Agent)는 훈련 데이터 셋을 통해 학습 알고리즘을 사용하여 학습한 뒤 결과를 도출한다. 표 1은 지도학습과 비지도학습의 알고리즘의 종류를 분류한 것이다.

Table. 1 Machine Learning Algorithms

Type	Tasks	Analysis methods/Algorithms
Supervised Learning	Classification	k-NN (k-Nearest Neighbor)
		Naive Bayes Classification
		SVM (Support Vector Machine)
		Decision Tree
	regression	Neural Networks
		Boosting Trees
Unsupervised Learning	Grouping	k-Means Clustering
		SOM(Self-Organizing Map)
	Dimension Reduction	PCA(Principal Component Analysis)
		SVD(Singular Value Decomposition)

2.2.1. k-최근접 이웃알고리즘(k-Nearest Neighbors)

k-NN 분류기는 주어진 데이터로부터 거리가 가까운 순서대로 k개의 데이터를 찾아서 그중에서 가장 많은 수의 데이터가 속한 그룹의 항목을 할당하는 방법이다. 데이터의 수가 늘어나면 분류기의 정확도가 높아질 수 있겠지만 그에 비례하여 계산하는 양과 컴퓨터 메모리 사용량도 함께 증가하는 문제점을 가지고 있다[7].

k-NN의 알고리즘은 비교적 간단하여 구현하기 쉽고 거리함수를 이용하는 학습 과정이 쉽고 데이터의 분포 모양에 크게 영향을 받지 않는다는 장점이 있다. 단점으로는 명목변수와 결과 예측 값에 대한 처리를 해야 하며, 속성 간의 관계성을 구하는데 어려움이 있다. 그림 3은 k=1 인 경우 k-NN 알고리즘을 나타낸 것이다.

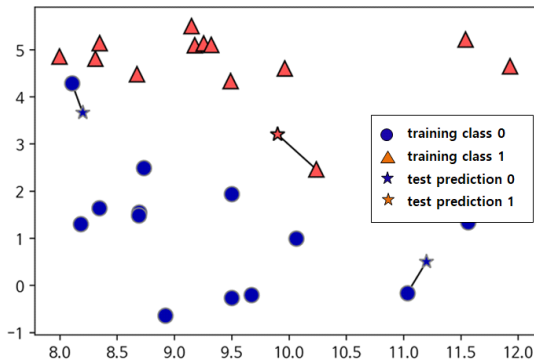


Fig. 3 k-NN Algorithm (k = 1)

2.2.2. SVM(Support Vector Machine)

SVM은 일반화 오차를 최소화 할 수 있는 방향으로 학습을 수행하는 선형분류기이며, 일반화 오차를 감소시키는 방향으로 학습이 이루어질 수 있도록 설계된 시스템이다. 낮은 복잡도의 선형분류기를 사용한 경우의 결정경계와 높은 복잡도의 선형분류기를 사용한 경우의 결정경계가 있다. 비선형분류기의 경우 학습데이터를 완벽히 분류하여 학습 오차가 없는 반면에 선형분류기의 경우 학습데이터를 제대로 구분하지 못하는 학습 오차가 발생한다. SVM은 독립변수의 차원이 큰 경우에 효율적이고 결정경계를 구한 후에는 분류할 데이터에 대하여 다시 처음부터 계산이 불필요하기 때문에 메모리 사용량이 적고 비선형분리를 위해서 커널법 사용이 가능한 장점이 있다[8][9][10]. Gamma 매개변수는 가우시안 커널 폭의 역수에 해당한다. Gamma 매개변수가 하나의 훈련 샘플이 미치는 영향의 범위를 결정한다.

2.2.3. 의사결정 트리(Decision Tree)

의사결정 트리는 분류 또는 예측의 과정이 트리 구조에 의한 추론 규칙에 의해 표현되기 때문에 신경망 알고리즘, 회귀 분석 등에 비하여 연구자가 그 과정을 쉽게 이해하고 설명할 수 있다는 장점을 가지고 있다. 기본적으로 결정 트리는 결정에 다다른 예/아니오 질문을 이어 나가면서 학습한다[11].

반복된 프로세스는 각 노드가 테스트 하나씩을 가진 이진 트리를 만든다. 각 테스트는 하나의 축을 따라 데이터를 둘로 나누는 것으로 생각할 수 있다. 각 테스트는 하나의 특성에 대해서만 이루어지므로 나누어진 영역은 항상 축에 평행하다. 그림 4는 의사결정트리에서

깊이가 9인 경우를 나타낸다. 깊이를 9로하면 연속된 질문을 최대 9개로 제한하여 과대적합이 줄어든다. 깊이를 제한하지 않으면 트리는 무한정 깊어지고 복잡해질 수 있다.

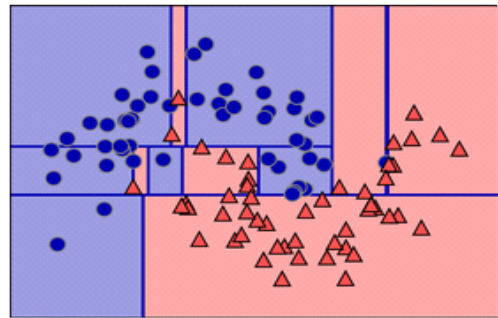


Fig. 4 Decision Tree Algorithm (depth = 9)

III. OpenPose와 머신러닝을 이용한 아동 학대 판단시스템

아동학대의 유형은 신체적 학대, 성적 학대, 심리적 학대, 방치 등이 있다. 제한하는 시스템에서는 신체적 학대를 판단 범위로 한다. 신체적 학대는 성인이 아동에게 직접적이고 물리적인 공격을 가하는 것을 말하며, 신체적 학대를 판단하려면 물리적인 공격을 판단하기 전에 성인과 아동을 판단하는 모델이 필요하다.

그림 5는 본 논문에서 제안하는 아동학대 판단 시스템의 구성도이다. 전체 시스템은 세부적으로 성인과 아동을 분류하는 모델(Adult Child Model)과 이 모델을 이용하여 아동학대를 판단하는 모델(Child Abuse Model)로 구성되어 있다.

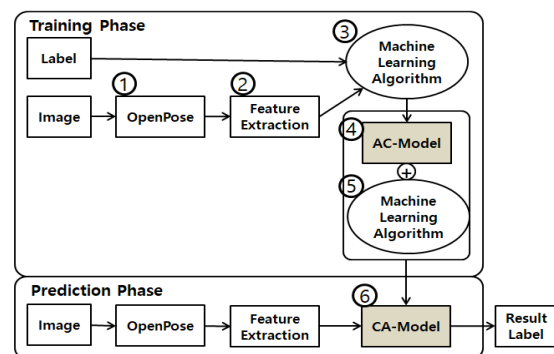


Fig. 5 A Child Abuse Judgment System

첫 번째 단계는 그림 5의 ①로서 입력 이미지로부터 openpose를 이용하여 사람의 관절을 추출한다. 두 번째 단계는 ②로서 앞에서 추출한 관절의 길이를 뽑아 성인과 아동을 분류할 수 있는 특징과 아동학대를 판단할 수 있는 특징을 만들어내는 과정이다. 이 과정에서 선별된 특징들은 머신러닝기반의 모델을 학습하기 위해 사용된다. 세 번째 단계는 ③으로서 앞에서 선택된 주요 특징들을 사용하여 머신러닝을 통해 ④의 성인과 아동 분류 모델을 생성한다. 다섯 번째 단계 ⑤는 앞에서 생성한 AC-Model과 ②에서 아동 학대판단에 사용할 특징을 함께 사용하여 머신러닝을 통해 ⑥인 아동 학대 판단 모델을 생성한다.

3.1. 성인과 아동 분류 모델 (AC-Model)

그림 6은 Open Pose를 이용하여 찾을 수 있는 관절의 위치 값을 보여준다[12].

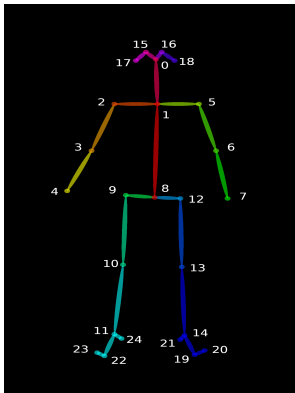


Fig. 6 Joint Point using OpenPose

각 프레임은 OpenPose를 이용하여 이미지의 pose 정보를 추출하여 json 파일로 결과 값이 저장된다. json 파일에는 인체의 주요 관절 25곳을 저장한다. 각 KeyPoint는 X, Y, C로 구성된다. X와 Y는 픽셀 값을 나타내며, C는 0.0부터 1.0 사이의 신뢰도를 나타낸다. 1.0으로 갈수록 신뢰성이 높다는 것을 보여준다[13]. 그림 7은 저장되는 json 파일의 일부분이다. 이 결과 값을 통해 각 관절의 좌표와 신뢰도를 확인할 수 있다. 그림 7에서 보여 주듯이 모든 관절의 값이 나오는 것은 아니며, 영상에서 확인되는 부분만 저장된다. 또한, 신뢰도가 0.3 이상의 값인 경우에 정확도를 신뢰할 수 있다.

```
"pose_keypoints_2d":
[
  982.003,394.245,0.11556,
  1003.59,419.765,0.342146,
  985.932,406.029,0.247897,
  962.405,435.451,0.0545821,
  0,0,0,
  1013.37,433.48,0.409336,
  993.748,474.591,0.29784,
  .....
]
```

Fig. 7 A json file for pose information

관절의 길이는 각 관절과 관절 사이의 거리를 구해서 알아낼 수 있다. 척추의 길이를 알아내기 위해서 목(1번)의 좌표 $A(X_1, Y_1)$ 와 엉덩이(8번)의 좌표 $B(X_2, Y_2)$ 를 정의하여 계산할 수 있다. 유클리드 거리 방법을 이용하여 척추의 길이(A와 B의 거리)를 계산하는 수식은 식 1과 같다.

$$\text{length} = L = \sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2} \quad (1)$$

그러나 입력되는 이미지의 크기가 다르고 사람이 가까이 있거나 멀리 있는 경우에 따라 같은 사람이라도 길이가 달라진다. 따라서 유클리드 거리 방법을 이용해서 나온 출력 값을 바로 적용할 수 없다. 이것은 절대 좌표로는 성인과 아이를 분류할 수 없음을 의미한다. 그러므로 본 연구에서는 관절의 비율을 계산한 수치 데이터를 이용하여 성인과 아동을 분류하는 모델(AC-Model)을 고안하였다. 성인과 아동의 신체적 특성은 머리의 비율이 성인이 될수록 작아지고 팔과 다리의 길이가 길어지는 특성을 가지고 있다. 그러나 그림 8의 (a)처럼 뒷 모습인 경우에는 눈, 코, 입의 좌표를 확인할 수 없기 때문에 머리의 크기를 측정할 수 없다.

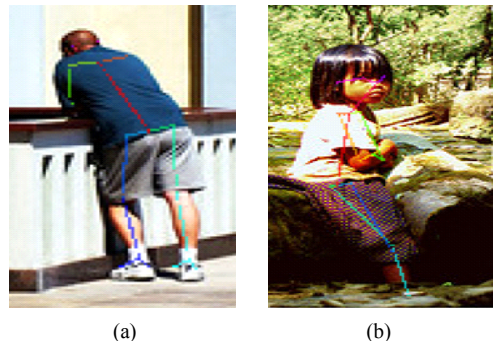


Fig. 8 Rear image of an adult and side image of a child

그림 8 (b)의 그림처럼 측면인 경우 신체 중 가로 의 길이(어깨, 엉덩이)의 값이 작아지므로 비율로 사용하 기에 부적절한 것을 확인할 수 있다. 관절 중 세로(척추, 팔, 다리)의 값은 정면일 때와 측면일 때에도 추출이 잘 되며 팔과 다리의 왼쪽과 오른쪽 중 한쪽만 측정되더라 도 사용할 수 있다.

식 2는 성인과 아동을 분류할 때 사용한 input data를 나타낸다. 이 모델은 몸 전체길이(팔+다리)의 길이를 척추의 길이로 나눈 비율을 사용하며 관절의 길이는 왼쪽 과 오른쪽 중에서 신뢰도 C 값이 큰 값을 선택하였다. 아동보다 성인인 경우 값이 커지며 머신러닝을 통해 학습하도록 하였다. 팔의 길이는 관절 2와 3의 거리와 3과 4의 거리를 더한 값을 사용하고, 다리의 길이는 9와 10 의 거리와 10과 11의 거리를 더한 값을 사용하였다. 척추의 길이는 1과 8의 거리를 사용한다.

$$Input\ data = \text{팔 길이} + \text{다리 길이} / \text{척추 길이} \quad (2)$$

3.2. 아동학대 판단 모델 (CA- Model)

본 연구에서는 신체적 학대 중 발로 학대를 가하는 경 우를 판단하였다. 그림 9의 이미지는 성인이 발로 아동을 때리는 장면을 Open Pose를 사용하여 관절을 추출한 이 미지이고, 그림 10은 일반 어린이집에서 활동하는 이미 지를 OpenPose를 사용하여 관절을 추출한 이미지이다.



Fig. 9 Child abuse image

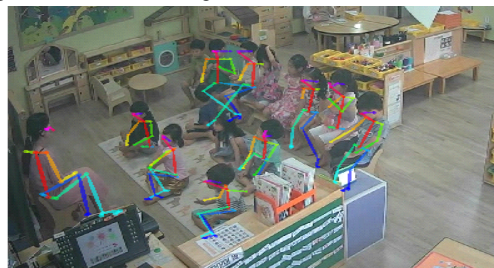


Fig. 10 An ordinary children's image

아동 학대 판단 모델을 학습시키기 위해서 세 개의 인 자 값을 사용하였다. 첫째, 성인과 아이를 분류하는 AC-Model을 이용하여 성인을 찾고 성인의 발과 가까이 에 있는 아이를 찾는다. 둘째, 발로 학대를 가할 때 항상 한쪽 다리가 들리는 포즈를 취하게 되므로 양쪽 발의 거 리가 멀어지는 특징을 찾아낸다. 양쪽 다리의 거리도 절 대적 길이는 적용하지 못하기 때문에 양쪽 발 사이의 거 리를 척추 길이로 나눈 비율을 사용하였다. 발로 학대를 가하게되면 일반적인 이미지보다 척추로 나눈 비율의 값이 커지게 된다. 셋째, 아이는 학대를 당할 때 웅크리 는 자세를 취하게 되는데 웅크리게 되면 척추와 발의 거 리가 짧아지게 된다. 역시 절대적 길이를 적용하지 못하 기 때문에 척추(1번)와 신뢰도가 높은 발(11번 또는 14 번) 사이의 거리를 몸길이(척추 길이 + 다리 길이)로 나 누 비율을 계산하였다. 학대를 당할 때 일반적인 이미지 보다 척추로 나눈 비율의 값이 작아지게 된다. 이렇게 세 가지의 특징으로 머신러닝을 이용하여 학습하였다.

IV. 실험 및 성능평가

4.1. 실험 데이터 평가모델

본 연구에서는 첫 번째 성인과 아동 인식 모델 생성 및 평가를 위해 imageNet[14]에서 제공하는 이미지 데 이터를 수집하였다. 아동은 만 5세 이하의 아동 사진을 사용하였다. 두 번째 아동학대 판단 모델 생성 및 평가 를 위하여 YouTube의 영상 중 발로 아동 학대를 가하는 68장을 아동학대 training set으로 훈련하였고, 일반 어 린이집에서 촬영한 영상 453장을 일반 영상 training set 으로 실험하였다. 표 2는 케이스 별 수집한 이미지의 수 이다.

Table. 2 Number of Images for training set

AC-Model		CA-Model	
Adult	Child	Normal	Abuse
300	377	453	68

표 3은 본 연구에서 시스템 성능을 검증하기 위해 사 용한 실험 환경을 보여준다. 머신러닝을 훈련하기 위해 파이썬 기반 오픈 소스 scikit-learn을 사용하였다[13].

Table. 3 Experimental Environment

	Configuration	Details
HW	CPU	CORE i7-1 860
	RAM	DDR3 4G PC3-12800*2
	VGA	GTX660 D5 2GB
	HDD	320GB Caviar Blue WD3200AAJS
SW	OS	ubuntu-16
	Python	3.5.2
	gcc	5.4.0
	Jupyter	4.4.0

성능평가는 식(3)의 정확도(Accuracy), 식(4)의 True Positive Rate(TPR), 식(5)의 False Positive Rate(FPR)의 평가 기준을 사용한다.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

$$True\ Positive\ Rate = \frac{TP}{TP + FN} \quad (4)$$

$$False\ Positive\ Rate = \frac{FP}{FP + TN} \quad (5)$$

4.2. 성인과 아이 분류 모델 평가

성인과 아동을 분류하는 모델을 만들기 위해 머신러닝 알고리즘 중 k-NN, SVM, Decision Tree를 각각 학습하여 모델을 생성하고 이 중에 더 정확하게 분류하는 모델을 아동 학대 판단 모델에서 사용하였다.

세 가지의 알고리즘을 사용하여 시험한 결과 k-NN은 k값이 1일 경우 88%로 가장 높은 인식률을 보였고 SVM은 gamma값이 3일 때 90%를 그리고 Decision Tree는 깊이가 15일 때 87.8%의 인식률을 보였다(그림 11).

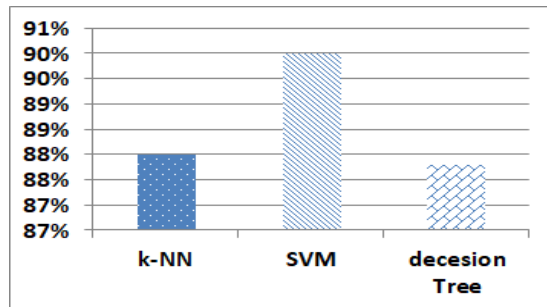


Fig. 11 Performance results of the three algorithms for classifying adult and child (AC-Model)

표 4는 세 가지 머신러닝 알고리즘의 성능결과를 보여준다. 정확도는 90%이하로 비슷한 결과 값을 보여주고 있으나 SVM으로 학습한 결과가 실제 성인을 인식하는 TPR값이 높고 아동을 성인으로 오인하는 FPR은 세 가지 머신러닝이 비슷한 값을 보이고 있다.

Table. 4 Performance comparison result of the proposed AC-Model

	k-NN	SVM	decision tree
Accuracy	87.8%	89.5%	87.8%
TPR	86.4%	90.9%	85.2%
FPR	11.2%	11.4%	10.3%

4.3. 아동학대 판단 모델 평가

아동 학대 판단 모델을 학습하기 위해서 성인과 아동 분류 모델은 인식률이 높게 나온 SVM을 적용하고 아동 학대 판단 모델에 k-NN, SVM, Decision Tree를 학습하였다.

그림 12는 AC-Model에 SVM을 사용하고 이를 적용하여 아동학대 판단 모델 알고리즘을 적용한 결과를 나타낸 것이다. k-NN은 k값이 2일 때 95%, SVM의 93.5%, Decision Tree는 깊이가 9일 때 92.2%의 학습 결과를 보였다.

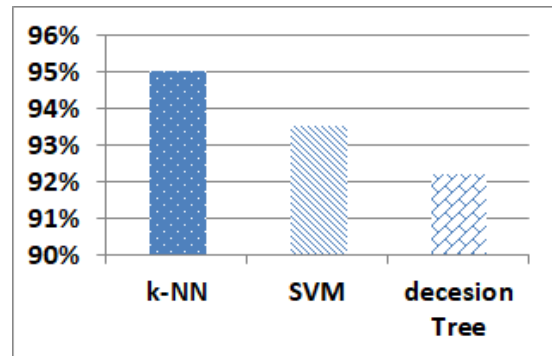


Fig. 12 Performance results of the three algorithms for abuse decision when SVM AC-Model is used

표 5는 세 가지 머신러닝 알고리즘의 성능결과를 보여준다. 정확도는 비슷한 값을 보여주고 있으나 학대를 인식하는 TPR 값이 k-NN과 SVM보다 decision Tree가 가장 좋은 성능을 보인다.

Table. 5 Performance comparison result of the proposed CA-Model when SVM AC-Model is used

	k-NN	SVM	decision tree
Accuracy	94.5%	94.5%	93.8%
TPR	50%	50%	91%
FPR	1.4%	1.4%	5.9%

V. 결론 및 향후 과제

본 논문에서는 어린이집이나 유치원 등 교육기관에서 아동학대를 감시하기 위한 시스템을 개발하였다. 아동학대 판단 모델을 구현하기 위해서 성인과 아동을 판단하는 모델을 제안하고 그 모델을 활용하여 아동학대가 발생했는지 판단하는 모델의 시스템을 구현하여 성능 평가를 실험하였다. 실험 결과 성인과 아동을 분류하기 위한 모델은 SVM으로 학습한 결과의 TPR 값이 높은 것을 알 수 있으며, 이것을 적용하여 아동학대 판단 모델에서는 Decision Tree로 학습한 결과가 높은 TPR 값을 보여주었다.

이 시스템을 활용하여 실시간 CCTV를 통해 아동학대가 발생는지 실시간으로 인지할 수 있고 감독관에 의해 아동학대로 판단되면 신속히 대응 할 수 있다.

본 연구에서는 아동학대를 구분하기 위해서 성인이 발로 가해하는 경우를 판단했다. 아동학대의 영상이 대부분 모자이크 처리되어 정확한 이미지 처리가 어려웠다. 모자이크 처리되지 않은 영상이라면 손동작의 모양을 활용하여 꼬집는 형태, 쓰다듬는 형태 등 손동작을 판별하여 좀 더 세세한 아동학대를 판별할 수 있는 모델을 구축할 수 있을 것이다.

ACKNOWLEDGEMENT

This research was supported in part by the Daejeon University Research Grants (2018) and the leading human resource training program of regional neo industry through the National Research Foundation of Korea(NRF) funded by the Ministry of Science, ICT and future Planning (grant number) (NRF-2016H1D5A1909154).

References

[1] Central Child Protection Agency, “Child Abuse & Near Korea,” 2018.

[2] B. W. Yoon, “A study on video based child and adult classification with biometry,” M.S. thesis, Department of Electrical and Electronic Engineering, University of Kyungshung, Feb. 2015.

[3] Z. Cao, T. Simon, S. E. Wei, and Y. Sheikh, “Realtime multi-Person 2D Pose Estimation using Part Affinity Fields,” in *Proceeding of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7291-7299, 2016.

[4] S. E. Wei, V. Ramakrishna, T. Kanade, and Y. Sheikh, “Convolutional Pose Machines,” in *Proceeding of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4724-4732, 2016.

[5] S. Qiao, Y. Wang, and J. Li, “Real-time human gesture grading based on OpenPose,” in *Proceeding of the 2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics*, pp. 1-6, 2017.

[6] K. S. Ahn, “Comparative Experiment and Evaluation on Machine Learning based k-Nearest Neighbor and Support Vector Machine,” M.S. thesis, Department of Information and Communication Engineering, University of Dankook, Dec. 2016.

[7] T. M. Cover, and P. E. Hart, “Nearest neighbor pattern classification,” *IEEE Transactions on Information Theory*, vol. it-13, no. 1, pp. 21-27, 1967.

[8] Y. Tang, “Deep learning using linear support vector machines,” *Workshop on Challenges in Representation Learning, ICML*, 2013.

[9] C. C. Chang, and C. J. Lin, “LIBSVM: A library for support vector machines,” *Transaction on Intelligent Systems and Technology*, vol. 2(3), no. 27, Nov. 2011.

[10] S. Tong, and D. Koller, “Support Vector Machine Active Learning with Applications to Text Classification,” *Journal of Machine Learning Research*, vol.2(1), pp. 45-66, Nov. 2002.

[11] A. Papagelis, and D. Kalles, “Breeding Decision Trees Using Evolutionary Techniques,” in *Proceeding of the Eighteenth International Conference on Machine Learning*, pp. 93-400, 2001.

[12] CMU-Perceptual-Computing-La. OpenPose [Internet]. Available: <https://github.com/CMU-Perceptual-Computing-Lab/openpose>.

[13] A. C. Muller and S. Guido, *Machine Learning with Python*, 1st ed. Germany, O'Reilly Media Inc, 2016.

[14] Image Set. [Internet]. Available: <https://image-net.org/>.



유혜림(Hye-Rim Yoo)

2005년 대전대학교 정보통신공학과(학사)
2007년 대전대학교 대학원 정보통신공학과(석사)
2011년 대전대학교 대학원 정보통신공학과(박사수료)
※ 관심분야 : 딥러닝, 머신러닝, 영상처리, 네트워크보안 등



이봉환(Bong-Hwan Lee)

1985년 서강대학교 전자공학과졸업(학사)
1987년 연세대학교 대학원 전자공학과 졸업(석사)
1993년 Texas A&M 대학교 대학원 전기 및 컴퓨터공학과 졸업(박사)
현재 대전대학교 전자정보통신공학과 교수
※ 관심분야 : 클라우드 컴퓨팅, 사물인터넷, 네트워크보안 등