

# CNN-LSTM 기반 낙상 감지 시스템 구현

이영욱\*, 박재한\*, 신수용°

## Implementation of Fall Detection Based on CNN-LSTM

Yeong-wook Lee\*, Jae-han Park\*, Soo-young Shin°

요약

본 논문에서는 영상 데이터 처리 모델인 CNN(convolutional neural networks)과 시퀀스 데이터 처리 모델인 LSTM(long short term memory)을 결합하여 낙상 감지 시스템 모델을 구현한다. 비접촉 적외선 열상 카메라 모듈(thermal array sensor)을 이용하여 인간의 보편적인 3가지 자세(서 있는 자세, 앉은 자세, 누운 자세)의 데이터셋을 수집하였다. 수집된 데이터를 CNN 알고리즘의 입력 값으로 사용하여 나온 출력 값을 LSTM 알고리즘의 입력 값으로 취하는 CNN-LSTM 결합형 알고리즘을 사용하였다. 본 논문에서는 비접촉 적외선 열상 카메라 모듈을 이용하여 수집한 비디오 데이터셋을 통해서 CNN-LSTM, ResNetCNN-LSTM 두 가지 딥 러닝 모델에서 낙상 감지 시스템을 구현한 후 각각 정확도를 비교 분석하였다. 실험 결과를 통한 정확도를 기반으로 향후 연구 방향에 대해 논의한다.

**Key Words** : Deep learning, CNN, RNN, LSTM, Fall Detection

ABSTRACT

In this paper, we implement a fall detection system through combining CNN(convolutional neural network) and LSTM(long short term memory) that are kinds of image data processing and sequence data processing, respectively. A dataset of three common human postures(standing posture, sitting posture, and lying posture) was collected using a non-contact infrared thermal array sensor. The output is obtained by employing a CNN-LSTM combination method, in which the sample data is first fed into the CNN, and the output of the CNN is then utilized as an input for the LSTM network. In this paper, the accuracy was compared and analyzed after implementing fall detection systems in two deep learning models,(CNN-LSTM, ResNetCNN-LSTM), through video datasets collected using non-contact infrared thermal array sensor. Future research directions are proposed based on the accuracy of the experimental data.

### 1. 서론

고령화 문제는 우리나라뿐만 아니라 전 세계가 당면한 심각한 문제이다. 인구 층에서 고령인구가 차지

하는 비율이 높아질수록 여러 가지 사회적인 문제가 발생하지만, 그중 고령인구의 건강 문제는 우리가 해결해야 할 과제 중 하나다. 우리나라 노인 인구 층에서 가장 빈번히 일어나는 안전사고는 가정, 의료시설

\* 이 논문은 2021년도 정부(교육과학기술부)의 재원으로 한국연구재단의 대학중점연구소 지원 사업으로 수행된 연구임 (2018R1A6A1A03024003).

♦ First Author : Kumoh National Institute of Technology, lyw1836@naver.com, 학생회원

° Corresponding Author : Kumoh National Institute of Technology, wdragon@kumoh.ac.kr, 종신회원

\* Kumoh National Institute of Technology, qkrwogks7094@kumoh.ac.kr, 학생회원

논문번호 : 202111-322-D-RU, Received November 23, 2021; Revised December 13, 2021; Accepted December 15, 2021

을 막론하고 낙상사고가 가장 큰 비중을 차지하고 있다. 2019년 의료기관평가인증원에서 발간한 ‘2019년 환자 안전 통계 연보’에서 2016년부터 2019년까지 4개년의 환자안전사고의 추이, 종류, 위해 정도 등에 대한 비교 분석을 하였다. 통계를 근거로 낙상사고는 가장 큰 비중인 44.3%를 차지하였고 매해 늘어나는 추세다<sup>1)</sup>. 이에 따라 본 논문에서는 의료시설이나 가정에서 실시간으로 사용자의 형상을 비접촉 적외선 열상 카메라 모듈(thermal array sensor)을 통해 촬영하여 딥 러닝 알고리즘을 통하여 낙상을 감지할 수 있는 시스템을 구현한다.

일반적인 카메라는 객체를 육안과 같은 수준으로 인식할 수 있지만, 적외선 열상 카메라 모듈은 객체를 픽셀 배열로 나타내기 때문에, 식별하기 어려운 문제점이 있다. 그런데도 본 논문에서 적외선 열상 카메라 모듈을 사용한 이유는 일반적인 카메라는 사용자의 사생활이 그대로 노출되기 때문에 사생활 보호 측면에서 어려움이 있으나 적외선 열상 카메라 모듈은 픽셀 배열로 표시되는 측면으로 인한 사생활 보호 측면에서 우수한 면을 보여주기 때문이다<sup>2)</sup>. 화질과 관련된 문제점은 딥 러닝을 통하여 학습시켜 정확도를 개선하는 실험을 진행하였다.

CNN(convolutional neural networks)은 최근 이미지 분석을 위한 딥 러닝 구조에 가장 많이 사용되는 모델이다. 이 구조는 다양한 컨볼루션 필터를 사용하여 이미지의 특징을 추출하고 추출된 특징들을 사용하여 객체 탐지(object detection), 이미지 분할(semantic segmentation), 이미지 분류(classification)를 수행한다. CNN 구조의 성능을 향상하는 방법으로는 일반적으로 신경망을 더 깊게 만들거나 입력 이미지의 크기를 키우는 방법이 있다. 그림 1은 CNN의 기본적인 구조를 나타내고 있으며 특징 추출 후 추출된 특징에 근거하여 이미지를 분류하는 구조를 보인다<sup>3)</sup>.

RNN(recurrent neural networks)은 시간적인 순서가 있는 딥 러닝 모델이다. 이러한 시간적인 순서를

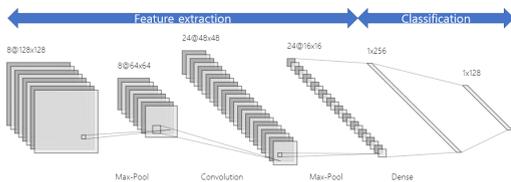


그림 1. 기본적인 CNN 구조  
Fig. 1. Basic CNN structure

시퀀스(sequence)라고 부른다. RNN에서는 시퀀스를 입력과 출력 값으로 받아 처리하는 네트워크 구조이며 주로 자연어 처리(NLP : natural language processing)에 활용된다. 대부분의 딥 러닝 모델들은 은닉 층(Hidden layer)에서 출력된 값이 출력 층(Output layer)의 입력 값으로 전달되었지만, RNN은 은닉 층의 출력 값이 출력 층으로 전달되는 것은 물론 다시 은닉 층의 입력으로 전달되는 특징을 가지고 있다. 이는 그림 2에 나타난 RNN의 구조에서 볼 수 있듯이 이전 은닉 층의 출력 값을 그대로 다음 은닉 층의 입력 값으로 받음으로써 시간적인 순서를 고려할 수 있게 된다<sup>4,5)</sup>.

본 논문에서는 두 가지 딥 러닝 모델을 사용하였다. 앞서 언급한 기본적인 CNN과 RNN, 두 가지 딥 러닝 모델을 응용하여 CNN-LSTM, ResNetCNN-LSTM을 구현한다. 각 모델의 개요와 구조는 다음 장에 기술된다.

낙상 감지 시스템은 다양한 방식으로 연구된 바 있다. 산업현장에서 낙상 사고 발생 시 신속한 사고 발견 및 조치를 위하여 산업현장에 적용 가능한 IMU 센서 기반 낙상 감지 시스템을 구현하여 사용자의 움직임에 따른 데이터를 기반으로 낙상 감지 분류 알고리즘을 통해 낙상 상태를 판단하는 연구도 진행되었다<sup>6)</sup>. 또한 RGB 이미지 기반 데이터셋과 가속도계 센서를 이용하여 낙상을 감지하는 다중기능의 컨볼루션 신경망 기법이 연구되었다. 낙상 감지와 관련된 여러 가지 연구가 진행되었지만, 객체를 식별하기 어려운 비접촉 적외선 열상 카메라 모듈을 이용한 경우는 전례가 없다. 또한 비디오 형식의 데이터를 이용한 딥 러닝 모델 또한 구현하기가 힘들다. 본 논문에서는 객

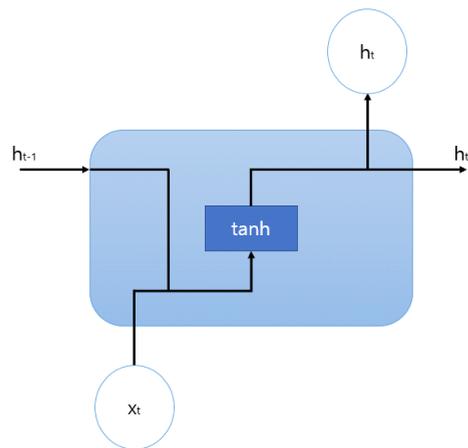


그림 2. 기본적인 RNN 구조  
Fig. 2. Basic RNN structure

체를 식별하기 어려운 비접촉 열상 카메라 모듈과 비디오 데이터를 이용하여 두 가지 낙상 감지 시스템 모델을 구현하고 성능을 확인한다.

논문 구성은 다음과 같다. 제 II장에서는 낙상 감지 시스템에 대한 설명과 본 논문에서 이용된 두 가지 딥러닝 모델에 대하여 설명하고, 제 III장에서는 실험 환경 및 실험방법에 관하여 기술하며 실험 결과를 이용하여 정확도 및 각 모델의 특징을 비교, 분석한다. 그리고 IV장에서 결론을 도출해낸다.

## II. 본 론

### 2.1 낙상 감지 시스템

#### 2.1.1 개요

본 논문에서 제안하는 낙상 감지 시스템은 열상 카메라 모듈을 통해 수집한 시각적 데이터세트와 딥러닝 기법을 이용하였다. 제 I 장에서 언급한 기존 낙상 감지 시스템과는 달리 사용자의 신체에 센서가 탑재된 IoT 장비를 부착하지 않고 실내 공간의 천장에 설치하여 실시간으로 낙상을 감지하는 시스템을 구현한다. 또한 일반카메라의 경우 사용자의 낙상 정보 외에도 불필요한 정보가 수집되어 사생활 침해의 우려가 있다. 하지만 본 논문에서 사용된 열상 카메라 모듈을 이용하였을 경우 객체를 열 배열의 형태로 나타내기 때문에 사생활 보호 측면에서 우수한 성능을 보이고 불필요한 정보를 수집하지 않는다. 그림 3은 본 논문에서 구현하는 낙상 감지 시스템의 대략적인 개요도를 보여준다. 천장에 설치된 열상 카메라 모듈을 이용하여 사람의 3가지 자세를 탐지하여 데이터세트를 수

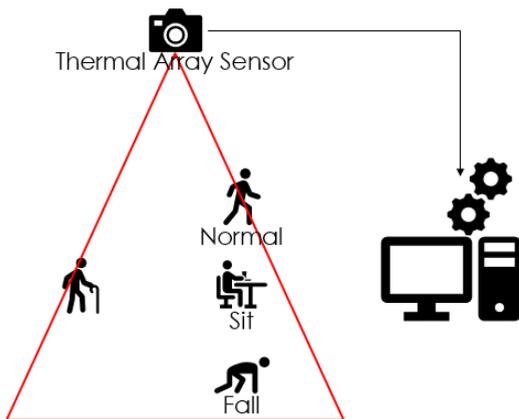


그림 3. 낙상 감지 시스템 개요도  
Fig. 3. Fall detection system overview

집한다.

#### 2.1.2 시스템 구조

낙상 감지 시스템을 구현하기 위해 NVIDIA사의 젯슨 나노 보드와 비접촉 열상 카메라 모듈을 i2c 인터페이스를 통하여 연결하여 실시간으로 데이터세트를 수집하였다. 수집된 데이터세트는 전처리 과정에서 데이터 증강(augmentation)을 통하여 임의로 각도를 회전시켜 데이터세트의 양을 증가했다. 이후 본 논문에서 제안한 시스템 (CNN-LSTM, ResNetCNN-LSTM) 모델에 데이터를 적용해 결과를 도출해낸다.

### 2.2 CNN-LSTM(convolutional neural networks-long short term memory)

#### 2.2.1 개요

제 I 장에서 언급한 두 가지 딥러닝 모델은 인간의 행동 인식 구현에서 각각의 장단점을 지니고 있다. CNN 모델의 경우 공간적 특징을 학습할 때 좋은 성능을 보이지만 하나의 프레임만을 입력으로 받기 때문에 이전 프레임들의 정보를 반영하기가 어렵다. 그로 인해 낙상과 일상생활에서의 높은 동작 등과 구별이 어려운 문제가 있다. RNN 모델은 시간적 특징을 학습할 때 좋은 성능을 보이지만 영상 프레임과 같은 데이터를 입력으로 처리할 때 어려움이 있다.

위와 같은 두 딥러닝 모델의 단점을 보완하기 위해 두 모델을 결합하여 CNN-LSTM 모델을 구현하였다. 영상 프레임을 CNN의 입력 값으로 취하여 나오는 출력 값을 LSTM의 입력 값으로 받음으로써 시공간적 특징 모두 학습을 할 수 있게 된다<sup>[8]</sup>.

#### 2.2.2 네트워크 구조

그림 4는 CNN-LSTM 네트워크의 구조를 보여준다. CNN-LSTM 모델은 CNN 인코더와 LSTM 디코더의 쌍으로 이루어져 있다. 인코더 역할을 하는 CNN 네트워크는 4개의 컨볼루션 레이어와 fully connect

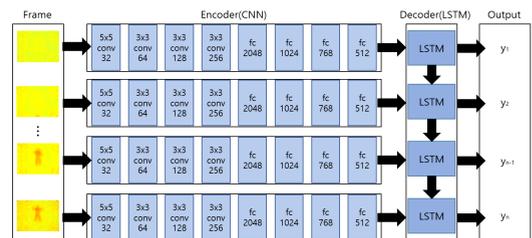


그림 4. CNN-LSTM 네트워크 구조  
Fig. 4. CNN-LSTM network structure

layer로 구성되며 모든 입력 2D 이미지를 1차원의 벡터로 인코딩한다. CNN 모듈에서 출력된 결과 값은 다시 디코더 역할을 하는 LSTM 네트워크에서 CNN 인코더로부터 시퀀스 입력 벡터를 수신하고 또 다른 1D 시퀀스를 출력한다. 결과 예측을 위해 마지막 단계에서 완전히 연결된 신경망(fully-connected neural network)이 연결된다.

### 2.3 ResNet CNN-LSTM

#### 2.3.1 개요

ResNet은 2015년 ILSVRC에서 우승을 차지한 CNN 모델이다. 마이크로 소프트웨어에서 개발된 네트워크 구조이다. 이전까지의 딥 러닝 모델들은 신경망의 깊이(Depth)가 깊어질 때 기존보다 성능이 향상될 것 같았지만 실제로 신경망의 깊이를 늘렸을 때 일정 부분까지 상승하다가 그 이후부터는 gradient vanishing, exploding과 같은 문제가 발생하여 오히려 성능을 떨어뜨리는 문제를 초래하였다. 이러한 현상을 degradation problem이라고 한다. ResNet 구조는 degradation 현상을 해결한 새로운 네트워크 구조이다<sup>19)</sup>. ResNet의 구조에서는 이전까지의 신경망에서 볼 수 없었던 residual 구조를 제안하였고 이로 인하여 신경망이 깊어질수록 신경망 초기에서 추출된 특징들이 점점 소실되는 문제점을 해결하였다<sup>19)</sup>.

#### 2.3.2 네트워크 구조

ResNet의 경우 일반 네트워크에서는 사용하지 않던 residual 구조를 제안하고 사용하였다. 이 구조로 인하여 기존의 앞쪽 네트워크에서 추출하였던 특징가중치들이 뒤쪽 네트워크로 갈수록 소실되는 문제점을 해결할 수 있었다. 기존의 신경망은 입력값  $x$ 를 출력 값  $y$ 로 매핑하는 함수  $H(x)$ 를 얻는 것이 목적이었다. 그러나 ResNet은  $F(x)+x$ 를 최소화하는 것을 목적으로 한다.  $x$ 는 현시점에서 변할 수 없는 값이므로  $F(x)$ 를 0에 수렴하게 만드는 것이 목적이다.  $F(x)$ 가 0이 되면 출력과 입력이 모두  $x$ 로 같아지게 된다.  $F(x)=H(x)-x$ 이므로  $F(x)$ 를 최소로 만드는 것은  $H(x)-x$ 를 최소로 만드는 것과 동일하다. 여기서

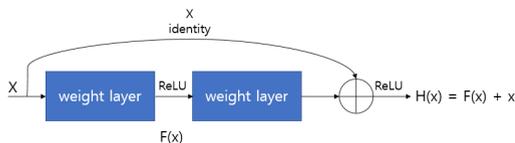


그림 5. Residual block 구조  
Fig. 5. Residual block structure

$H(x)-x$ 를 잔차(residual)라고 한다. 즉 잔차를 최소로 만드는 것이기에 ResNet이라 이름이 붙게 되었다. ResNet CNN-LSTM 모델의 경우 기존 CNN-LSTM 모델과 비교하였을 때 인코더 부분인 CNN의 네트워크 구조가 변경되고 디코더 부분인 LSTM에서는 구조가 동일하다.

## III. 실험

### 3.1 실험방법

본 논문에서는 CNN-LSTM, ResNetCNN-LSTM 두 가지 모델을 구현하고 각 모델의 특징과 정확도를 이용한 비교 분석을 진행하였다. 실험에 이용할 데이터셋을 수집하기 위해 비접촉 열상 카메라 모듈(MLX90640)을 NVIDIA사의 젯슨 나노 보드와 I2C 방식으로 연결하여 직접 데이터를 수집하였다. 구성된 데이터 수집용 장치는 그림 6과 같다.

비접촉 적외선 열상 카메라 모듈(thermal array sensor)은  $32 \times 24$  IR 열상 센서를 탑재하고 있으며 I2C 인터페이스를 통해 총 768개의 온도 측정값을 반환한다. 또한 최대 16Hz의 프레임으로  $-40^{\circ}C$ 에서  $300^{\circ}C$ 까지 온도 측정이 가능하며  $110^{\circ} \times 70^{\circ}$ 의 시야각을 특징으로 한다.

데이터셋의 형식은 7~100프레임까지 다양한 길이의 비디오와 RTX Titan GPU를 이용하여 딥 러닝 학습을 하였다. 데이터셋은 인간의 보편적인 자세(서 있는 자세, 앉아있는 자세, 누운 자세) 총 세 가지 클래스로 나누었고 그림 7은 열상 카메라 모듈에서 수집한 세 가지 자세의 예를 보여준다. 본 실험에서

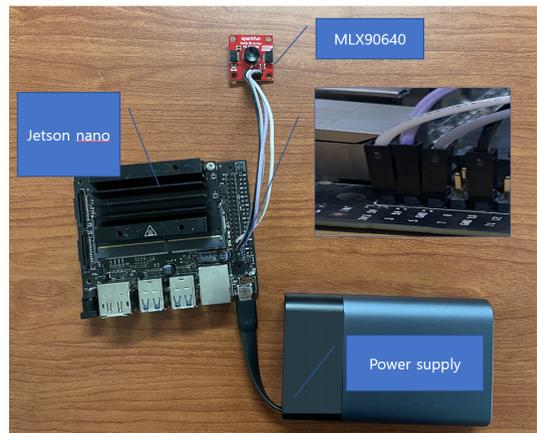


그림 6. 데이터 수집용 장치  
Fig. 6. Device for data collection

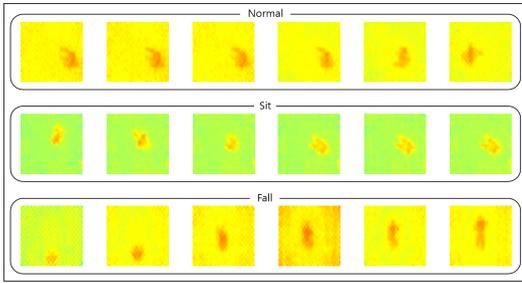


그림 7. 실험에 사용된 클래스와 데이터세트의 예  
Fig. 7. Examples of classes and date sets used in the experiment.

분석하는 두 가지 딥 러닝 모델 모두 학습을 1,010회로 고정했으며 매 100번째의 학습 가중치 파일을 저장하도록 설정하였다. 모델별 입력 파라미터는 표 1에 나타나 있다. 데이터세트는 총 1,643개의 비디오를 사용하였으며 세 가지 클래스(Normal, Sit, Fall)가 각각 551, 592, 500개로 구성되어있다.

보드와 연결된 카메라 모듈은 2.3m 높이의 천장에 설치되어 데이터를 수집하고 데스크톱으로 그 결과를 실시간으로 확인하였다.

표 1. 각 모델별 입력 파라미터  
Table 1. Input parameters for each model.

Model	Parameter	Value
	Number of epoch	1010
	Learning rate	$10^{-4}$
	Dropout rate	0.5
CNN-LSTM	Input dim,	512
	Output dim	256
ResNetCNN-LSTM	Input dim	512
	Output dim	256

표 2. 실험장비  
Table 2. Experiment equipment

Company	Sparkfun	NVIDIA	NVIDIA
Model name	MLX 90640	Jetson nano	RTX Titan

표 3. 실험에 사용된 클래스와 데이터세트의 수  
Table 3. The number of classes and datasets used in the experiment.

Class	Number of dataset
Normal	551
Sit	592
Fall	500
Total	1643

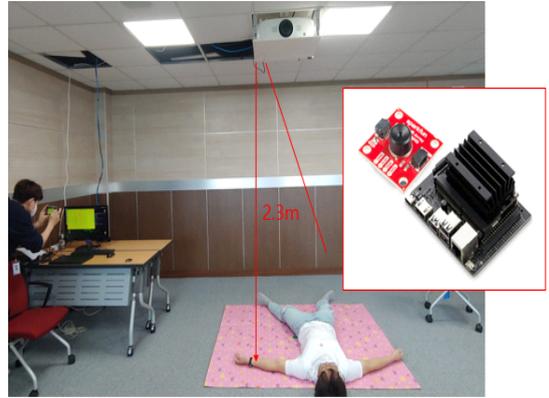


그림 8. 실험 환경  
Fig. 8. Experiment environment

### 3.2 실험 결과

#### 3.2.1 CNN-LSTM

CNN-LSTM의 실험 결과는 그림 9에 나타나 있다. 위의 그래프는 학습 결과의 손실 그래프를 나타내고 아래의 그래프는 학습 결과의 정확도 그래프를 나타

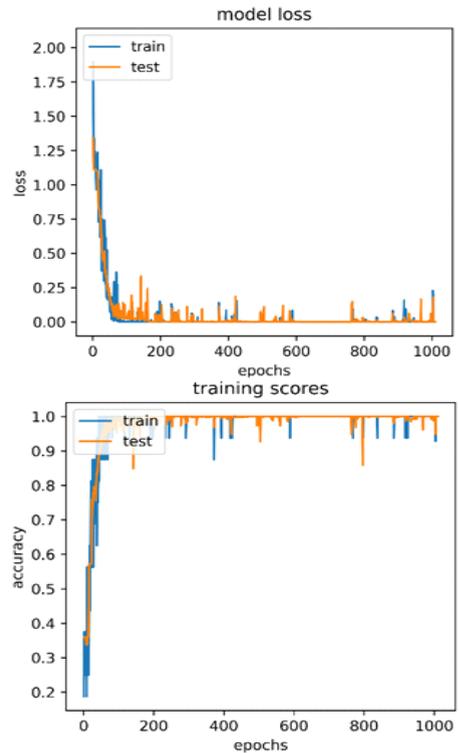


그림 9. CNN-LSTM의 실험 결과(상 : Loss 하 : Accuracy)  
Fig. 9. CNN-LSTM's result of experiment(upper : Loss lower: Accuracy)

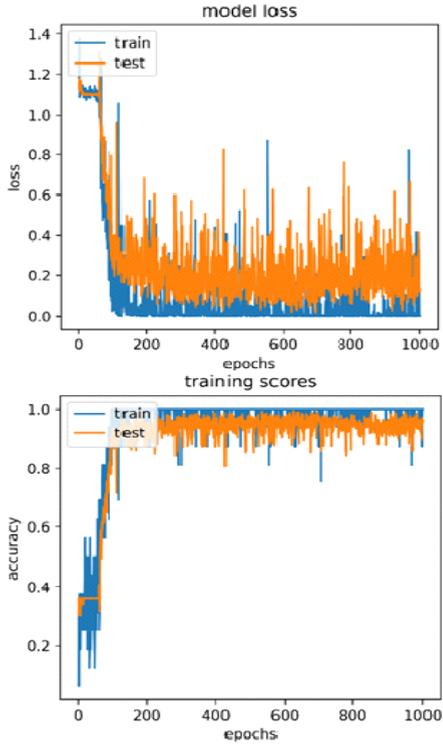


그림 10. ResNetCNN-LSTM의 실험 결과(상:Loss 하: Accuracy)  
 Fig. 10. ResNetCNN-LSTM's result of experiment  
 (upper : Loss lower : Accuracy)

낸다. 손실도는 점차 감소하다가 약 100회째의 학습부터 안정화되는 것을 볼 수 있다. 정확도 또한 점진적으로 상승하여 100회째의 학습부터 100에 가까운 정확도에 수렴하는 양상을 보여준다.

### 3.2.2 ResNetCNN-LSTM

그림 11은 ResNetCNN-LSTM의 실험 결과를 보여준다. CNN-LSTM의 결과와는 달리 약 200회 이후의 학습에서부터 손실도와 정확도가 높은 성능을 보이는 것을 확인할 수가 있다. CNN-LSTM의 경우 손실도와 정확도 모두 일정 수치에서 유지되는 경향을 보였지만 ResNetCNN-LSTM은 학습이 끝날 때까지 큰 폭으로 손실도와 정확도 모두 변화가 있었다. 표 4는 두 시스템 모델의 평균 정확도를 나타낸다. CNN-LSTM 모델이 손실도와 정확도 측면 모두 ResNetCNN-LSTM 모델보다 우수한 성능을 보여주었다. 또한 ResNetCNN-LSTM 역시 마찬가지로 높은 성능을 보여주었다. 그림 11과 12는 각각의 모델이 예측한 결과를 보여준다. CNN-LSTM은 1,643개의 데이터셋 중 1,640개를 정답으로 예측하였으며 3개의

	filename	y	y_pred
0	v_Fall_g01_c415	Fall	Fall
1	v_Sit_g01_c435	Sit	Sit
2	v_Sit_g01_c288	Sit	Sit
3	v_Sit_g01_c313	Sit	Sit
4	v_Fall_g01_c317	Fall	Fall
...	...	...	...
1638	v_Normal_g01_c93	Normal	Normal
1639	v_Sit_g01_c117	Sit	Sit
1640	v_Sit_g01_c455	Sit	Sit
1641	v_Normal_g01_c434	Normal	Normal
1642	v_Fall_g01_c98	Fall	Fall

1640 rows × 3 columns

그림 11. CNN-LSTM 예측 결과  
 Fig. 11. CNN-LSTM prediction result

	filename	y	y_pred
0	v_Fall_g01_c415	Fall	Fall
1	v_Sit_g01_c435	Sit	Sit
2	v_Sit_g01_c288	Sit	Sit
3	v_Sit_g01_c313	Sit	Sit
4	v_Fall_g01_c317	Fall	Fall
...	...	...	...
1638	v_Normal_g01_c93	Normal	Normal
1639	v_Sit_g01_c117	Sit	Sit
1640	v_Sit_g01_c455	Sit	Sit
1641	v_Normal_g01_c434	Normal	Normal
1642	v_Fall_g01_c98	Fall	Fall

1613 rows × 3 columns

그림 12. ResNetCNN-LSTM 예측 결과  
 Fig. 12. ResNetCNN-LSTM prediction result

오답을 예측하였다. ResNetCNN-LSTM의 경우 1,643개의 데이터셋 중 1,613개를 정답으로 예측하였으며 30개의 오답을 예측하였다. 예측 결과를 바탕으로 한 정확도는 표 4에 나타나 있다.

표 4. 시스템 모델의 정확도 비교  
Table 4. System Model Accuracy Comparison

Model	Accuracy
CNN-LSTM	99.81%
ResNetCNN-LSTM	98.17%

#### IV. 결 론

본 논문에서는 인간의 낙상을 감지하기 위하여 두 가지 딥 러닝 모델과 데이터셋을 이용한 학습을 통해 낙상을 감지할 수 있는 시스템을 구현하고 각 모델의 정확도와 성능을 비교 분석하였다.

실험을 통해 두 가지 모델 모두 높은 정확도와 성능을 보이는 것을 확인하였다. 실험에 적용한 1,010회의 학습의 경우 CNN-LSTM이 ResNetCNN-LSTM보다 더 나은 성능을 보여주었으며 딥 러닝 학습의 기울기 소실 등과 같은 문제는 발생하지 않았으므로 CNN-LSTM과 ResNetCNN-LSTM 모델 모두 낙상 감지에 적합한 모델인 것을 확인할 수 있었다. 예상과 달리 기울기 소실과 같은 문제가 발생하지 않은 이유는 디코더인 LSTM 모듈에서 -1과 1사이 값을 출력하는 tanh 활성화수를 사용하여 역전파로 학습하는 과정에서 가중치가 발산하거나 0으로 수렴하게 되는 현상을 보완하였기 때문이다.

추후 본 시스템을 기반으로 더욱 다양한 인간의 자세를 기반으로 한 클래스와 데이터를 이용하여 낙상 뿐만 아니라 통합사고 감지 시스템을 구현하고 이미 지 기반인 2D CNN과 3D CNN의 구현 가능성을 검토하고 연구하여 기존 시스템 모델과 비교 및 분석을 통한 확장 연구가 필요하다.

#### References

[1] Ministry of Health and Welfare, Korea Institute for Healthcare Accreditation, *Korean Patient Safety Incident Report 2019*, Seoul: Korea Institute for Healthcare Accreditation, 2020.

[2] Y. W. Lee, S. Y. Shin. Fall Detection System using the Thermal Array Sensor. *KICS 2021*, pp. 1124-1125, Jeju, Korea, Jun. 2021.

[3] Y. LeCun, et al., "Gradient-based learning applied to document recognition," in *Proc. IEEE*, vol. 86, no. 11, pp. 2278-2324, 1998.

[4] S. Liao, et al., "Logsig-RNN: A novel

network for robust and efficient skeleton-based action recognition," arXiv preprint arXiv: 2110.13008, 2021.

[5] A. H. Khan, S. Li, and X. Cao, "Tracking control of redundant manipulator under active remote center-of-motion constraints: An RNN-based metaheuristic approach," *Sci. Chin. Inf. Sci.*, vol. 64, no. 3, pp. 1-18, 2021.

[6] S. P. Yadav, et al., "Survey on machine learning in speech emotion recognition and vision systems using a recurrent neural network (RNN)," *Archives Comput. Meth. in Eng.*, pp. 1-18, 2021.

[7] J.-J. Kim, et al., "Implementation of the fall detection system based on IMU sensor applicable to industrial sites," *J. KIICE*, pp. 33-36, 2021.

[8] F. Elmaz, et al., "CNN-LSTM architecture for predictive indoor temperature modeling," *Building and Environ.*, vol. 206, no. 108327, 2021.

[9] K. He, et al., "Deep residual learning for image recognition," in *Proc. IEEE Conf. CVPR*, 2016.

[10] S. Patnaik, S. Kumari, and S. Das Mahapatra, "Comparison of deep CNN and ResNet for handwritten devanagari character recognition," *2020 IEEE 1st ICCE*, Kolkata, India, Sep. 2020.

#### 이 영 욱 (Yeong-wook Lee)



2020년 2월 : 금오공과대학교 전자공학부 졸업  
2020년 3월~현재 : 금오공과대학교 IT융복합공학과 석사과정  
<관심 분야> 머신러닝, 딥러닝, 영상 처리

[ORCID:0000-0003-2180-6161]

박 재 한 (Jae-han Park)



2020년 2월 : 금오공과대학교  
전자공학부 졸업  
2020년 3월~2021년 12월 : 금  
오공과대학교 IT융복합공학  
과 석·박사 통합과정  
<관심 분야> 딥러닝, 영상 처  
리

[ORCID:0000-0002-3856-5846]

신 수 용 (Soo-young Shin)



1999년 2월 : 서울대학교 전기  
공학부 졸업  
2001년 2월 : 서울대학교 전기  
공학부 석사  
2006년 2월 : 서울대학교 전기  
공학부 박사  
2010년~현재 : 금오공과대학교  
전자공학부 교수

<관심 분야> 5G/B5G 무선 접속 기술, 드론 응용,  
혼합 현실, 블록체인, 머신러닝, 딥러닝

[ORCID:0000-0002-2526-2395]