

논문 2020-57-1-9

인공지능 기반의 자세 추정에 따른 모바일 헬스케어 동작 연구

(Analysis on the Mobile Healthcare Behavior using an Artificial Intelligence Based Pose Estimation)

하 태 용*, 이 후 진**

(Taeyong Ha and Hoojin Lee[©])

요 약

스마트폰의 디바이스를 활용하여 개인의 건강을 실시간 관리하는 트렌드 확산과 헬스케어 관련 디바이스가 많은 관심을 받고 있지만 아직은 실제 시장으로 확산하기에는 쉽지 않을 것 같다. 현재는 사용자가 비디오 콘텐츠나 휘트니스 앱을 통한 강사 따라하기형에 많이 의존하고 있다. 본 연구에서는 LSP(Leeds Sports Pose) 데이터셋에서 사용되는 데이터를 기반으로 인공지능(Artificial Intelligence: AI) 합성신경망(Convolutional Neural Network: CNN)의 PoseNet을 활용하여 자세 추정 학습 및 인체 좌표를 분석하여 그 결과값으로 손뼉치기를 구현하고자 한다. 개발 언어는 Node.js 프로그램을 사용하여 최종 운동량 결과를 모바일 헬스케어로 구현하고자 한다. 이를 통해 자세 추정에 대한 동작 결과가 직관과 오류의 의존성에서 벗어나는 동기가 마련된다. 인체 좌표를 통한 보다 정량적이고 신속한 자세 추정에 근거한 동작 분석이 가능하게 되었다. 이런 점에서 PoseNet은 헬스케어 플랫폼의 많은 기능 중 하나로 서비스 되는 개인 휘트니스 프로그램을 개발하고자 하는 우리의 목적에 상당히 부합하는 모델이다.

Abstract

Trends in managing personal health using smartphone devices and healthcare-related devices are receiving a lot of attention, but it is not easy to spread them to the real market for some time. Currently, users rely heavily on instructor training through video content or fitness applications. In this study, we use PoseNet of artificial intelligence(AI) convolutional neural network(CNN) based on the data used in the Leeds Sports Pose(LSP) dataset to analyze pose estimation and analyze human coordinates, and to applaud the result as a clap. The development language uses the Node.js program to implement the final momentum results into mobile healthcare. This provides motivation for the behavioral results of pose estimation to deviate from the dependency of intuition and error. Motion analysis based on more quantitative and faster pose estimation through body coordinates is now possible. In this regard, PoseNet can be a model that fits our purpose of developing a personal fitness program, being served as one of the many features of the healthcare platform.

Keywords : AI, Deep Learning, CNN, Fitness, PoseNet

I. 서 론

최근 인공지능(Artificial Intelligence: AI) 시대에 일반인의 건강에 대한 관심 증가로 인해 모바일 헬스케어와 그에 관련된 콘텐츠가 주목받고 있다^[1]. 스마트폰의

*학생회원, 한성대학교 스마트융합컨설팅학과
(Department of Smart Convergence Consulting, Hansung University)

**평생회원, 한성대학교 스마트융합컨설팅학과
(Department of Smart Convergence Consulting, Hansung University)

[©]Corresponding Author(E-mail : hjlee@hansung.ac.kr)

Received ; October 24, 2019 Revised ; November 4, 2019

Accepted ; November 7, 2019

디바이스를 활용하여 개인의 건강을 실시간 관리하는 트렌드 확산과 헬스케어 관련 디바이스가 많은 관심을 받고 있지만 아직은 실제 시장으로 확산은 한동안 쉽지 않을 전망이다. 단순히 신체 정보만을 수집하는 것이 아니라 실질적으로 사용 고객에게 어떤 효용성을 제공하는가에 따라 헬스케어 디바이스의 성공 여부를 판단할 수 있으며, 이를 반영한 다양한 헬스케어 디바이스가 시장에 나와 있지만 사용자의 건강 데이터 가운데 일부분만을 단편적으로 측정하고 있는 상황이다^[2]. 이런 문제점을 보다 효율적으로 해결하고자 하는 노력의 일환으로써 최근에는 딥러닝(Deep Learning: DL)을 비롯한 다양한 인공지능 기술의 적용이 활발해지고 있다. 특히

컴퓨터 비전의 이미지 인식 분야에서 인공지능 기술 기반으로 뛰어난 성능을 보여 주는 객체 인식 및 검출 알고리즘이 계속 발표되어지고 있으며, 기존 방법론으로는 해결이 어려운 문제나 기술들을 인공지능 딥러닝 기술을 활용함으로써 실제적이고 효과적으로 해결하는 성과를 거두고 있다^[3].

본 연구는 사용자의 편의성이 효과적으로 반영된 모바일 헬스케어 애플리케이션 개발을 위한 동작 연구로서 향후 모바일 헬스케어 시스템 구성에 도움을 주고자 한다. 이를 위한 구체적 개발 방법으로 인공지능의 딥러닝 뼈대인식 기술 기반으로 실시간 자세 추정 기법을 적용하여 운동자에게 카운트 및 자세교정 정보를 전달하고 실제 트레이너와 함께 운동하는 효과를 얻게 하여, 궁극적으로는 이러한 개인 운동 트레이너 서비스의 구현이 피트니스 애플리케이션 시장에서 우수한 경쟁력을 갖춘 응용 분야가 될 것으로 예상된다.

II. 선행 연구 및 한계

1. 인공지능

가. 인공지능, 기계학습, 딥러닝 관계

인공지능(AI), 기계학습(Machine Learning: ML), 딥러닝(DL)의 관계를 간략하고 명료하게 그림 1에서 보여주고 있다. 딥러닝은 규모가 큰 신경망을 사용한다는 점과 방대한 대용량의 데이터를 기반으로 학습한다 특징으로 기계학습 알고리즘으로 분야에 속하는 하나의 방식이다.^[4]

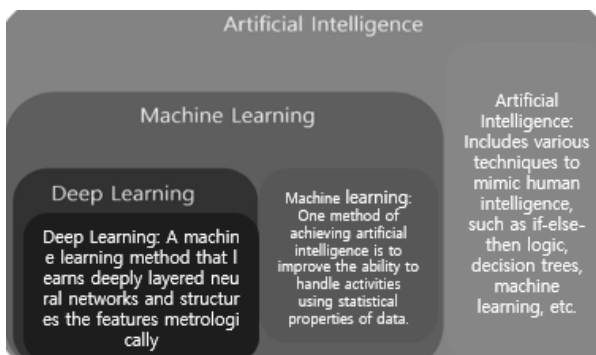


그림 1. 인공지능, 기계학습, 딥러닝 관계

Fig. 1. Artificial Intelligence, Machine learning, Deep learning Relation^[4].

나. CNN 주요 모델

이번에는 Convolutional Neural Network(CNN)의 주요

모델들에 대해 각 논문을 참고해 살펴보면 아래의 표 1과 같이 정리가 된다^[5~7].

표 1. CNN 모델의 종류

Table1. Types of CNN Models.

Year	Model	Contents	Author
2012 AlexNet	Neural Net Begins	ImageNet Classification with Deep Convolutional Neural Networks	Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton ^[8]
2014 GoogLeNet	Deeper Network Depth and Width	Going Deeper with Convolutions	Christian Szegedy ¹ , Wei Liu, Yangqing Jia ¹ , Pierre Sermanet ¹ , Scott Reed ^[9]
2015 VGGNet	Small Receptive Field and Network Depth	Very Deep Convolutional Networks For Large-Scale Image Recognition	Karen Simonyan & Andrew Zisserman ^[10]
2016 ResNet	8 Times More Depth than VGGNet	Deep Residual Learning for Image Recognition	Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun ^[11]
2017 MobileNet	Light and Efficient	MobileNetV2: Inverted Residuals and Linear Bottlenecks	Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, Liang-Chieh Chen ^[12]

다. PoseNet

(1) 소개

PoseNet의 경우 Tensorflow.js로 개발되어 크롬과 같은 브라우저에서 실시간으로 구동이 가능하기 때문에 다른 모델들에 비해 활용 가능성이 높다. 실제로 MobileNet 버전 모델은 PoseNet이 모바일 디바이스에서도 구동시키기 위해 고안된 모델이다. PoseNet은 키 바디 조인트의 위치를 추정하여 이미지 또는 비디오에서 사람의 포즈를 추정하는 데 사용할 수 있는 비전 모델이다. Google Creative Lab에서 출시했으며 TensorFlow.js에서 실행되므로 브라우저에서 실시간 포즈 추정이 가능하다. PoseNet을 사용하여 단일 포즈 또는 다중 포즈를 추정할 수 있으며 각 알고리즘마다 다른 알고리즘이 사용된다.

여기서는 단일 포즈 추정 알고리즘을 기반으로 자세 추정 학습을 하고자 한다^[13].

(2) 방법

그림 2와 같이 PoseNet 자세 추정의 제 1단계에서, 정 사각형 이미지가 사전 훈련된 모델을 통해 공급된다. 사용자는 MobileNet v1과 ResNet50 사이에서 모델 아키텍처를 선택할 수 있다. MobileNet 모델이 더 빠르므로 실시간 응용 프로그램에 더 적합하지만 ResNet50 모델보다 정확도가 떨어진다. 로드된 모델은 입력 이미지에서 사람을 감지하고 17개의 키포인트 (눈, 코, 귀, 어깨, 팔, 다리 등)에 대한 신뢰도 점수 및 위치 뿐만 아니라 포즈 신뢰도 점수를 반환한다. PoseNet은 이미지를 처리 할 때 사용자가 선택한 출력 단계에 따라 원본 이미지를 축소한다. 확장된 회선은 모델이 더 큰 출력 해상도를 만들어야 할 때 사용된다. 모델 출력은 키포인트 히트 맵 (키 포인트 당 하나의 채널) 및 오프셋 벡터 (키 포인트 당 두 개의 채널 (x, y))이며, 오프셋 벡터는 키 포인트 현지화 정확도를 향상시킨다. 키 포인트를 얻기 위해 각 신체체의 최고 점수 히트 맵 x 및 y 지수에 출력 보폭을 곱한 다음 해당 오프셋 벡터에 추가한다. 마지막으로 각 키 포인트 신뢰 점수는 히트 맵 위치의 신뢰 점수로서 자세 신뢰도 점수는 키 포인트 점수의 평균이다^[13~15].

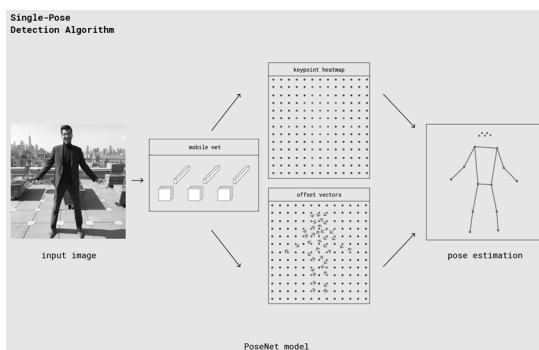


그림 2. PoseNet의 단일 포즈 Estimation 알고리즘
Fig. 2. Single-Pose Estimation Algorithm of PoseNet^[13].

2. 모바일 헬스케어

가. 헬스케어 어플리케이션 유형별 분류

모바일 헬스케어 분야는 일반적으로 모바일 기기인 의학용 기기와 센서와 더불어, 건강 정보 및 기기와 연결된 어플리케이션과의 효과적인 건강관리를 위해 제공되는 다양한 개인 건강 지침 시스템 등을 포함한다. 즉, 표 2와 같이 모바일 헬스케어에는 센서와 기기를 포함

하는 하드웨어, 정보의 제공을 위한 어플리케이션과 같은 소프트웨어적인 요소를 모두 포함한다^[16, 17].

표 2. 모바일 헬스케어 유형별 분류

Table2. Classification Of Mobile Healthcare by Type.

Forms & Examples	Biosensors & Examples
- Smart Devices - Smart phone - Smart Pad - Wearable Device - Bluetooth Headsets - Smart Watches - Smart Clothes - Smart Band - Smart Glasses - Smart Shoes - Patch & Tattoos - Smart Implants - Other	Activity Monitors - GPS - Acceleration Sensor - Vibration Motor - Detection Sensor Physical Index Monitors - Electrocardiogram (ECG) (e.g. AliveCor) - Electromyography (EMG) - Blood Oxygen Saturation and Pulse Measurement - Blood Glucose easurement

나. 헬스케어 어플리케이션 기능별 분류

모바일 헬스케어 어플리케이션은 표 3과 같이 운동법 제공, 행동추적 및 신체 정보 모니터링 등과 같은 기능별 분류가 가능하다^[17, 18].

표 3. 헬스케어 어플리케이션 기능별 분류

Table3. Healthcare Application Functional Classification.

Division	Main Function	Application Example
Offer Exercise	Determining the Type of Exercise that You Want, or Acquiring/Managing Exercise Courses	'Cody', 'Nike Training Club', 'Coco's Workout World', 'Runtastic Six Pack Abs', 'Stacked - Your New Personal Trainer', etc.
Activity Tracker	Calorie Consumption, Number Of Steps, Travel Distance, Sleep Monitoring, Muscle Movement, pose, etc.	'Fitbit', 'Kinetic GPS', 'Nike+(Kinetic, Training, Running, etc.)', 'NFL Play 60', 'Runtime-simple Run Tracking', 'Datalove', etc.
Physical Index monitors	Heart Rate, Temperature, Skin Conduction, Breathing, Glucose Level, Blood Oxygen Level, Heart Rate Measurement of Blood Pressure, etc.	'Fitbit', 'Adidas Fit Smart Fitness Tracker', 'Iriver on', 'Hexoskin', 'Angel Wristband', 'AskMD', etc.

다. 실용적 모델 사례

수작업 대체를 위해 자동화 작업에서 작업 대상의 물체 인식과 자세 추정 그리고 비주얼 서보잉 기술을 이용해 물체 조립 작업에 적용할 수 있다. 산업 현장의 제조 공정에 실제로 적용하기 위해서는 먼저, 실시간으로 물

체 검출이 안정적으로 이루어져야 한다. 하지만 기존 전통방식으로 동적 물체조립에 적용할 경우에는 정확한 물체 검출과 자세 추정 정밀도가 저하되는 문제점이 발생하였다. 따라서 인공지능 딥러닝의 CNN 기반으로 하여 실시간 물체를 검출하고 내부의 특징을 추출하여 대상 물체에 대한 영역 검출 및 정밀한 자세를 추정할 수 있다^[19].

III. 연구 방법

본 연구는 그림 3과 같이 ① LSP 데이터셋에서 사용되는 데이터를 기반으로 ② 인공지능 딥러닝 중 CNN 알고리즘을 이용하여 자세를 추정하고, 이 추정 모델을 Node.js 알고리즘으로 분석하여 ③ 최종 운동량 결과를 파악하여 모바일 헬스케어 구현을 하고자 한다.

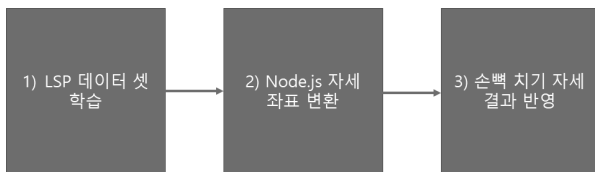


그림 3. 자세 추정에 따른 동작 연구 모형
Fig. 3. Motion Study based on the Pose Estimation.

1. Leeds Sports Pose(LSP) Dataset

LSP Dataset은 운동하는 사람의 이미지 2,000개로 이루어져 있다. LSP Extended Training Dataset은 이보다 좀 더 많은 10,000개의 이미지로 이루어져 있다. 이미지는 사람의 크기가 대략 150픽셀 정도 되도록 스케일링된 상태이다. 양 발목, 무릎, 엉덩이, 손목, 팔꿈치, 어깨, 그리고 목과 머리 끝까지 총 14개의 관절이 라벨링된 것이 특징이다. LSP Extended Training Dataset 전체와 LSP Dataset의 처음 1,000개의 이미지는 훈련 데이터셋으로 쓰였고, LSP Dataset의 나머지 1,000개의 이미지가 모델 성능 평가에 사용되었다.

2. PoseNet 학습 방법

PoseNet은 LSP나 LSP Extended Dataset과는 무관하게 학습된 모델이나 다른 모델들과 성능을 비교하기 위해 LSP Dataset만으로 테스트를 진행하였다. Single - Person Pose Estimation 알고리즘을 사용하였고, 모델 구조에 따른 설정 값은 표 4 및 그림 4와 같다.

표 4. PoseNet 모델 구조에 따른 파라미터 정보

Table4. Parameter Information According to PoseNet Model Structure.

Architecture	MobileNet	ResNet
Output Stride	16	32
Input Resolution	513	257
Multiplier	0.75	x
QuantBytes	2	2



그림 4. PoseNet의 Single-Person Pose Estimation
Fig. 4. Single-Person Pose Estimation on PoseNet.

PoseNet은 목과 머리 끝 포인트 대신 코, 눈, 귀 포인트를 예측하여 결과적으로 한 포즈 당 총 17개의 키포인트를 출력한다. 행렬로 표현하자면, 2(x, y 좌표) × 17(키포인트) × 1,000(테스트 이미지 개수). 그래서 PoseNet의 성능을 LSP Dataset으로 가능해보기 위해서는 결과데이터 포맷에 대한 추가적 처리가 필요하다.

PoseNet의 Key Points의 출력값을 .json 파일로 리다이렉트 하기 위해서는 LSP Dataset의 Annotation 포맷인 .mat와 같이 MatLab에서 읽을 수 있는 변수로 만들기 위해 고안한 방법은 그림 5와 그림 6과 같이 출력 값을 모두 .json 형식으로 변환시켜 저장하는 것이다. MatLab은 json 형식 텍스트를 디코딩하는 함수 'jsondecode'를 지원하기 때문이다.

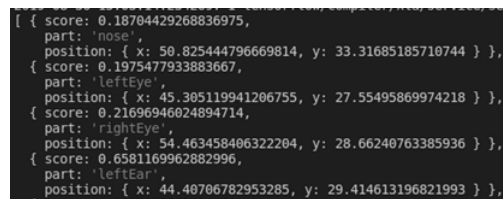


그림 5. PoseNet의 출력 값 형식 예시
Fig. 5. Example output value format for PoseNet.

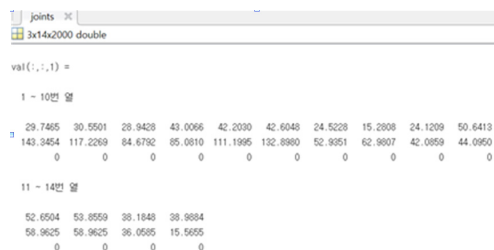


그림 6. LSP Dataset의 Annotation 포맷 .mat 예시
Fig. 6. Example of .mat Annotation Format of LSP Dataset.

PoseNet의 출력 값을 콘솔창으로 프린트 하는 대신에 임의의 json파일로 리다이렉트한다. 그 다음 json 파일을 열어 파일의 가장 앞과 뒤에 각각 '['와 ']'를 추가해주고, 각 이미지에 대한 Key Point 값이 ','로 구별되도록 ')' 뒤에 ','를 추가해주면 MatLab에서 열 수 있는 json 파일 형식이 그림 7과 같이 완성된다.

```
{ "keypoint": [[50.825444796669814,33.31685185710744],[456.077515954637335,31.926233699813427],[46.78921725499491,[57.89891586600575,77.27075124065236],[60.72728684718507,78[71.87809973097033,112.71612085720908],[38.07283589997644,1[127.76633509773225,50.67996672990257],[123.02816974328185,[95.68328263694673,62.997578186747624],[140.33906905474828,[110.19262481570706,112.85476833002112],[89.48216597857642,[47.18329558019972,128.2690556745121]]],[ "keypoint": [[68.7072795467154,41.96233050275869],[55.25903257971143,54.[55.63679047224586,80.49761541344311],[103.313973749658,81.[106.81957105347153,102.13446460530915],[64.80403006587046,
```

그림 7. PoseNet의 출력 값을 json 파일로 저장한 예시
Fig. 7. Example of saving PoseNet output to a .json file.

3. 자세 추정 파악

PoseNet 모델을 사용하여 사람의 포즈 추정을 위한 TensorFlow Lite 애플리케이션을 활용한다.

PoseNet은 주요한 인체 부위들의 위치를 감지하여 이미지 또는 동영상으로 부터 사람의 포즈를 추정하는 비전 모델이다. 포즈 추정 모델은 이미지에 나와 있는 사람이 누군지 식별하는 것이 아니라, 주요 인체 부위의 위치만 식별하고, TensorFlow Lite 디바이스의 카메라를 이용하여 사람의 주요한 인체 부위를 실시간으로 감지 및 표시하는 Android용 애플리케이션을 제공한다.

4. 눈, 코, 귀 포인트로 머리 끝 포인트 구한 방법

PoseNet을 활용하여 눈, 코, 귀 포인트를 머리 끝으로 인식하게 하는 구조로 아래의 그림 8과 같이 알고리즘으로 구성된다.

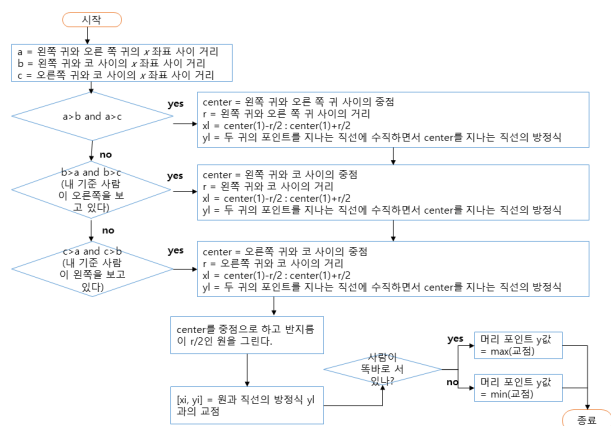


그림 8. 머리 포인트 좌표값 구하는 순서도
Fig. 8. Flow Chart to obtain Head Point Coordinate Value.

5. 손뼉 치기 파악

PoseNet을 활용하여 사람의 운동 동작 중 가장 단순한 편에 속하는 ‘손뼉 치기’ 동작을 인식해 보았다. 사람이 카메라 앞에서 손뼉을 치면 PoseNet을 통해 얻은 관절들의 위치 정보들을 가공하여 손뼉 친 횟수를 세고 이를 화면에 표시해 주는 구조로 아래의 그림 9와 같이 알고리즘으로 구성된다.

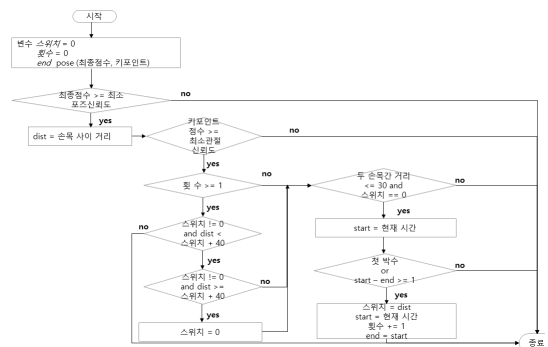


그림 9. 손뼉 횟수 카운트 순서도
Fig. 9. Applause Count Flowchart.

IV. 연구 결과 및 의미

첫 번째 실험을 통해 PCK@0.2 값들을 구체적으로 살펴보면, ResNet50 버전 모델이 MobileNet 버전 모델보다 전체적으로 10% 이상 높은 성능을 보인다. 레이어가 많아 모델 자체가 무겁고 느린 대신에(i5-7200U CPU, Ubuntu 16.04에서 MobileNet 버전은 모델 로딩에 14.8초, 이미지 1,000장 처리에 70.1초가 소요. ResNet50 버전은 모델 로딩에 43.2초, 이미지 1,000장 처리에 238.7초 소요) 정확도는 상대적으로 높은 ResNet 구조 자체의 특성이 잘 보여지는 결과값이다.

표 5. PoseNet의 ResNet50 모델과 MobileNet 모델의 성능 비교 (LSP, MPII PCK@0.2)

Table5. Performance Comparison of PoseNet's ResNet50 and MobileNet Models. (LSP,MPII PCK@0.2)

Method	PoseNet MobileNet	PoseNet ResNet50	Wei et al., CVPR'16	Newell et al., ECCV'16	Insafutdinov et al., ECCV'16
Head	33.45	39	97.8	98.2	97.45
Sho	74.2	85.2	92.5	96.3	92.7
Elb	63.1	77.65	87.0	91.2	87.55
Wri	58.3	71.35	83.9	87.1	84.45
Hip	76.1	83.8	91.5	90.1	91.45
Knee	70.3	83.25	90.8	87.4	89.9
Ank	59.4	80.65	89.9	83.6	87.2
PCK Mean	62.12	74.41	90.5	90.9	90.1
AUC	36.53	46.9	65.4	65.2	66.08

표 5를 근거로 관절 부분에 따른 인식 정확도의 양상을 보면 다른 모델들과 크게 다르지 않다. 다른 모델들이 상대적으로 높은 정확도로 인식한 어깨나 골반 부분은 PoseNet도 다른 부분들에 비해 잘 인식했다. 무릎과 발목 부분에 대한 성능은 다른 모델과 거의 차이 나지 않거나 조금 높다(1% 안팎). 하지만 표 5와 그림 10의 눈, 코, 귀 포인트를 목과 머리 끝 포인트로 변환하는 연구방법으로 나온 결과값은 정식으로 예측되는 부분이 아니기 때문에 PoseNet이 PCK의 Head 부분에서 다른 모델보다 낮은 점수에는 큰 의미를 둘 필요가 없다고 본다. 그래도 간단히 살펴보자면, 목 포인트는 왼쪽과 오른쪽 어깨 포인트의 중점으로 지정했다. 머리 끝 포인트를 정하기 위해서는 눈과 귀, 코 포인트를 이용한다. 머리 끝 포인트는 얼굴의 다섯 포인트에 대해 예측된 값을 최대한 활용하여 계산하였는데, 이미지에서 사람이 앞/뒤를 보고 있을 경우, 왼쪽을 보고 있을 경우, 오른쪽을 보고 있을 경우로 나누어 계산해 주었다.

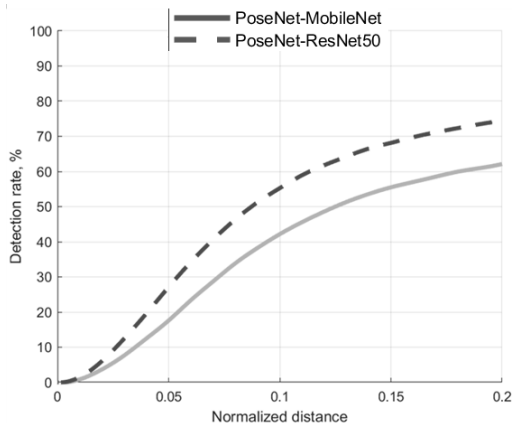


그림 10. PoseNet의 ResNet50 모델과 MobileNet 모델의 성능 비교 (LSP, PCK@0.2)

Fig. 10. Performance Comparison of PoseNet's ResNet50 and MobileNet Models. (LSP, PCK@0.2)

두 번째 손뽁 동작 구현 방식으로 모든 코드는 PoseNet의 데모 코드 파일 camera.js에 추가된다. 인식한 포즈의 키포인트와 그에 따른 골격을 화면에 그려주는 방식으로, 필요한 변수는 미리 선언해 주었음을 전제한다.

손뽁치는 동작을 인식하는 방법을 요약한 세 가지 키워드는 그림 11과 같이 손목 사이의 거리, 스위치 변수, 시간차 계산이다. 손목 간 거리 정보만 가지고 손뽁 친 횟수를 세려고 할 때에 손바닥을 맞댄 채 계속 가만히 있어도 손뽁 친 횟수가 증가하는 문제가 발생했다. 그래서 손바닥이 한번 맞대 진 후에 일정 거리 이상 떨어졌는지 확인하는 스위치 변수를 두는 동시에, 손바닥이 포개지는

순간을 포착한 후 손뽁 횟수를 증가시키는 그 사이의 시간차를 계산하도록 하여 이중으로 이런 문제가 발생하지 않도록 하였다.

동작 인식으로 PoseNet을 활용하여 사람의 운동 동작 중 가장 단순한 편에 속하는 '손뽁 치기' 동작을 인식 결과로서 그림 12와 같이 화면 하단에 손뽁 친 횟수가 표시된다.

```
if (keypoints[7].score >= minPartConfidence && keypoints[8].score >= minPartConfidence && keypoints[9].score >= minPartConfidence && keypoints[10].score >= minPartConfidence) {
    if(count >= 1) {
        if (_switch != 0 && dist < _switch + 40)
            return;
        if (_switch != 0 && dist >= _switch + 40)
            _switch = 0;
    }

    if(dist <= 80 && _switch == 0) {
        start = new Date();
        var audio = document.getElementById("myAudio");
        if (count == 0 || (start - end >= 1000)) {
            _switch = dist;
            count = count + 1;
            document.getElementById("clap").innerHTML = '박수 ' + count;
            myAudio.play();
            myAudio.currentTime = 0;
            audio.play();
            end = start;
        }
    }
}
```

그림 11. 손뽁 동작 인식 소스 예

Fig. 11. Applause Gesture Recognition Source Example.

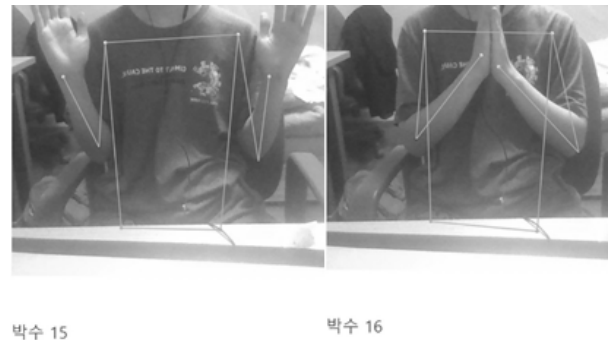


그림 12. 손뽁 동작 인식 화면 예

Fig. 12. Applause Gesture Recognition Screen Example.

V. 결 론

본 논문에서는 LSP 데이터셋을 가지고 PoseNet으로 자세 추정 학습을 해 보았으며, 그 결과 눈, 코, 귀 키 포인트에 대한 결과가 타 모델과 해석의 차이로 다소 차이가 나지만, 나머지 부위는 성능이 떨어지지 않음을 확인할 수 있었다. 또한, 이 결과값을 이용하여 손뽁 치기라는 하나의 동작에 따른 운동량 분석하는 프로그램을 구현하였다.

PoseNet은 플랫폼에서 제공하는 하나의 서비스로서 활용하기가 수월하다. 자바스크립트 프로그램 코드 몇 줄 작성으로 인해 이 기능을 사용할 수 있으며, 또한 카메라 기능이 지원되는 디바이스가 있다면 어디서든 브라우저로 통해 실시간 자세 추정을 통한 동작 분석을 통해

굉장히 단순한 방법이지만 사용자 효과는 생각보다 클 것으로 예상된다.

향후 PoseNet은 헬스케어 플랫폼의 많은 기능 중 하나로 서비스되는 스마트 헬스케어시스템을 개발하고자 하는 우리의 목적에 상당히 부합하는 모델이라는 판단된다.

REFERENCES

- [1] I. J. LEE, "Current Status and Direction of Business Promotion in ICT Industry for Mobile Health Market", Weekly Technology Trend, NIPA, pp. 1-8, July 2011.
- [2] S. H. Lee, "The Future of Smart Healthcare Devices", ICT Promotion Center, vol. 17, pp. 46-47, 2015.
- [3] K. E. Ko and K. B. Shim, "Trend of Object Recognition and Detection Technology Using Deep Learning", Journal of Control Robotics and Systems, vol. 23, no. 3, pp. 17-24, June 2017.
- [4] H. I. Ku, "Artificial Intelligence and Deep Running Trends", The world of electricity, vol. 67, no. 7, pp. 7-12, 2018.
- [5] L. Jin, Z. Chen, et al, "Object Detection Free Instance Segmentation With Labeling Transformations", Computer Vision and Pattern Recognition(CVPR), Nov. 2016.
- [6] J. Tompson, A. Jain, et al, "Joint Training of a Convolutional Network and a Graphical Model for Human Pose Estimation" In NIPS, 2014.
- [7] A. Jain, J. Tompson et al, "Learning Human Pose Estimation Features with Convolutional Networks" In ICLR, 2014.
- [8] A. Krizhevsky, I. Sutskever et al, "ImageNet Classification with Deep Convolutional Neural Networks", In ILSVRC, 2012.
- [9] C. Szegedy and W. Liu et al, "Going Deeper With Convolutions" IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2015.
- [10] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks For Large-Scale Image Recognition", Conference at ICLR, 2015.
- [11] K. He. and X. Zhang, et al, "Deep Residual Learning for Image Recognition", Computer Vision and Pattern Recognition(CVPR), 2015.
- [12] M. Sandler and A. Howard, et al, "MobileNetV2: Inverted Residuals and Linear Bottlenecks", The IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 2018.
- [13] Real-time Human Pose Estimation in the Browser with TensorFlow.js, <https://medium.com/tensorflow/>
- [14] G. Papandreou, T. Zhu, et al, "PersonLab: Person Pose Estimation and Instance Segmentation with a Bottom-Up, Part-Based, Geometric Embedding Model", 2017.
- [15] G. Papandreou, T. Zhu, et al, "Towards Accurate Multi-person Pose Estimation in the Wild" IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.
- [16] S. H. Lee and S. S. Yoo, "Mobile Healthcare Application Status and Forecast. Information and Communication Broadcasting Policy", vol. 26, no 17, pp. 1-23, 2014.
- [17] T. Y. HA and H. LEE, "Presenting Direction for the Implementation of Personal Movement Trainer through Artificial Intelligence based Behavior Recognition", Journal of the Korea Convergence Society, vol. 10, no. 6, pp. 235-242, 2019.
- [18] S. J. Lee, "Electronics and Telecommunications Trends", ETRI vol. 33, no. 4, pp. 33-42, Aug. 2018.
- [19] J. M. Cho, et al, "Object Recognition and Pose Estimation Based on Deep Learning for Visual Servoing", Journal of Korea Robotics Society, vol. 14, no. 1, pp. 1-7, 2019.

저 자 소 개



하 태 용(학생회원)
1996년 동아대학교 수학과 학사 졸업.
2013년 한성대학교 지식서비스
건설링학과 석사 졸업.
2017년~현재 한성대학교 스마트융
합건설링학과 박사 과정.
<주관심분야: 스마트 팩토리, 인공지능, 헬스케어,
보건의료>



이 후 진(평생회원)
1997년 서울대학교 전기공학부
학사 졸업.
2002년 The University of Texas
at Austin ECE 석사 졸업.
2007년 The University of Texas
at Austin ECE 박사 졸업.
2009년~현재 한성대학교 IT융합공학부 및 스마트
융합건설링학과 교수
<주관심분야: 통신 및 네트워크, 멀티미디어 신호
처리, 사이버보안>