JKSCI

한국컴퓨터정보학회논문지
Journal of The Korea Society of Computer and Information
Vol. 30 No. 1, pp. 53-63, January 2025
https://doi.org/10.9708/jksci.2025.30.01.053

A Practical Approach to Fall Detection: Lightweight AI for Low-Cost Implementation

Minhyung Ryu*, Jaegoo Shim**

*Researcher, R&D Center, NEARNETWORKS Co., Ltd., Daegu, Korea
**Professor, Dart. of Radiological Science, Daegu Health College, Daegu, Korea

[Abstract]

This paper proposes the development of a lightweight AI-based fall detection system that can operate efficiently in low-spec hardware environments. Existing fall detection research has often relied on high-performance GPU environments or focused solely on detection accuracy, without adequately addressing processing speed and hardware limitations. To overcome these challenges, this study utilizes MediaPipe, a GPU-independent framework, combined with data analysis-based machine learning models, to design a system that requires minimal computational resources while ensuring fast and accurate fall detection. Experimental results demonstrate that the proposed model outperforms GPU-based models in terms of training and inference speed, while also delivering competitive detection performance. This research presents a promising approach to reducing dependency on high-cost hardware, thereby maximizing the practicality of fall detection systems in real-world scenarios.

▸ **Key words**: fall detection, lightweight AI, low-spec hardware, machine vision, health care

[요 약]

본 논문은 저사양 하드웨어 환경에서도 효율적으로 동작할 수 있는 경량 AI 기반 낙상 탐지 시스템을 개발하고 이를 제안한다. 기존의 낙상 탐지 연구는 주로 고성능 GPU 환경에 의존하거나 탐지 정확도에만 중점을 두어 처리 속도와 하드웨어 제약 문제를 충분히 고려하지 못했다. 이에 본 연구에서는 GPU에서 자유로운 미디어파이프와 데이터 분석기반의 머신러닝 모델을 활용하여, 낮은 연산 자원을 요구하면서도 신속하고 정확한 낙상 탐지가 가능한 시스템을 제안하였다. 제안된 시스템은 병원, 요양원 등 다양한 실내 환경에 설치되어 낙상 사고를 실시간으로 탐지하고, 관리자에게 즉시 알림을 전송함으로써 골든타임 내 대응이 가능하도록 한다. 실험 결과, 제안된 모델은 GPU를 사용하는 모델과 비교하여 학습 및 추론 속도에서 우위를 보였으며, 탐지 성능 측면에서도 경쟁력 있는 결과를 나타냈다. 이러한 연구는 고비용 하드웨어 의존성을 줄이고 실제 환경에서의 실용성을 극대화한 낙상 탐지 시스템의 새로운 가능성을 제시한다.

▸ **주제어**: 낙상탐지, 경량 AI, 저사양 하드웨어, 머신 비전, 헬스케어

• First Author: Minhyung Ryu, Corresponding Author: Jaegoo Shim
*Minhyung Ryu (tkj06039@aaf.co.kr), R&D Center, NEARNETWORKS Co., Ltd.
**Jaegoo Shim (sjg2428@dhc.ac.kr), Dart. of Radiological Science, Daegu Health College
• Received: 2024. 12. 06, Revised: 2024. 12. 23, Accepted: 2024. 12. 27.

Copyright © 2025 The Korea Society of Computer and Information
http://www.ksci.re.kr pISSN:1598-849X | eISSN:2383-9945
www.kci.go.kr

I. Introduction

통계청이 발표한 '세계와 한국의 인구현황 및 전망'에 따르면 세계 인구 중 고령인구 구성비는 2022년 9.8%에서 2070년 20.1%로 증가하는 반면, 한국의 인구 중 고령인구 구성비는 2022년 17.5%에서 2070년 46.4%로 증가하는 것으로 전망됐다[1]. 이는 한국이 세계 인구에 비해 압도적으로 고령 인구 구성비가 높으며, 미래에 빠른 속도로 노인의 인구수가 증가할 것이라고 예견된다. 노인 인구의 증가에 따라 사회적 문제로 대두되고 있는 또 하나의 문제는 낙상이며, 낙상은 '사람이 실수로 바닥 또는 다른 낮은 수준으로 떨어지는 사건'으로서 노인은 나이가 들면서 근력과 자세균형 능력이 감소하게 되어 낙상에 쉽게 노출될 수 있다[2]. 실내외를 막론하고 낙상사고를 당해 병원을 찾는 환자들이 증가하고 있다. 특히 골밀도가 낮은 노인의 경우 낙상 사고 발생 시, 초동 대처가 미흡하거나 도움을 요청할 수 없는 상황이 발생하게 되면 생명을 잃게 되는 위험한 상황을 초래하게 된다. 또한, 노인층뿐만 아니라 장년층에서도 뇌출혈 등으로 쓰러지는 경우가 다수 있으나, 발견이 늦어져 골든타임을 놓치고 사망하는 사례가 빈번하다[3-4]. 이에 따라, 낙상 사고를 검출하기 위한 연구가 활발히 진행되고 있다. 하지만 이와 같은 기술은 감지 센서를 착용해야만 하는 번거로운 문제점이 있다. 따라서 낙상을 탐지하고 위험 상황을 외부에 자동으로 알릴 수 있는 낙상 탐지 장치가 필요한 실정이다.

이에 따라 낙상을 탐지하기 위한 연구들이 현재까지 활발히 진행되고 있으며 최근에는 AI 기술을 활용한 형태의 낙상 탐지 기법이 주목받고 있다[5-7]. 그러나 대부분의 연구가 탐지 정확도에만 중점을 두고 있어 신속한 처리속도를 고려한 연구도 고려될 필요성이 있다. 또한 기존연구의 처리속도는 고사양의 GPU 서버를 활용하기 때문에 사양이 떨어지는 저사양 하드웨어에서는 AI 모델의 탑재가 어려우며, GPU 서버와의 통신 오버헤드 역시 처리속도에 큰 문제를 영향을 준다. 이러한 상황에서는 저사양의 하드웨어에서 동작 가능한 경량 AI 모델이 요구되며 이를 활용한 연구사례들이 최근 발표되었다[8-10]. 또 다른 접근법으로 엣지컴퓨팅(edge computing)을 이용한 방법이 고려될 수 있다[11-13].

본 논문에서는 저사양 하드웨어에서도 구동 가능한 경량 AI 모델을 개발하고, 이를 활용한 빠른 처리속도와 학습 시간을 갖춘 AI 머신비전 기반 낙상 탐지 시스템을 제안한다. 제안된 시스템은 병원이나 요양원과 같은 환경에 설치되어, 신체 능력이 저하된 환자나 노인의 낙상을 실시간으로 탐지하고, 이를 관리자에게 알리거나 경고 알람을 발생시켜 긴급 상황에 신속히 대처할 수 있도록 설계되었다. 본 논문의 기여는 아래와 같다.

- 경량 AI 모델 설계를 통해 저사양 하드웨어에서도 동작 가능하도록 구현함으로써, 기존 고사양 GPU에 의존하는 시스템의 한계를 극복함.
- 낙상 상황을 신속히 탐지하고 경고할 수 있도록 처리속도를 최적화한 AI 머신비전 시스템 개발함
- 병원, 요양원 등 다양한 실내 환경에서 쉽게 설치 및 운영 가능한 시스템으로 활용성을 극대화하여 실용적 설치 환경 제공함
- 낙상 탐지시 관리자 알림 기능을 통해 골든타임 내 응급 대응 가능성을 높여 노인 및 환자들의 안정성을 향상시킴.

이하 본 논문의 구성은 다음과 같다. 2장에서 배경 및 관련 연구를 소개하고 3장에서 머신 비전 기반 낙상탐지 시스템을 제안한다. 그리고 4장에서 제안 모델에 대한 실험 평가를 통해 성능을 검증하고 마지막 5장에서 결론을 맺는다.

II. Background and Related Works

1. Background

1.1 MediaPipe

미디어파이프(MediaPipe)는 Google이 개발한 크로스 플랫폼의 오픈소스 라이브러리로, 효율적이고 경량화된 머신러닝 기반 애플리케이션 개발을 지원한다. 이 프레임워크는 실시간 데이터 스트리밍과 고성능 처리를 가능하게 하는 그래프 기반 아키텍처를 제공하며, 특히 비디오, 오디오, 이미지 등의 멀티모달 데이터 처리를 효율적으로 수행한다[14]. 미디어파이프는 다양한 플랫폼(CPU, GPU, 모바일 디바이스)에서 실행 가능하며, TensorFlow Lite와 같은 경량화된 모델을 활용하여 자원 제한적인 환경에서도 높은 성능을 보장한다.

미디어파이프의 핵심은 모듈화된 처리 그래프로, 이를 통해 데이터 입력에서부터 머신러닝 모델 실행 및 출력 생성까지의 모든 과정이 일관되고 유연하게 관리된다. 이러한 아키텍처는 낙상 탐지와 같은 신체 움직임 분석 애플리케이션에 적합하며, 객체 탐지, 얼굴 인식, 자세 랜드마크 탐지 등 다양한 기능을 제공한다.

미디어파이프는 경량성과 실시간 처리를 특징으로 하며, IoT 디바이스와 같은 하드웨어 자원이 제한된 환경에서도 실행 가능하도록 설계되었다. 특히, Pose Detection 모델은 신체의 주요 관절에 대한 2D 또는 3D 좌표 데이터를 추출하여 사람의 신체 움직임을 정확하게 추적하고 분석하는 데 매우 유용하다.

이러한 경량화된 접근은 고사양 GPU 없이도 실시간 데이터 처리가 가능하게 하며, 병원, 요양원 등에서의 실용적 응용을 가능하게 한다[15].

또한, 미디어파이프는 확장성과 호환성을 강조하며, Python, C++, JavaScript와 같은 다양한 프로그래밍 언어와 호환된다. 이를 통해, 연구자와 개발자는 다양한 환경에서 AI 응용을 손쉽게 개발하고 배포할 수 있다. 특히, 엣지 컴퓨팅 환경에서의 운영은 통신 지연 및 오버헤드를 최소화하며, 데이터의 실시간 처리 능력을 극대화할 수 있다.

따라서, 미디어파이프는 경량화된 AI 모델과 실시간 처리 능력을 통해, 제한된 하드웨어 환경에서 효과적으로 작동할 수 있는 최적의 프레임워크로 자리 잡고 있다. 낙상 탐지와 같은 응용에서도 미디어파이프는 포즈 추정 및 키포인트 기반 데이터 분석을 통해, 정확성과 효율성을 모두 제공할 수 있는 강력한 도구로 활용될 수 있다.

1.2 YOLO-pose

YOLO(You Only Look Once)는 딥러닝 기반 객체 탐지(Object Detection) 알고리즘으로, 하나의 신경망을 통해 이미지 내 객체의 위치와 클래스 정보를 동시에 예측하는 통합된 접근 방식을 제안한다[16]. 기존의 객체 탐지 알고리즘은 객체의 위치와 클래스 분류를 별도로 처리하며, 슬라이딩 윈도우나 지역 제안 방식(region proposal)을 사용하는 복잡한 처리 구조로 인해 연산 비용이 높고 속도가 느린 단점이 있었다. 이에 비해 YOLO는 단일 패스(single pass)로 이미지를 처리하여, 탐지 속도를 급격히 향상시키고 실시간 객체 탐지에 적합한 구조를 제공한다[17].

YOLO의 대표적인 장점은 첫째로 실시간 탐지가 가능하다는 점이다. 단일 신경망 구조로 설계되어, 고속 처리가 가능하며 초당 다수의 프레임을 처리할 수 있다. 이는 자율주행, 감시 시스템 등 실시간 응용에 적합하다. 둘째, 단순한 구조와 학습 효율성을 제공한다. YOLO는 단일 모델을 통해 이미지 내 모든 객체를 처리하므로, 복잡한 전처리나 추가 단계가 요구되지 않는다. 마지막으로 전체 이미지를 고려한 탐지가 가능하다. 이는 지역 제안 방식보다 적은 false positive(잘못된 탐지)를 발생시키는 데 기여한다.

그러나 YOLO는 작은 객체 탐지에 대한 성능 한계가 있다. 그리드 기반 접근법으로 인해, 작은 객체가 그리드 셀에 포함되지 않거나 다른 객체와 합쳐질 경우 탐지 성능이 저하될 수 있다[18]. 그리고 복잡한 배경에서의 오탐지 가능성이 있다. 단일 패스 방식의 탐지는 세밀한 지역 정보를 반영하지 못해 복잡한 이미지 환경에서 정확성이 감소할 수 있다. 셋째, 고해상도 이미지 처리의 제약이 있을 수 있다. YOLO는 고해상도 이미지를 처리할 때 메모리와 연산 속도 면에서 제약이 발생할 수 있어, 하드웨어 성능에 따라 처리 효율이 영향을 받을 수 있다. 즉, 고사양 GPU가 탑재된 하드웨어가 요구된다.

2. Related works

2.1 Fall detection system using PoseNet and GRU

강윤규 외(2021)는 딥러닝 기술을 활용한 낙상 탐지 시스템을 제안하며, PoseNet과 GRU(Gated Recurrent Unit)를 결합하여 인체 관절 데이터(Skeleton Keypoints)를 기반으로 낙상을 탐지하는 기법을 연구하였다[7]. 이 연구에서는 저비용 2D RGB 카메라를 활용하여 PoseNet으로 인체의 17개 주요 관절 키포인트를 추출한 후, GRU를 통해 시간적 변화를 분석하여 낙상 여부를 판단하였다. PoseNet은 영상으로부터 신체 관절 좌표를 추출하는 CNN 기반 기술로, 연구에서는 이를 통해 낙상 시 발생하는 머리와 어깨의 위치 변화, 하강 속도, 신체 폭 대 높이 비율 등의 특징을 계산하여 정확도를 높였다. 특히 GRU는 시간적 데이터 처리에 강점을 가진 딥러닝 기법으로, 학습 시간 단축과 메모리 효율성을 통해 낙상 여부를 높은 정확도로 예측하였다. 실험결과, 낙상 탐지 정확도 99.8%를 달성하였으며, 신체의 자세 변화와 동작 패턴을 정밀히 분석하여 기존의 단순 센서 기반 탐지 시스템보다 높은 성능을 보였다. 또한, 비전 기반 접근 방식을 활용함으로써 센서 착용의 불편함을 제거하고, 기존 낙상 탐지 기술의 실용성을 강화하였다.

2.2 Fall detection system using YOLO

김정수의 연구에서는 지하공동구 작업자들의 낙상을 탐지하기 위해 딥러닝 기반의 객체 탐지 기술을 활용한 연구를 수행하였다[6]. 이 연구는 YOLOv5 모델을 기반으로 세 가지 접근법을 비교하며, 낙상 탐지 성능을 평가하였다. 첫 번째 접근은 MS-COCO 데이터셋으로 사전 학습된 YOLOv5 모델에 낙상 판단 규칙을 적용한 방식이고, 두 번째 접근은 지하공동구 데이터를 추가 학습한 후 동일한 규칙을 적용한 방식이다. 마지막 세 번째 접근은 작업자의 정상 상태와 낙상 상태를 구분할 수 있도록 YOLOv5를 학

www.kci.go.kr

습하였다.

이 연구의 세 번째 접근법은 낙상 자세를 학습하여 F1 점수 0.867을 기록하며, 다른 두 접근법에 비해 가장 높은 낙상 탐지 성능을 보였다. 이는 낙상을 직접 학습한 딥러닝 모델이 규칙 기반 모델보다 우수한 성능을 제공한다는 점을 입증한다. 또한, 연구에서는 지하 공동구의 저조도 환경과 작업자의 다양한 자세를 반영한 데이터 셋을 구축하여 실험에 활용하였다.

결론적으로, 이 연구는 딥러닝 기반 낙상 탐지 모델의 성능 비교를 통해 낙상 자세를 학습한 모델의 유효성을 입증하고, 규칙 기반 접근법의 한계를 보완할 수 있는 데이터 중심 학습 접근법의 중요성을 강조하였다.

2.3 Fall detection system using Bi-LSTM

신병근의 연구에서는 MS사의 키넥트 RGBD 카메라로 추출한 사람의 키포인트를 활용하고, 2-Stacked Bi-LSTM 모델을 결합하여 낙상을 탐지하는 시스템을 제안하였다[19]. 이 연구에서는 기존의 OpenPose 또는 PoseNet과 같은 딥러닝 모델로 신체 관절 좌표를 추출한 후, 순환신경망 모델인 LSTM이나 GRU를 활용하여 낙상을 탐지하는 접근법을 개선하고자 하였다. 이를 위해 키넥트 카메라를 사용해 관절 정보를 직접 추출하고, 중심 관절의 움직임 가속도와 중심 관절과 바닥 사이의 거리 데이터를 시계열 특징으로 활용하였다. 2-Stacked Bi-LSTM 모델은 양방향(Bi-directional) LSTM을 층층이 구성하여 학습하며, 시간적 전후 관계를 모두 반영하여 낙상 여부를 판단한다. 연구 결과, 기존 GRU 및 단층 LSTM 모델 대비 탐지 정확도, 정밀도, 재현율, F1 Score 등 모든 지표에서 성능이 개선되었음을 검증했다. 특히, 키넥트 카메라에서 제공하는 데이터의 25개 관절 중 신뢰도가 높은 주요 관절만을 선별하여 중심 관절 정보를 계산함으로써, 불필요한 연산을 줄이고 관절 검출의 불확실성을 감소시킬 수 있음을 보였다.

이 연구는 딥러닝 기반 낙상 탐지 기술이 비전 데이터를 효율적으로 활용할 수 있음을 입증했으며, 스마트홈 환경에서 낙상과 같은 응급 상황을 실시간으로 탐지할 수 있는 가능성을 제시하였다.

그러나, 앞서 언급한 선행 연구들은 CNN, PoseNet, YOLO, Bi-LSTM 등 고성능 GPU가 요구되는 딥러닝 알고리즘을 활용하여 높은 성능을 달성했지만, 경량 AI 모델이나 저사양 하드웨어 환경에서의 실행 가능성은 다루지 않았다. 현실적인 설치 비용과 제품 상용화를 고려하면, 이러한 한계를 극복하고 최적화할 필요성이 있다.

III. The Proposed System

본 장에서는 미디어파이프와 머신러닝을 기반으로 한 낙상 탐지 시스템의 개발 과정을 설명한다. Fig 1은 제안된 시스템의 개요를 흐름도로 나타낸 것이다. 먼저, 카메라 또는 CCTV를 사용하여 실시간으로 프레임을 촬영한다. 이후, 미디어파이프 객체 탐지 모델을 활용해 각 프레임에서 사람을 탐지하고, 탐지된 각 사람의 영역을 새로운 프레임으로 추출한다. 이는 미디어파이프가 한 프레임에서 1명의 신체 골격 데이터만 추출할 수 있기 때문에 필요한 과정이다. 이후, 추출된 프레임으로부터 사람의 관절 데이터를 얻는다.

각 사람의 관절 데이터 추출 이후, SORT 알고리즘을 이용하여 탐지된 사람들을 추적하기 위한 고유 번호를 할당한다. 프레임에 등장한 모든 사람을 순서대로 사람의 관

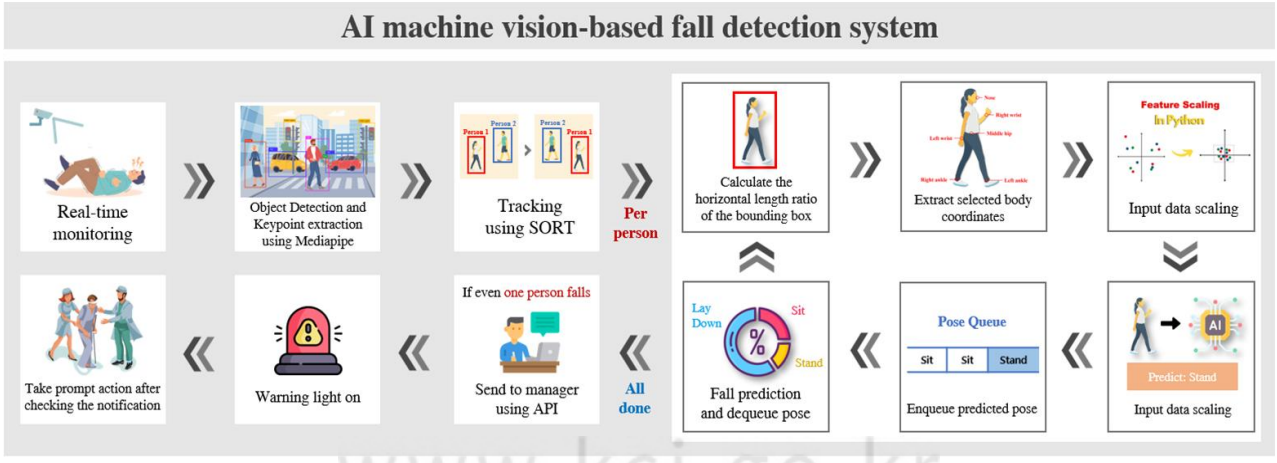


Fig. 1. System Overview

www.kci.go.kr

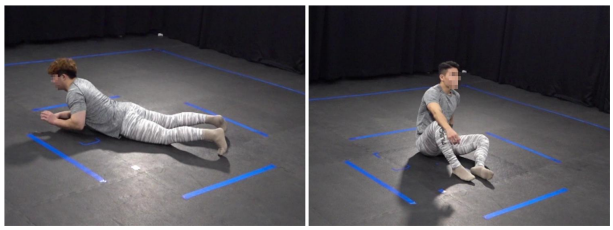


Fig. 2. Example of collected data

절 데이터를 전처리하고 사전학습된 분류 모델(Pre-trained Model)을 통해서 자세를 탐지한다. 각 프레임별 자세 탐지가 가능해지면, 과거로부터 지속적으로 탐지되는 각 사람의 자세의 시계열 패턴을 통해서 낙상 여부를 식별한다. 영상 속 사람들 중 한 명이라도 낙상으로 식별되면, 해당 상황에 대한 정보가 서버로 전송되고, 관리자에게 관련 정보를 전송한다. 이 정보에는 낙상 발생 시각, 영상에 등장한 인원수, 낙상한 인원수, 그리고 포착된 이미지 등이 포함된다. 이를 통해 관리자는 낙상 상황을 신속히 확인하고, 포착된 이미지를 바탕으로 적절한 조치를 취하여 현장의 문제를 빠르게 해결할 수 있다.

1. Pose detection training dataset construction

본 절에서는 자세 탐지 모델을 구현하기 위한 학습데이터 구축을 위해서 데이터 수집과 분석과정을 설명한다. 데이터는 AI Hub의 사람 인체/자세 3D 데이터에서 자세별로 각 20,000장의 이미지를 수집하였고 이를 기반으로 데이터 분류 및 분석을 수행하였다. Fig 2는 수집된 이미지의 예시를 보여준다. 수집된 데이터는 선 자세(stand), 앉은 자세(sit), 그리고 누운 자세(Lay Down)로 분류되었으며, 미디어파이프를 통해 각 자세의 신체 관절 좌표를 추출하였다. 추출된 데이터는 JSON 형식으로 저장되었고, 이후 파이썬의 시각화 도구를 활용하여 분석을 진행하였다.

자세 탐지를 위한 주요 특징을 찾기 위해 분석해본 결과, Fig 3에 보이는 것처럼, 같이 선 자세, 앉은 자세, 그리고 누운 자세의 바운딩 박스(bounding box)에 대해서 유의미한 차이점을 보이고 있다. 각 자세별로 바운딩 박

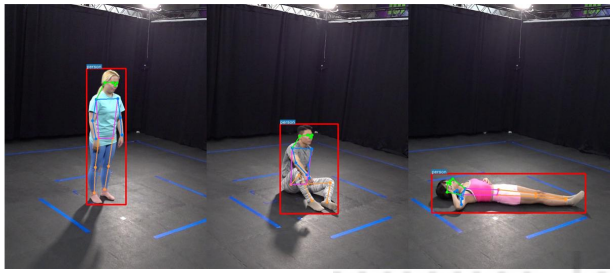


Fig. 3. Bounding box for standing, sitting and lying images

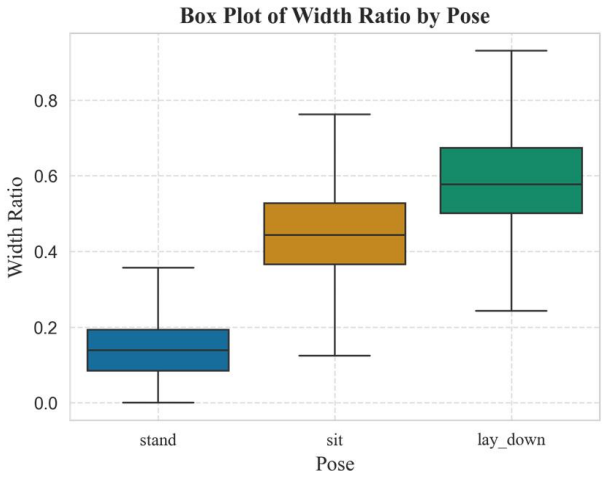


Fig. 4. Distribution of bounding box width ratio distribution by pose

스의 너비(width), 높이(height)의 비율에 대해 명확한 패턴이 확인된다. 선 자세의 경우 바운딩 박스의 너비의 비율이 낮고 누운 자세에서는 높게 나타남을 알 수 있다.

이러한 현상을 검토하기 위해, 모든 이미지에 대해 높이-너비 비율에 대한 통계를 확인하였다. Fig 4는 각 자세별로 바운딩 박스의 너비 비율을 시각화한 것이다. 시각적 분석과 유사하게, 선 자세에서는 낮은 분포가, 누운 자세에서는 높은 분포가 나타남을 알 수 있다. 앉은 자세의 경우, 두 그룹의 분포 사이에 위치함을 확인할 수 있다. 이러한 박스 플롯 분포를 통해 바운딩 박스의 너비 비율로 선 자세와 누운 자세는 명확히 구분할 수 있으나, 앉은 자세는 구별이 어려운 것으로 보인다. 또한, 카메라 각도의 영향을 크게 받을 것으로 판단된다.

바운딩 박스 외에 다른 주요 정보를 추출하기 위해 미디어파이프를 사용하여 관절 좌표를 추출하고, 자세별로 분포를 확인하였다. Fig 5는 선택된 관절 데이터의 자세별 2차원 좌표 분포를 시각화한 박스 플롯 중 자세 별로 유의미한 차이를 보이는 주요 정보만을 추출한 결과를 나타낸 것이다. 시각적으로 유의미한 변수는 코, 좌우 손목, 엉덩이, 양 발목 등 총 6가지인 것으로 확인되었다.

따라서, 선정된 관절 좌표 6개 항목과 바운딩 박스의 너비 비율을 활용하여 학습 데이터셋을 구성한 후, 이를 기반으로 SVM 및 Random Forest와 같은 경량 머신러닝 모델을 학습시켜 자세 탐지 모델을 구현한다.

2. Fall detection algorithm

본 절에서는 실시간으로 사람의 자세를 탐지하고, 자세 변화에 따라 낙상 여부를 판단하는 알고리즘에 대해 서술한다. 여기에서 설명하는 Algorithm 1은 비디오 스트림의

www.kci.go.kr

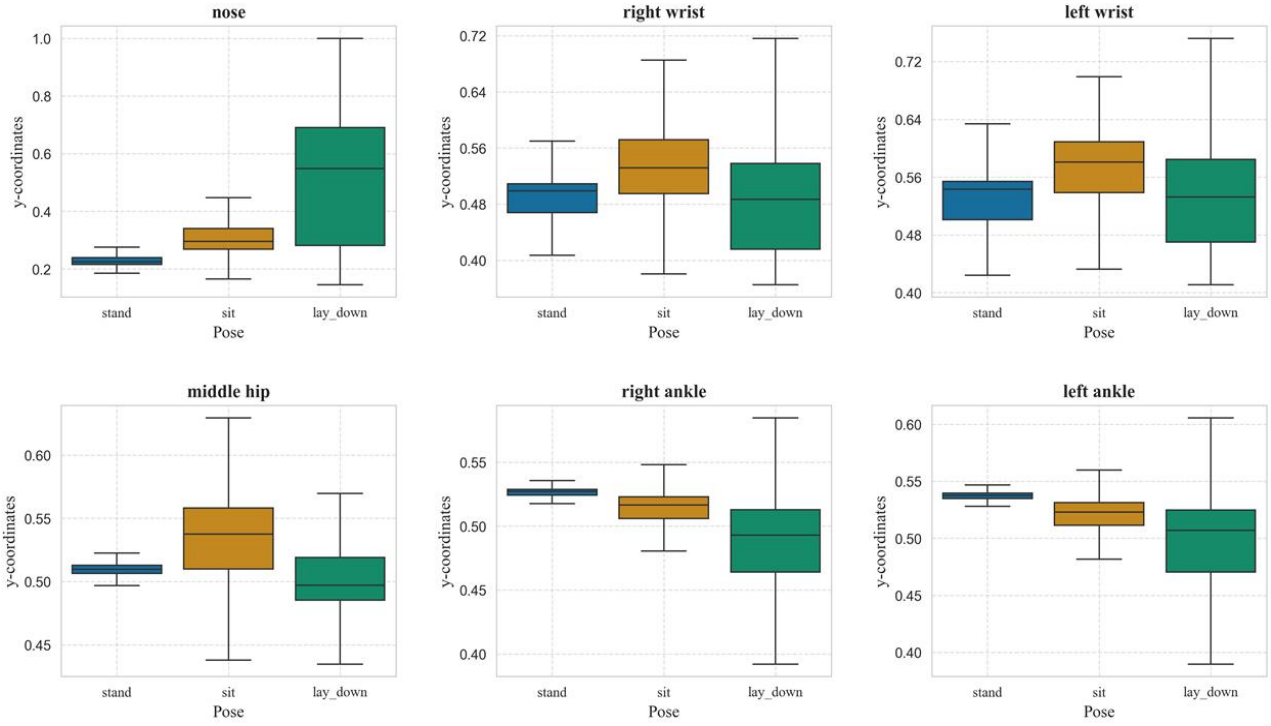


Fig. 5. Distribution of key features

각 프레임에서 지속적으로 탐지되는 자세 변화를 분석하여 낙상을 탐지하는 프로세스를 포함한다. 효율적인 설명을 위해 이미지 전처리, 데이터 스케일링과 같은 세부적인 전처리 과정은 생략하였으나, 실제 구현에서는 이러한 전처리가 정확도를 높이는 데 중요한 역할을 한다.

Algorithm 1 : Detecting Fall Down

```
1  input threshold
2  stanceQueue(200)
3  while isLastFrame() != true :
4     currentFrame = GetCurrentFrame()
5     hFrames = GetHumanFrame(currentFrame)
6     for frame in hFrames
7        keypoints, width = GetFeatures(frame)
8        input = merge(keypoints, width)
9        stance = model.predict(input);
10       stanceQueue.add(stance)
11       laydownRate = GetLayDownRate(stanceQueue)
12       if laydownRate >= threshold :
13          Warnning()
```

threshold는 200개의 프레임 내에서 누운 자세의 비율이 특정 임계값 이상일 경우 이를 낙상으로 판단하기 위한 기준으로 사용된다. 사용자는 초기에 이 값을 설정하거나 중간에 수정할 수 있다(1번 줄). stanceQueue는 최대 200개의 자세 데이터를 저장할 수 있는 큐로, 새 데이터가 추가될 경우 큐가 가득 차 있으면 가장 오래된 데이터를 자동으로 삭제하여 항상 최신 데이터를 유지한다(2번 줄). 이를 통해 메모리 효율성을 유지하며, 실시간으로 자세 변화를 분석할 수 있는 구조를 제공한다.

알고리즘의 본격적인 처리는 3번 줄부터 시작된다. 현재 프레임(currentFrame)을 기준으로, 미디어파이프 객체 탐지 기능을 활용해 프레임 내의 개별 사람을 탐지하고 이를 개별 영역(hFrames)으로 분할한다(4-5번 줄). 각 영역에 대해 자세 탐지를 수행하며, 이 과정에서 자세의 정확도를 높이기 위해 바운딩 박스의 너비(width)와 6가지 주요 관절 좌표를 추출한다. 이 정보는 한 사람의 자세를 설명하는 데이터로 취합되어 모델의 입력값(input)으로 사용된다.

이 입력 데이터는 사전에 학습된 자세 탐지 모델에 의해 분석되며, 현재 프레임에 출현하는 사람의 자세가 예측된다. 예측된 자세는 큐에 저장되며, 큐 내의 누운 자세 비율이 사용자가 설정한 threshold를 초과할 경우, 이를 낙상으로 판단하여 즉시 알람을 전송한다. 이러한 알람은 의료진이나 보호자에게 실시간으로 전달되어 빠른 대응이 가능하게 한다(7-13번 줄). 한 번의 처리가 끝난 후, 아직 처리되지 않은 나머지 hFrames에 대해서도 동일한 과정을 반복적으로 수행하며, 이러한 절차를 통해 영상 전체에서 모든 사람의 자세를 분석할 수 있다.

www.kci.go.kr

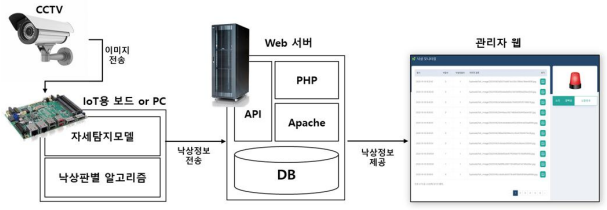


Fig. 6. Proposed System Overview

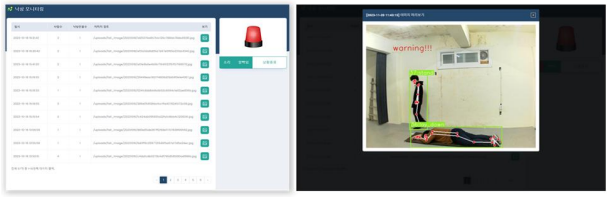


Fig. 7. Web software for manager

3. System architecture and web software

Fig 6은 제안된 낙상 탐지 시스템의 개요를 보여준다. CCTV와 IoT 보드는 물리적으로 직접 연결되거나 통신을 통해 연동되어 있으며, IoT 보드에는 머신러닝 기반 자세 탐지 모델과 낙상 탐지 알고리즘이 탑재되어 있다. 탐지된 정보는 웹서버에 전송되고, PHP 및 Apache 기반의 웹 애플리케이션이 이를 관리하며, 직관적인 웹 인터페이스를 통해 사용자에게 정보를 제공한다. 데이터는 API를 통해 데이터베이스에 저장되며, 이를 통해 장기적인 데이터 관리와 분석이 가능하다. 특히 고사양 GPU 서버를 활용하지 않아도 된다는 점이 가장 강력한 장점이다.

Fig 7은 낙상 탐지 시스템의 전체적인 동작과 인터페이스를 보여주는 시각화된 웹 기반 소프트웨어를 설명하고 있다. 이 소프트웨어는 낙상 탐지와 관련된 데이터를 실시간으로 관리하고, 관리자에게 즉각적인 피드백을 제공하기 위해 설계되었다. 시스템은 주로 좌측의 낙상 이력 화면과 우측의 낙상 관련 정보 화면으로 구성된다. 좌측에 위치한 낙상 이력화면은 시스템이 탐지한 모든 낙상 이벤트를 시간 순서대로 정리하여 표시한다. 이 화면에는 발생 시간, 탐지된 전체 인원수, 낙상으로 판단된 인원수, 그리고 해당 프레임의 이미지를 저장한 파일 경로가 명확히 나타난다. 이 경로는 시스템에 저장된 이미지를 기반으로 하며, 보기 버튼을 통해 관리자가 해당 이미지를 확인할 수 있는 기능을 제공한다. 이를 통해 관리자는 탐지된 상황의 정확성을 시각적으로 확인하거나, 필요할 경우 추가 조치를 취할 수 있다.

우측에 배치된 낙상 관련 정보 화면은 실시간으로 알림 기능과 상황 종료 기능을 포함한다. 알림 기능은 소리와 깜빡임으로 구성되어 있으며, 각각의 알림은 별도의 버튼을 통해 온, 오프 설정이 가능하다. 소리는 주변 환경에 즉각적으로 경고를 전달하며, 깜빡임은 시각적인 주의를 끌기 위해 설계되었다. 이러한 알림 기능은 낙상이 발생했을 때 현장에서 빠른 대처가 가능하도록 도울 수 있다. 추가적으로, 상황 종료 버튼은 모든 알림을 종료하고 현재 발생한 낙상 이벤트에 대한 데이터를 API를 통해 데이터베이스에 저장할 수 있는 기능을 제공한다.

IV. Experiment and evaluation

본 논문에서 제안하는 시스템의 우수성을 검증하기 위해 탐지정확도, 추론 및 학습시간 측면에서 여러 모델들과의 비교실험을 통해 검증한다. 실험을 위해 총 60,000장의 이미지 중 57,000장을 학습하였고 3000장을 검증데이터로 활용하였다. 실험에 사용된 하드웨어 사양은 Table 1에 나타내었다.

Table 1. Hardware Enviroment

Item	specification
CPU	Intel i7-10700 CPU @ 2.90Ghz
RAM	32GB
GPU	NVIDIA GeForce RTX 3060 ti

1. Evaluation of pose detection accuracy

본 연구에서 제안하는 낙상 탐지 기법의 정확도를 높이기 위해서는 우선적으로 자세 탐지의 정확도가 중요하다. 이를 위해, 분석을 통해 도출된 주요 요인들로 구성된 학습 데이터를 활용하여 다양한 머신러닝 모델로 3가지 자세 탐지 성능을 평가하였다. Table 2는 5가지 모델을 활용하여 성능을 리포팅 한 것으로 자세 탐지 모델에 대한 성능 평가 결과를 나타낸다. 평가에 사용된 지표로는 정확도(accuracy), 정밀도(precision), 재현율(recall), 그리고 F1-스코어 (F1-score)로 구성되어 있으며, 각 모델의 분류 성능을 종합적으로 분석할 수 있다. 분석결과, 로지스틱 회귀를 제외한 모든 모델에서 0.95 이상의 정확도, 정밀도, 재현율, F-1 스코어를 기록했으며 매우 우수하고 안정적으로 균형잡힌 성능을 내는 것을 증명하였다. 이것은 분석된 주요 요인들이 자세 탐지에 매우 유의했다는 것을 증명한다. 특히, 랜덤 포레스트(Random Forest)의 경우, 정확도 0.9835, 정밀도 0.9836, 재현율 0.9825, F1-점수 0.9825로, 모든 지표에서 가장 높은 성능을 기록했다. 따라서 성능이 가장 우수한 랜덤 포레스트를 본 연구의 제안 모델에 가장 적합한 머신러닝 기법으로 채택하였다.

www.kci.go.kr

Table 2. Evaluation for each model

	acc.	pre.	rec.	F1
ANN	0.9547	0.9547	0.9547	0.9547
SVM	0.9525	0.9527	0.9525	0.9525
Random For.	0.9835	0.9836	0.9825	0.9825
Logistic Reg.	0.8763	0.8758	0.8763	0.8758
KNN	0.9815	0.9815	0.9815	0.9815
Gradient Boost	0.9620	0.9623	0.9620	0.9621

Table 3. All features vs. Key features

all features	precision	recall	f1-score	samples
lay_down	0.98	0.99	0.98	975
sit	0.97	0.98	0.97	984
stand	0.99	0.98	0.98	1,041
macro f1-score			0.98	3,000
key features	precision	recall	f1-score	samples
lay_down	0.98	0.99	0.98	975
sit	0.95	0.99	0.97	984
stand	0.99	0.99	0.99	1,041
macro f1-score			0.98	3,000

Table 3은 제안 모델을 이용하여 모든 특징(33개 키포인트 + 바운딩 박스 너비 비율)을 모두 학습한 모델과 주요 요인(key features)만을 학습한 모델의 성능을 비교한 결과를 보여준다. 분석 결과, 두 모델 간 성능 차이가 크지 않은 것으로 나타났다. 이는 모든 데이터를 사용하지 않고 주요 요인만으로도 충분한 성능을 낼 수 있다는 것을 의미한다. 따라서 주요 요인만 사용하여 입력 데이터를 구성하면 데이터 크기가 줄어들어 학습 및 처리 속도를 최적화할 수 있다.

3장에서 설명한 알고리즘을 이용한 낙상탐지 성능평가를 수행하기 위해 AI-hub의 사람 인체/자세 3D 데이터에서 수집하고 임의로 추출한 100개의 영상과 직접 촬영한 100개의 영상을 활용하였다. Table 4에 의하면 수집 데이터와 촬영데이터 모두 97% 이상의 높은 성능을 보였다.

Table 4. Evaluation of fall detection of proposed system

data	input file count	success count
AI-hub	100	97
self video	100	99

2. Evaluation of training and inference time

이미지 기반 자세 탐지 모델에는 합성곱 신경망(Convolutional Neural Network, CNN)을 활용할 수 있으며, CNN은 매우 높은 성능을 제공하지만 GPU를 필요로 한다는 한계가 있다. 본 실험에서는 제안 모델과 CNN 모델을 대상으로 학습 시간, 추론 시간, 및 추론 성능을 비교하는 실험을 수행하였다. CNN 모델을 학습시키기 위한 레이어 설계와 파라미터 설정값을 Table 5에 나타내었다.

Table 5. CNN model design and parameter settings

type	values or contents
Conv2D	1-convolution layer: 32 filters, filter size (3, 3), ReLU 2-convolution layer: 64 filters, filter size (3, 3), ReLU 3-convolution layer: 128 filters, filter size (3, 3), ReLU
Max Pooling 2D	Reduces data size and emphasizes features. Pooling size: (2, 2)
Flatten	Converts 2D data into 1D format for passing to the Dense layer.
Dense	1-Dense layer: 128 nodes, ReLU 2-Dense layer (output layer): 3 nodes, softmax
Loss	sparse categorical crossentropy
Epochs	5
Optimizer	Adam

Table 6. Parameter setting for proposed model

paramters	values
max_depth	10
min_samples_leaf	1
min_samples_split	2
n_estimators	100

Table 6은 제안 모델의 학습 알고리즘에 대한 학습 파라미터 값이다. 언급되지 않은 기타 파라미터들은 사이킷런에서 제공되는 디폴트값을 활용하였다.

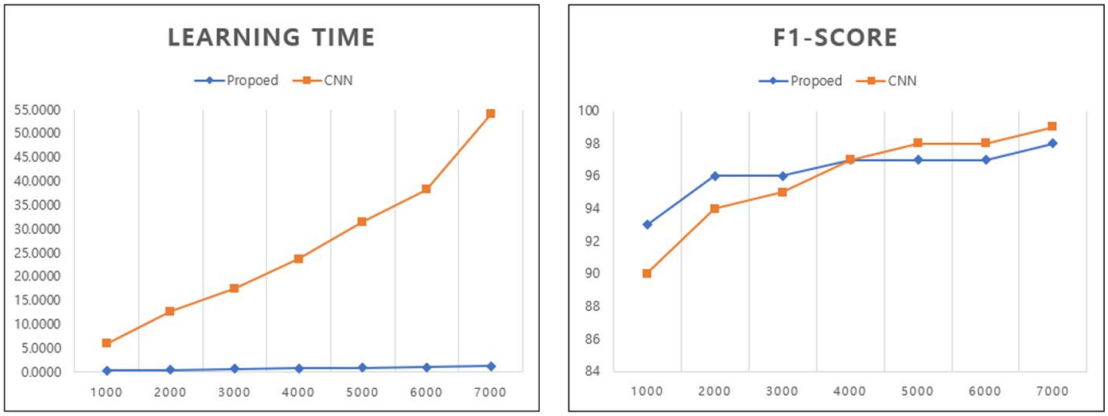
Fig 8은 데이터 학습량에 따른 CNN과 제안 모델 간의 성능을 학습 시간과 탐지 성능(F1-Score)으로 나타낸 것이다. Fig 8-a를 통해 학습 시간이 지남에 따라 두 모델 간의 시간 차이가 매우 크게 벌어짐을 알 수 있다. 제안 모델의 경우, 데이터 학습량이 7,000건에 이르러도 학습 시간은 2초를 넘지 않았다. 반면, CNN 모델은 동일한 7,000건 학습에서 40초 이상이 소요되었으며, 학습 데이터가 증가할수록 더 많은 시간이 요구될 것으로 예상된다.

Fig 8-b는 학습량에 따른 두 모델의 성능을 나타낸 것이다. 두 모델 간에 성능 차이는 크지 않았지만, 주요 특징 위주로 학습을 수행한 제안 모델이 더 빠르게 95% 이상의 성능을 달성하여, 상대적으로 적은 데이터 양으로도 충분히 적절한 모델을 구성할 수 있음을 보여주었다.

Table 7. Comparison of inference time between the proposed model and CNN

models	inference (3,000 samples)
CNN(GPU)	164.55 sec
Proposed Model based RF(CPU)	73.46 sec

www.kci.go.kr



(a) Learning Time (b) F1-Score

Fig. 8. Performance Comparison Between Proposed Model and CNN Based on Learning Effort

Table 7은 데이터 7000건을 학습한 두 모델의 학습량에 따른 추론 시간을 나타낸 것이다. 여기서 측정된 추론 시간은 이미지 입력부터 결과 출력까지 두 모델이 요구하는 모든 전처리 과정을 포함한다. 총 3,000건의 검증용 샘플로 추론 수행 시, CNN은 164.55초가 소요된 반면, 제안 모델은 73.46초 만에 추론을 완료하여, 제안 모델이 추론 시간 면에서도 우위를 점하고 있음을 입증하였다.

Table 8의 매크로 F1-스코어를 보면 CNN 모델이 매우 근소하게 높은 성능을 보였으나, 성능 차이는 미미한 수준임을 확인할 수 있다. 이를 통해 CNN 모델은 GPU 가속을 활용하여 복잡하고 정확한 탐지가 요구되는 작업에 적합할 수 있지만, 제한된 환경에서는 제안 모델이 더욱 효율적이고 실용적임을 의미한다.

Table 8. Comparison of stance detection performance between the proposed model and CNN

CNN	precision	recall	f1-score	samples
lay_down	0.98	0.99	0.99	975
sit	0.97	0.97	0.98	984
stand	0.99	0.99	0.99	1,041
		macro f1-score	0.99	3,000
Proposed	precision	recall	f1-score	samples
lay_down	0.98	0.99	0.98	975
sit	0.95	0.99	0.97	984
stand	0.99	0.99	0.99	1,041
		macro f1-score	0.98	3,000

V. Conclusions

본 논문에서는 저사양 하드웨어 환경에서도 실용적으로 적용 가능한 경량 AI 기반 낙상 탐지 시스템을 제안하였다. 기존 연구들이 주로 고성능 GPU에 의존하거나 탐지 정확도에만 초점을 맞춘 반면, 본 연구는 미디어파이프와 데이터 분석 기반 머신러닝 모델을 활용하여 낮은 연산 자원으로도 신속하고 정확한 낙상 탐지가 가능하도록 설계되었다. 제안된 시스템은 병원, 요양원 등 다양한 실내 환경에서 실시간으로 낙상을 탐지하고, 관리자에게 알림을 전송하여 신속한 응급 대응을 가능하게 한다는 점에서 실용성을 강조한다.

실험 결과, 제안 모델은 GPU를 활용하는 CNN 모델에 비해 학습 및 추론 속도에서 현저한 우위를 보였으며, 탐지 성능 측면에서도 경쟁력 있는 결과를 나타냈다. 특히, 주요 특징만을 활용한 데이터 입력 방식을 통해 불필요한 연산을 줄이면서도 모델 성능을 유지할 수 있음을 입증하였다. 이러한 최적화는 학습 및 처리 속도를 획기적으로 개선하여 자원이 제한된 환경에서도 낙상 탐지 시스템을 효과적으로 운영할 수 있는 가능성을 보여주었다.

본 연구는 고사양 하드웨어에 대한 의존성을 줄이고, 실제 환경에서의 적용 가능성을 극대화한 경량 낙상 탐지 시스템의 새로운 방향성을 제시하였다. 그러나 여전히 개선이 필요한 부분이 존재한다. 특히, 다양한 조명과 카메라 각도에서의 시스템 성능을 보완하고, 외부 환경 변화에 대한 적응성을 강화할 필요가 있다. 또한, 더욱 폭넓은 데이터셋과 다양한 머신러닝 알고리즘을 통합함으로써 탐지 정확도를 향상시키고 응용 범위를 확장할 수 있을 것이다.

향후 연구에서는 시스템의 신뢰성과 확장성을 높이는 데 중점을 두고, 스마트홈 및 공공시설 등 다양한 분야에서 실용적으로 활용할 수 있는 범용 낙상 탐지 솔루션 개발을 수행할 계획이다. 이는 안전사고 예방 및 의료 지원 시스템의 발전에 기여할 수 있을 것으로 전망된다.

www.kci.go.kr

REFERENCES

[1] Statistics Korea, "Population status and outlook of the world and Korea reflecting future population projections as of 2022," Statistics Korea, Sep. 2024, https://kostat.go.kr/board.es?mid=a10301010000&bid=207&act=view&list_no=432825

[2] HS. Kim and EH. Choi, "Validation on Adult Fall Assessment Tools: Focusing on Hospitalized Patients in a General Hospital," J. Muscle Jt. Health, Vol. 31, No. 2, pp. 65-74, Aug. 2024. DOI: 10.5953/JMJH.2024.31.2.65

[3] EM. Choi and YJ. Park and SH. Bang and JH. Jeong, "A Study on the Main Diagnostic Code according to the Analysis of the Frequency of Fall Patients by Case-Centered Damage External Code," J. Korea Inst. Inf. Electron. Commun. Technol., Vol. 16, No. 6, pp. 533-539, Dec. 2023, DOI: 10.17661/jkiiect.2023.16.6.533

[4] YS. Choi, "The Effect of Patient-related and Environment-related Characteristics on the Risk of Falling in Inpatient Care Unit - A Case-Control Study to identify Inpatient Fall Risk Factors," J. Korea Inst. Healthc. Archit., Vol. 27, No. 4, pp. 61-70, Dec. 2024, DOI: 10.15682/kjiha.2021.27.4.61

[5] S. Park and K. Kim, "Bayesian Onset Measure of sEMG for Fall Prediction," J. Korea Robot. Soc., Vol. 19, No. 2, pp. 213-220, June. 2024, DOI: 10.7746/jkros.2024.19.2.213

[6] J. Kim, "Comparative Study on Deep-learning Object Detection Based Recognizing Worker Falls in Underground Utility Tunnels," J. Korea Acad.-Ind. Coop. Soc., Vol. 25, No. 8, pp. 22-30, Aug. 2024, DOI: 10.5762/KAIS.2024.25.8.22

[7] YK. Kang and HY. Kang and DS. Weon, "Human Skeleton Keypoints based Fall Detection using GRU," J. Korea Acad.-Ind. Coop. Soc., Vol. 22, No. 2, pp. 127-133, Feb. 2021, DOI: 10.5762/KAIS.2021.22.2.127

[8] EB. Ko and IW. Jeong and CH. Lee, "Design and Implementation of Platform for Small-Scale Smart Farm using Deep Learning," J. Korean Soc. Manuf. Process Eng., Vol. 22, No. 11, pp. 59-67, Nov. 2023, DOI: 10.14775/ksmpe.2023.22.11.0059

[9] D. Lee and J. Oh and Y. Choi and D. Lee and H. Lee and SB. Shim and Y. Ha, "Development of YOLO-based apple quality sorter," Korean J. Agric. Sci., Vol. 50, No. 3, pp. 373-382, Sep. 2023, DOI: 10.7744/kjoas.500307

[10] YW. Lee and JH. Park and SY. Shin, "Implementation of Fall Detection Based on CNN-LSTM," J. Korean Inst. Commun. Inf. Sci., Vol. 47, No. 2, pp. 340-347, Feb. 2022, DOI: 10.7840/kics.2022.47.2.340

[11] Q. Luo and S. Hu and C. Li and G. Li and W. Shi, "Resource scheduling in edge computing: A survey," IEEE Commun. Surv. Tutor., Vol. 23, No. 4, pp. 2131-2165, Aug. 2021, DOI: 10.1109/COMST.2021.3106401

[12] MS. Aslanpour and AN. Toosi and C. Cicconetti, "Serverless edge computing: vision and challenges," Proc. 2021 Australas. Comput. Sci. Week Multiconf., pp. 1-10, Feb. 2021, DOI: 10.1145/3437378.3444367

[13] J. Wu and F. Dong and H. Leung and Z. Zhu and J. Zhou, "Topology-aware federated learning in edge computing: A comprehensive survey," ACM Comput. Surv., Vol. 56, No. 10, pp. 1-41, June. 2024, DOI: 10.1145/3659205

[14] C. Lugaresi and J. Tang and H. Nash and C. McClanahan, "Mediapipe: A framework for building perception pipelines," arXiv preprint arXiv:1906.08172, pp. 1-9, June. 2019, DOI: 10.48550/arXiv.1906.08172

[15] S. Negi and M. Garg and H. Maindola and V. Kansal and U. Jain and S. Bhatla, "Real-Time Human Pose Estimation: A MediaPipe and Python Approach for 3D Detection and Classification," Proc. 2023 3rd Int. Conf. Technol. Adv. Comput. Sci., pp. 128-133, Jan. 2024, DOI: 10.1109/ICTACS59847.2023.10390506

[16] J. Redmon and S. Divvala and R. Girshick and A. Farhadi, "You only look once: Unified, real-time object detection," Proc. IEEE Conf. Comput. Vis. Pattern Recognit., pp. 779-788, Dec. 2016, DOI: 10.1109/CVPR.2016.91

[17] A. Bochkovskiy and CY. Wang and HYM. Liao, "Yolov4: Optimal speed and accuracy of object detection," arXiv preprint arXiv:2004.10934, pp. 1-17, Apr. 2020, DOI: 10.48550/arXiv.2004.10934

[18] CY. Wang and A. Bochkovskiy and HYM. Liao, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., pp. 7464-7475, Aug. 2023, DOI: 10.1109/CVPR52729.2023.00721

[19] BG. Shin and UH. Kim and SW. Lee and JY. Yang and W. Kim, "Fall Detection Based on 2-Stacked Bi-LSTM and Human-Skeleton Keypoints of RGBD Camera," KIPS Trans. Softw. and Data Eng., Vol. 10, No. 11, pp. 491-500, Nov. 2021, DOI: 10.3745/KTSDE.2021.10.11.491

www.kci.go.kr

Authors



Minhyung Ryu received a B.S. degree in the Department of Statistics from Yeungnam University, Gyeongsan, Korea, in 2021 and an M.S. degree in computer engineering from Yeungnam University, in 2023.

She is currently a Researcher in the R&D Center at NEARNETWORKS. She is interested in health care, vision systems, artificial intelligence, and data analysis.



Jaegoo Shim received an M.S. degree in the Department of Radiology Technology from Hanseo University, Seosan, Korea in 2010 and a Ph.D. degree in Department of Health from Hanseo University, in 2014.

He is currently a Professor in the Department of Radiology Technology from Daegu Health College. He is interested in healthcare, radiology technology, artificial intelligence, and data analysis.

www.kci.go.kr