

Gait Type Classification Using Multi-modal Ensemble Deep Learning Network

Hee-Chan Park*, Young-Chan Choi**, Sang-Il Choi***

*Researcher, Dept. of Computer Science and Engineering, Dankook University, Yongin, Korea
Researcher, AlcheraInc, Seongnam, Korea

**Researcher, Dept. of AI-based Convergence, Dankook University, Yongin, Korea

***Professor, Dept. of Computer Science and Engineering, Dankook University, Yongin, Korea
Professor, Dept. of AI-based Convergence, Dankook University, Yongin, Korea

[Abstract]

This paper proposes a system for classifying gait types using an ensemble deep learning network for gait data measured by a smart insole equipped with multi-sensors. The gait type classification system consists of a part for normalizing the data measured by the insole, a part for extracting gait features using a deep learning network, and a part for classifying the gait type by inputting the extracted features. Two kinds of gait feature maps were extracted by independently learning networks based on CNNs and LSTMs with different characteristics. The final ensemble network classification results were obtained by combining the classification results. For the seven types of gait for adults in their 20s and 30s: walking, running, fast walking, going up and down stairs, and going up and down hills, multi-sensor data was classified into a proposed ensemble network. As a result, it was confirmed that the classification rate was higher than 90%.

▶ **Key words:** Gait type, Deep learning, Ensemble network, Smart insole, Multi-modal sensor

[요 약]

본 논문에서는 멀티 센서가 장착된 스마트 인솔로 측정된 보행 데이터에 대해 앙상블 딥러닝 네트워크를 이용하여 보행의 타입을 분류하는 시스템을 제안한다. 보행 타입 분류 시스템은 인솔에 의해 측정된 데이터를 정규화하는 부분과 딥러닝 네트워크를 이용하여 보행의 특징을 추출하는 부분, 그리고 추출된 특징을 입력으로 보행의 타입을 분류하는 부분으로 구성되어 있다. 서로 다른 특성을 가지는 CNN과 LSTM을 기반으로 하는 네트워크를 독립적으로 학습하여 두 종류의 보행 특징 맵을 추출하였으며, 각각의 분류 결과를 결합하여 최종적인 앙상블 네트워크의 분류 결과를 도출하였다. 20~30대 성인의 걷기, 뛰기, 빠르게 걷기, 계단 오르기과 내려가기, 언덕 오르기과 내려가기의 7종류의 보행에 대해, 스마트 인솔을 이용하여 실측한 멀티 센서 데이터를 제안한 앙상블 네트워크로 분류해 본 결과 90% 이상의 높은 분류율을 보이는 것을 확인하였다.

▶ **주제어:** 보행 타입, 딥러닝 네트워크, 앙상블 네트워크, 스마트 인솔, 멀티 모달 센서

-
- First Author: Hee-Chan Park, Second Author: Young-Chan Choi, Corresponding Author: Sang-Il Choi
 - *Hee-Chan Park (hc.park@alcherainc.com), Dept. of Computer Science and Engineering, Dankook University; AlcheraInc
 - **Young-Chan Choi (njs04288@gmail.com), Dept. of AI-based Convergence, Dankook University
 - ***Sang-Il Choi (choisi@dankook.ac.kr), Dept. of Computer Science and Engineering, Dankook University; Dept. of AI-based Convergence, Dankook University
 - Received: 2022. 09. 30, Revised: 2022. 11. 07, Accepted: 2022. 11. 15.

I. Introduction

보행은 인간의 기본적인 일상 행동으로서, 신체의 움직임에 대한 다양한 정보를 포함하고 있다. 보행의 패턴은 개인마다 고유의 특성을 보이기 때문에 생체 인식을 위한 정보로도 사용될 수 있으며, 보행의 종류를 구분하여 일상의 운동 패턴을 모니터링하거나 의료 진단 [1]이나 헬스케어서비스를 위한 용도로도 사용될 수 있다.

보행의 패턴을 분석하는 다양한 방법들이 제안 되어 왔다. [2]에서는 CCTV를 이용한 보행자 비디오 분석을 통해 보행을 분석하였고, 쓰레드밀(treadmill)을 이용하거나 [3] 압력 매트[4]를 이용한 연구들도 수행되었다. 이러한 방법들은 보행의 시각적 분석[3]이나 보폭[4]이나 발바닥의 압력 측정[5]을 정확하게 할 수 있는 장점이 있으나, 보행 측정의 장소와 시간 및 환경 조건에 제약이 있어 일상적인 보행 패턴의 분석에 한계가 있다.

최근에는 저전력 기술이나 센서 소형화 기술에 의해 웨어러블 디바이스 기술이 발전하면서 IMU(Inertial Measurement Unit)를 이용한 보행 분석기법들이 소개되었다. [6]에서는 여러개의 IMU를 이용하여 측정된 가속도 데이터와 자이로 데이터를 사용한 보행패턴 분류 모델을 제시하였고, [7]에서는 스마트폰 혹은 스마트워치에서 측정된 IMU 데이터를 이용하여 여러 보행을 포함한 Human Activity Recognition을 수행하는 모델을 제안하였다. 보행 측면에서는 보행 관련 데이터를 쉽게 얻을 수 있는 스마트 인soles이 상용화되기 시작하였고, 이에 따라 스마트 인soles을 이용한 보행 패턴 분류 관련 연구들이[8,9] 진행되었다.

여러 가지 센서들을 이용하여 수집한 데이터는 컴퓨팅 알고리즘을 통해 여러 가지 용도로 분석된다. 여러 가지 머신러닝(machine learning) 기법들을[8,10,11,12,13] 이용한 보행 분석이 시도되었으며, 최근 십여 년 동안 딥러닝 기술이 빠르게 발전하면서 보행 패턴 분석에 딥러닝 기술을 활용한 연구들도[16,17,18,19,21] 제안되었다.

본 논문에서는 여러 가지 타입의 보행(걷기, 뛰기, 빠르게 걷기, 계단 오르기, 계단 내려가기, 언덕 오르기, 언덕 내려가기)을 멀티 모달 센서(압력센서, 가속도센서, 자이로 센서)가 장착된 스마트 인sole로 측정하여 자동으로 보행의 종류를 분류하는 방법을 제안한다[17]. 보편적인 일상 용품인 인soles을 이용하여 측정한 보행 데이터는 보행의 장기적인 패턴을 분석할 수 있으며, 사용자의 신체에 별도의 장비를 부착하지 않기 때문에 자연스러운 일상의 보행 패턴의 분석에 용이하다.

압력센서 데이터의 경우, 센서가 장착된 위치에 따른 센서 신호 간의 상관관계를 활용하기 위해 공간 필터인 합성곱 필터를 기반으로 하는 CNN을 이용하여 효과적으로 보행 특징을 추출할 수 있다. 한편, 가속도 센서와 자이로 센서는 시계열 데이터의 특성을 갖기 때문에 순환신경망(Recurrent Neural Network, RNN)[18]를 이용하여 또 다른 성질의 특징을 추출할 수 있다. 이에 제안한 방법에서는 [8]에서 사용자 식별을 위해 설계된 앙상블 네트워크의 CNN과 대표적인 RNN 계열의 네트워크인 LSTM(Long Short Term Memory)[19]을 기본 네트워크로 사용하여 보행의 종류를 분류할 수 있는 앙상블 딥러닝 네트워크를 구성하였다.

제안한 방법의 성능을 확인하기 위해, 본 연구에서는 일상 생활에서 이루어지는 7종류(보통걸음, 빠른걸음, 뛰는 걸음, 오르막길걸음, 내리막길걸음, 계단오르기, 계단내려가기)의 대표적인 보행 종류에 대한 실험을 수행하였다. 20~30대 성인 14명을 대상으로 7종류의 보행을 실측한 데이터에 대한 실험을 통해 제안한 방법의 우수한 분류 성능을 확인하였다.

II. Related Works

1. Gait Analysis Using Machine Learning

여러 가지 머신러닝 기법을 이용한 보행 분석 방법들이 제안되어 왔다. [10]에서는 커널(Kernel)을 이용하여 측정된 데이터를 고차원의 특징 공간으로 변환(transform)한 후에 주성분 분석(Principal Component Analysis, PCA) 방법[20]을 적용하는 Kernel PCA(Kernel principal component analysis)기법[21]을 이용하여 보행 패턴의 특징을 추출하였다. [11]에서는 분류 문제 해결을 위해 제안된 영공간 선형 판별 분석(Nullspace LDA)[22]를 이용하여 사용자 분류를 위한 보행 데이터 분석을 제안하였다. [12]에서는 은닉 마르코프 모델(Hidden Markov Model, HMM)[23]를 IMU 센서 데이터에 적용하여 보행 구간(gait phase)를 구분하였고, [8,13]에서는 IMU 센서 데이터로부터 사용자를 인식하는데 HMM을 사용하였다.

2. Gait Analysis Using Deep Learning

센서 데이터로부터 딥러닝 기술을 이용하여 보행 패턴을 분석하기 위한 방법들도 제안되었다. [8]에서는 스마트 인sole에 장착된 압력센서, 가속도 센서, 자이로 센서를 이용하여 측정된 보행 데이터에 앙상블 딥러닝 네트워크를

적용하여 사용자를 식별하였고, [9]에서는 동일한 스마트 인솔 데이터에 보행의 종류를 구분하기 위한 합성곱(convolution) 기반의 네트워크(convolutional neural network, CNN)[14]을 제안하였다. [15]에서는 흉부와 허리, 오른쪽 손목, 무릎 및 발목에 각각 IMU 센서를 부착하여 보행 정보를 수집하였고, CNN 기반 예측 모델을 이용하여 사용자를 식별하였다[24]. 이 방법은 보행이 측정된 시간 도메인의 데이터와 그것을 주파수 도메인으로 변환한 데이터를 입력으로 사용하였다. [16]에서는 보행 정보를 수집하기 위해 스마트 폰에 설치된 IMU 센서를 사용하였다. 참가자가 바지 주머니에 스마트 폰을 넣고 걸으며 데이터를 수집한 다음 CNN의 혼합 모델과 SVM[25]를 사용하여 개인을 식별하였다.

III. The Proposed Scheme

제안한 방법은 크게 보행 데이터의 전처리 단계와 특징 추출 단계, 그리고 분류 단계로 구성되어 있다. 보행은 연속적인 동작으로 이루어지는데, 보행 데이터를 샘플 데이터로 저장하기 위해서는 데이터를 일정 구간으로 분할해야 한다. 보행 주기는 스윙 구간(swing phase, SW)와 지면 구간(stance phase, ST)으로 구분되는데[26], 전처리 단계에서는 보행 데이터를 SW와 ST를 기준으로 분할하여 걸음 단위의 샘플 데이터를 생성한다[8]. 특징 추출부를 구성하고 있는 CNN과 LSTM은 전처리 단계를 통해 생성된 각 센서별(압력 센서, 가속도 센서, 자이로 센서) 보행 데이터를 가지고 독립적으로 학습된다.

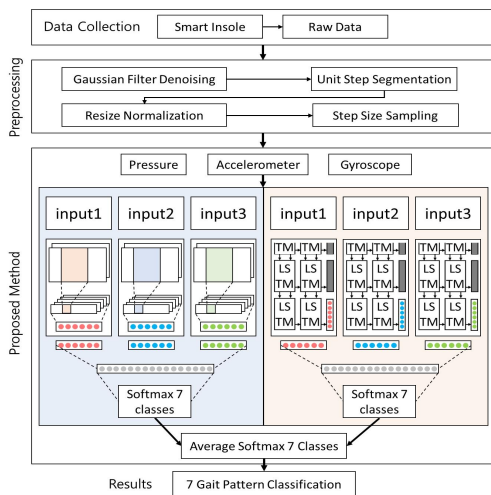


Fig. 1. Overall Structure

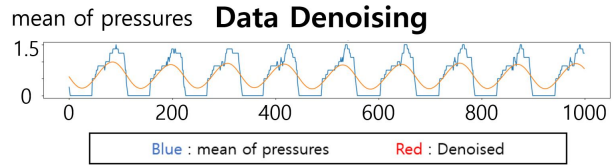


Fig. 2. Data Denoising

그런 다음, 네트워크로부터 출력된 CNN과 LSTM의 특징 맵(feature map)은 완전연결 네트워크(fully connected network)로 구성된 분류기를 통해 Softmax[27] 형태로 결과를 출력하고, CNN과 LSTM의 각 출력에 대한 평균 Softmax를 통해 최종적으로 보행의 종류를 결정한다. Fig. 1은 제안한 보행 종류 분류 방법의 전체 흐름도이다.

1. Data Pre-processing

사람의 보행은 시간에 따라 연속적인 동작이 반복되는 주기적인 특성을 갖는다. 보행 주기는 한쪽 발의 움직임을 기준으로 크게 2개의 구간으로 구분하고[3, 4], 각 구간은 다시 사람의 관절, 몸통 위치, 발 뒤꿈치 등의 변화에 따라 'heel strike', 'foot flat', 'mid stance', 'heel off', 'toe off', 'mid swing', 'late swing' 구간으로 구분된다. 한쪽 발을 기준으로, ST는 'heel strike'에서부터 'toe off'에 해당한다. heel strike'는 발 뒤꿈치가 지면에 닿기 시작하는 순간을 말하며 보행 주기의 시작이다[28]. 'Foot Flat'은 발바닥이 지면에 닿게 되는 순간이고, 'mid stance'는 발이 지면과 완전히 닿아서 다리가 지면과 수직인 상황이다. 그리고 이때 지면과 떨어져 있던 반대 다리가 앞으로 움직인다. 'heel off'에서 발 뒤꿈치부터 지면에서 떨어지면서 ST 구간의 마지막 순간인 'Toe off'에서 발 대부분이 지면에서 떨어진다. SW 구간에서는 발이 공중에 떠 있는 상태로 있는데, 'Mid Swing'은 지면과 떨어진 발이 'mid stance'에서 반대 다리가 지면과 수직일 때, 앞으로 나아가는 과정을 포함한다. 마지막 'Late Swing'에서는 바로 다음 ST 구간이 시작되기 전까지 발이 지면과 닿기 전 상황을 말한다.

이러한 보행 주기의 특성은 IMU (inertial measurement units) 센서 데이터에서도 나타나지만 [5] SW 구간과 ST 구간이 명확하게 구분될 수 있는 압력 센서 데이터에서 보다 확연하게 관찰된다. 제안한 방법에서는 보행 주기에서 나타나는 압력 센서 데이터의 특성을 바탕으로 연속적인 보행 데이터로부터 단위 걸음 구간을 자동으로 분할하였다[8]. 압력 센서는 발과 지면 사이의 압력을 측정하므로 SW 구간에서는 인솔에 장착된 8개의 압력

센서의 값이 0이 된다. Fig. 2에서의 파란색 값은 왼쪽 발을 기준으로 보통걸음에 대한 8개의 압력 센서 값들의 평균값을 의미한다. Fig. 2에서 압력 센서의 평균이 0이 되는 SW 구간이 주기적으로 나타나는 것을 볼 수 있는데, SW 구간이 시작되는 시점, 즉 압력 센서의 평균이 0이 아닌 상태에서 0으로 바뀌는 시점을 기준으로 단위 걸음을 분할 할 수 있다 [8, 11]. 그런데, 데이터를 실측 시에 인솔의 온도가 올라가거나 센서 간의 간섭 등의 예기치 않은 요인들로 인해 종종 SW 구간임에도 특정 센서의 값이 0이 되지 않는 경우가 발생할 수 있다. 이러한 센서 값 오류들의 영향을 줄이고 보다 정확하고 강인하게(robustly) 단위 걸음을 분할하기 위해, 본 연구에서는 식(1)의 가우시안 필터(Gaussian filter)($y(t)$)를 이용하여 단위 걸음에 해당하는 주기의 특징을 추출하였다[8]. 필터의 표준편차는 [8]에서와 같이 σ 는 0.2로 설정하였다.

$$y(t) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{t^2}{2\sigma^2}} \text{ where } \sigma = 0.2 \quad (1)$$

t 시점(time point)에서의 8개 압력 센서 값의 평균을 $x^p(t)$ 라고 할 때, 가우시안 필터링 된 결과 $z(t)$ (Fig. 4에서의 빨간색 선)는 식(2)와 같이 합성 곱 연산을 통해 구할 수 있다.

$$z(t) = (x^p(t) * y)(t) = \int_0^t x^p(\tau)y(t-\tau)d\tau \quad (14)$$

$z(t)$ 에서 단위 걸음은 $\frac{d}{dt}z(t) = 0$ 이면서 $\frac{d^2}{dt^2}z(t) > 0$ 이 되는 시점들을 기준으로 분할하였다[8].

Fig. 3는 보행 데이터를 단위걸음($s_i(t)$)으로 분할한 것을 시각화한 것이다. Fig. 3의 각 보행 타입 별 데이터에서, 가로축은 왼쪽 발과 오른쪽 발의 8개 센서(총 16개)의 측정값(0~2)이고, 세로축은 샘플링 된 시간 지점이다. 녹색 선은 $z(t)$ 의 미분 값을 기준으로 구분한 단위 걸음 경계선이다. Fig. 3에서, 같은 보행자가 동일한 타입의 보행을 하고 있음에도 각각의 단위 걸음의 길이가 다른 것을 볼 수 있다. 이처럼 같은 보행 종류에 대한 샘플마다 보행 속도가 다른 것은 보행의 종류를 구분하는 데에 오히려 방해 요소가 될 수 있다. 이에, 본 연구에서는 단위 걸음별로 분할된 $s_i(t)$ 들을 spline interpolation[29]을 이용하여 일정한 길이(d)로 (standard length) 리사이징 (resizing) 한 다음에 압력센서 값은 [0, 1], 가속도 센서, 자이로 센서값은 [-1, 1] 사이의 값을 갖도록 Min Max Scaler를 이용해 정규화하였다. 정규화 길이 d 는 분할된 $s_i(t)$ 중에서 가장 짧은 길이(여기서는 $d=63$)을 기준으로 설정하였다.

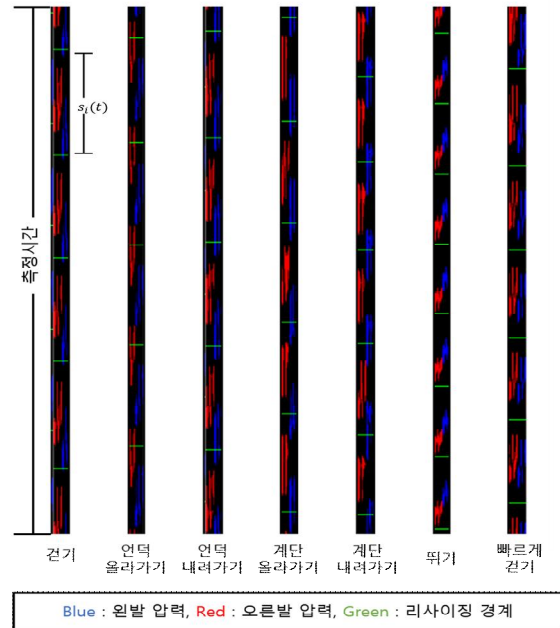


Fig. 3. Visualization before applying normalization

추가로, 딥러닝 네트워크가 보행 타입의 특징을 학습하기 위한 데이터 샘플을 몇 걸음(k)으로 구성하는 것이 적절한지를 확인하기 위해, 한 걸음에서부터 다섯 걸음까지 포함된 데이터 샘플들을 생성하였다 ($1 \leq k \leq 5$). 결과적으로 양 발에 대한 압력 센서 어레이, 3축 가속도 센서 어레이, 3축 자이로 센서 어레이로 측정하여 정규화 한 단위 걸음 샘플 $\mathbf{x}_p^l(k)$, $\mathbf{x}_a^l(k)$, $\mathbf{x}_r^l(k)$ 의 크기는 k 에 따라 각각, $(d \cdot k) \times (8 \cdot 2)$, $(d \cdot k) \times (3 \cdot 2)$, $(d \cdot k) \times (3 \cdot 2)$ 가 된다.

2. Design of Multi-modal Deep Learning Ensemble Network for Gait Type Classification

여러 종류의 센서 데이터를 이용하는 멀티 모달 학습은 멀티 모달 데이터를 한꺼번에 입력으로 사용하여 하나의 네트워크를 학습하는 방법과 모달 별로 네트워크를 학습한 후 각각의 결과를 결합하는 방법이 있다[30]. 제안한 방법에서는 이 두가지 접근법을 혼용하였다. 즉, 앙상블 네트워크를 구성하고 있는 CNN과 LSTM은 각각 독립적으로 학습하고 각 네트워크의 결과를 결합하여 최종 결과를 도출하였으나, CNN와 LSTM의 학습은 압력센서, 가속도 센서, 자이로 센서 데이터를 한꺼번에 입력으로 사용하여 종단 간(end-to-end) 학습을 수행하되, 내부 구조는 센서 별로 구분된 흐름(stream)을 별도로 갖도록 구성하였다 (Fig. 1).

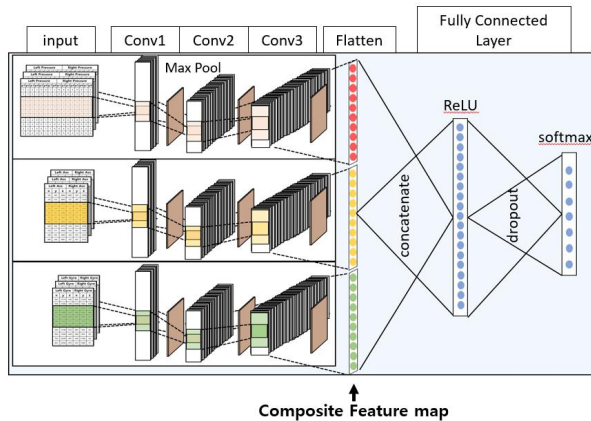


Fig. 4. Architecture of multi-modal deep CNN

2.1 Preliminaries

일반적인 CNN 기반의 분류 네트워크는 특징맵을 추출하는 합성곱 레이어(convolution layer)들과 분류기 역할을 하는 완전 연결 레이어(fully connected layer)로 구성된다. 그리고 합성곱 레이어를 세분화 하면 필터 레이어(filter layer)와 활성화 함수 레이어(non-linearity activation function layer), 그리고 풀링 레이어(pooling layer)로 구성되어 있다[31]. 합성곱 레이어는 필터를 이동하면서 입력 데이터에 적용하여 특징을 추출하는 방법으로, 지역적인 특성이 반영된 특징을 추출한다. 본 논문에서는 합성곱 레이어를 기존의 2차원이 아닌 1차원 데이터에 적용하였다. 이러한 1차원 합성곱 레이어의 기존에 사용되던 합성곱 레이어와 유사하지만 2차원 입력 데이터에서 필터를 이동하며 적용하는 대신 1차원 입력 데이터에서 시간 축을 기준으로 필터를 이동하면서 특징을 추출한다.

RNN은 데이터를 시간 순서대로 받아들여 메모리 셀에 저장된 데이터와 연산 후 다시 메모리 셀에 저장하는 형태의 네트워크이다. RNN은 그 구조로 인해 시계열 데이터 분석에서 순차적인 정보를 담은 고유의 특징을 추출할 수 있다. LSTM은 RNN에 게이트를 추가한 구조로, RNN에서 발생하는 기울기 소멸 문제(vanishing gradient problem)을 방지하기 위해 forget gate, input gate, output gate를 추가한 네트워크이다.

2.2 CNN Based Multi-modal Network

Fig. 4은 본 논문에서 사용한 멀티 모달 CNN 기반의 딥러닝 분류 네트워크의 구조를 보여준다. Fig. 4의 CNN은 3개의 합성곱 레이어와 각 합성곱 레이어는 합성곱 필터와 활성화 함수로 구성되었다. 네트워크의 학습을 위한 오차 역전파(error back-propagation) 알고리즘을 수행하는 과정에서 발생 할 수 있는 그레디언트 손실(vanishin

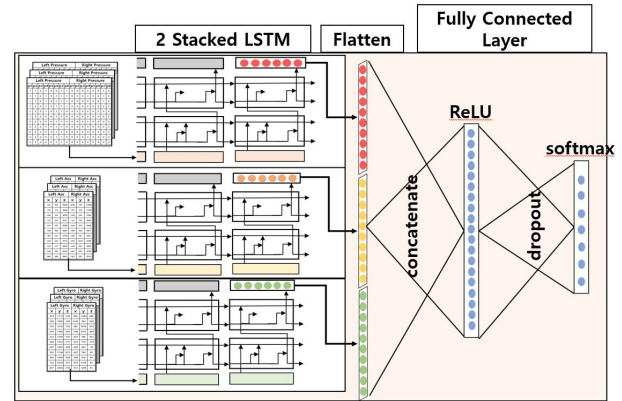


Fig. 5. Architecture of multi-modal LSTM

g gradient) 현상[32]을 방지하기 위해 활성화 함수는 ReLU(Rectifier Linear Unit)[33]을 사용하였다. 필터의 합성곱 연산 시 이동간격(stride)은 $k \leq 4$ 일 때는 1로, 그리고 $k=5$ 일 때는 2로 설정하였다. 첫 번째 합성곱 레이어에서는 $w \times 20$ 크기의 32개의 필터를 사용하였고(w 는 x_t^p 에 대해 16, x_t^a 와 x_t^v 에 대해서는 6), 두 번째와 세 번째 합성곱 레이어에서는 앞선 레이어에서 출력되는 특징 맵의 크기에 따라 각각 20×32 와 20×64 크기의 필터를 각각 64개와 128개를 사용하였다. Fig. 4에서와 같이, 각각의 센서마다 개별적인 네트워크를 구성하고 각 네트워크에서 출력된 특징 맵을 연결(concatenating)하여 256차원 벡터 형태의 복합 특징 맵(composite feature map)을 구성하였다.

2.3 LSTM Based Multi-modal Network

보행 데이터는 시간에 따른 연속적인 동작에 대한 측정 값이기 때문에 시간적 문맥(temporal context) 정보를 갖고 있다. 이에 제안한 방법에서는 CNN과 함께 시계열 분석에 유용한 모델인 LSTM을 사용하였다. 멀티 모달 CNN 기반의 분류 네트워크에서와 같이 멀티 모달 LSTM 기반의 분류 네트워크도 각 센서 데이터별로 동일한 구조의 개별적인 네트워크를 구성하였다[8] (Fig. 5). 센서별로 개별 네트워크들은 특징 추출을 위한 두 개의 연속적인 LSTM 레이어로 구성하였다. 각각의 LSTM 레이어들은 64개의 메모리 유닛을 가지며 출력 값의 형태는 many-to-many 형식으로 첫 번째 LSTM 레이어의 출력 값이 두 번째 LSTM 레이어의 시간적 관계를 갖는 입력으로 사용되도록 하였다. 과적합(overfitting) 문제를 피하고자, 순환 드롭아웃 비율(recurrent dropout ratio)은 0.2로 설정하였다. 첫 번째 LSTM 레이어에 대한 입력의 크기는 k 에 따라 x_t^p , x_t^a , x_t^v 에 대해 각각, $(d \cdot k) \times 8 \cdot 2$, $(d \cdot k) \times (3 \cdot 2)$, $(d \cdot k) \times (3 \cdot 2)$ 이고, 각 메모리 유닛

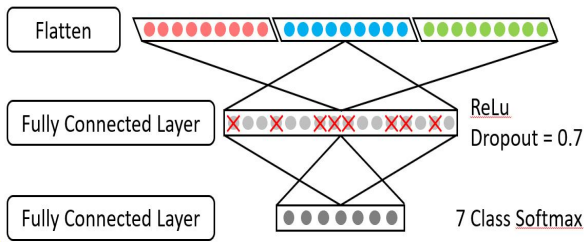


Fig. 6. Fully connected layer

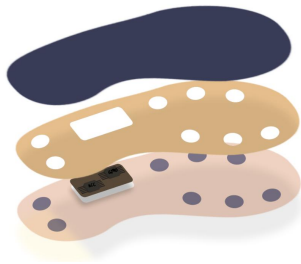


Fig. 7. Picture of Smart Insole

마다 출력된 스칼라 값들이 연결되어 출력되기 때문에 출력의 크기는 $(d \cdot k) \times 64$ 가 된다. 두 번째 LSTM 레이어에서의 출력 값은 many-to-one 형식으로 메모리 유닛 당 마지막 값만을 출력하기 때문에 64차원 벡터 형태의 특징 맵을 구성한다. 멀티 모달 CNN기반의 분류 네트워크에서와 마찬가지로 각각의 센서마다 개별적인 네트워크를 구성하고 각 네트워크에서 출력된 특징 맵을 연결(concatenating)하여 256차원 벡터 형태의 복합 특징 맵을 구성하였다.

2.4 Classification Network

앞서 언급한 대로, 멀티 모달 CNN 기반의 분류 네트워크와 LSTM 기반의 분류 네트워크는 각각 독립적으로 학습된다. 보행 종류의 분류를 위한 각 네트워크는 완전 연결 네트워크(fully connected network, FCN)를 갖는다(Fig. 6). FCN은 완전 연결 레이어와 softmax 레이어로 구성되어 있다. 완전 연결 레이어는 256개의 노드(node)로 구성되어 있으며 활성화 함수로는 ReLU를 사용하였다. 과적합(over-fitting) 문제를 방지하고 일반화(regularization) 성능을 향상시키기 위해 학습 시에 드롭아웃(dropout)[34]을 적용 하였으며 실험적 경험을 바탕으로 드롭아웃 비율은 0.7로 설정하였다. 각 FCN은 센서 별로 출력된 특징 맵을 연결한 복합 특징 맵을 입력으로 받고, 분류하고자 하는 보행 종류의 개수 만큼(본 논문에서는 7개)의 노드를 포함한 Softmax 레이어에서 각 노드에 대한 정규화 된 확률을 계산한다. 이렇게 멀티 모달 CNN 기반의 분류 네트워크와 LSTM 기반의 분류 네트워크

로부터 구한 2개의 Softmax 함수 결과 값은 Average Softmax를 이용하여 최종 보행 분류 결과를 산출한다.

IV. Experiments

1. Data Acquisition

본 논문에서 보행 데이터의 수집을 위해 사용한 스마트 인솔은 한쪽 발에 8개의 압력 센서와 3축 가속도 센서, 그리고 3축 자이로 센서를 내장하고 있다[9] (Fig. 7). 각 센서들은 100Hz의 샘플링율로 데이터를 측정하며, 한 시점에 양 쪽 인솔로부터 압력 센서는 16차원의 어레이(array)로, 가속도 센서와 자이로 센서는 각각 6차원의 어레이로 데이터를 저장한다(Fig. 8). 압력 센서의 경우, 센서가 장착된 발바닥 부위의 압력을 0, 1, 2의 값으로 측정하는데, 0은 압력이 없는 상태, 즉 SW 구간(지면에서 발이 떨어진 상태)이고, 1과 2는 ST 구간(발이 지면을 디디고 있는 상태)에서의 압력의 세기를 의미한다. 가속도 센서와 자이로 센서는 각각 가속도와 회전을 측정하는데 센서 값들은 모두 3D 공간에서 $-32,768$ 에서 $32,768$ 사이의 정수로 기록된다. 취득된 데이터는 블루투스 통신을 통해서 안드로이드 스마트폰 어플리케이션을 이용하여 데이터베이스 서버로 전송된다[9].

보행 데이터는 14인의 20~30대 성인이 동일한 조건에서 평지, 언덕, 계단에서 측정하였으며, 보통 걸음, 빠른 걸음, 뛰는 걸음은 3분 동안, 오르막길 걸음과 내리막길 걸음은 2분, 계단 오르기, 계단 내려가기는 1분 동안 측정하였다. Table 1은 본 연구에서 사용한 보행 데이터에 대한 정보를 보여준다.

2. Experimental Setting

실험에 사용한 환경은 CPU: intel i5-10400, GPU: GTX 1070 8GB, 메모리: 16GB로 구성된 컴퓨터에서 실험을 수행하였다. 스마트 인솔을 이용하여 측정한 14명에 대한 7종류의 보행 데이터는 앞서 III-1에서 기술한 전처리 방법을 이용하여 단위 걸음으로 분할하여 데이터 샘플을 생성하였다. 보행의 종류별 샘플 수는 $k=1$ 을 기준으로 714개에서 3612개이며, k 에 따른 샘플 수는 Table 2에 제시되어 있다.

학습과 테스트에 필요한 데이터는 전체 데이터 샘플 중에서 무작위로 각각 1000개씩을 샘플링하여 구성하였다. 성능 평가의 신뢰도를 확보하기 위해 이러한 과정을 20번 반복 수행하였으며, 20번의 테스트에서의 평균 분류율을

p1(L)	p2(L)	p3(L)	p4(L)	p5(L)	p6(L)	p7(L)	p8(L)	ACC X(L)	ACC Y(L)	ACC Z(L)	GYRO X(L)	GYRO Y(L)	GYRO Z(L)
0	1	2	2	0	2	2	2	-227	249	-8624	-400	-2384	97
0	1	2	2	0	2	2	2	-264	280	-8600	-368	-2534	93
0	1	2	2	0	2	2	2	-387	281	-8588	-261	-2612	45
0	1	2	2	0	2	2	2	-369	302	-8510	-189	-2567	-68
0	1	2	2	0	2	2	2	-470	288	-8491	-98	-2483	-180
0	1	2	2	0	2	2	2	-533	242	-8433	-47	-2345	-250
0	2	2	2	0	2	2	2	-591	226	-8426	-12	-2172	-259
0	2	2	2	0	2	2	1	-636	173	-8425	-20	-1989	-260
0	2	2	2	0	2	2	1	-696	179	-8424	-34	-1853	-265
0	2	2	1	0	2	1	1	-685	182	-8433	-74	-1770	-251
0	2	2	1	0	2	1	1	-736	271	-8473	-90	-1707	-259
1	2	2	1	0	2	1	0	-726	198	-8451	-40	-1690	-282
1	2	2	1	0	2	1	0	-768	208	-8546	-58	-1728	-258
1	2	2	1	0	2	1	0	-797	245	-8558	-96	-1764	-311
1	2	2	1	0	2	1	0	-844	313	-8530	-143	-1807	-395
1	2	2	1	0	2	1	0	-886	278	-8526	-220	-1813	-467
1	2	2	1	0	2	2	0	-965	291	-8531	-255	-1858	-521
1	2	2	0	0	2	2	0	-1056	254	-8558	-282	-1917	-566
1	2	2	0	0	2	2	0	-977	198	-8617	-321	-1978	-547
1	2	2	0	1	2	2	0	-977	240	-8732	-444	-2055	-556
1	2	2	0	1	2	2	0	-1027	267	-8991	-677	-2235	-574
1	2	2	0	1	2	1	0	-986	203	-9441	-1075	-2807	-533
2	2	2	0	1	2	1	0	-1091	-39	-9839	-877	-3681	-397
2	2	2	0	1	2	1	0	-1060	363	-9754	-532	-4494	-249
2	2	2	0	2	2	1	0	-1098	218	-9961	-297	-5283	-314
2	2	2	0	2	2	1	0	-1203	262	-10207	-233	-5928	-272
2	2	2	0	2	2	1	0	-1668	303	-9713	-662	-6749	-258
2	2	2	0	2	2	1	0	-1577	343	-9664	-819	-7423	-246
2	2	2	0	2	2	1	0	-1968	299	-11573	-1369	-8744	-336
2	2	2	0	2	2	1	0	-1114	1072	-11465	-1572	-9932	-621
2	2	2	0	2	2	0	0	-1911	1271	-10760	-1456	-11226	-926
2	2	2	0	2	2	0	0	-1807	1228	-11241	-1179	-12634	-1251
2	2	2	0	2	2	0	0	-2006	1366	-11894	-721	-14078	-1468
2	2	2	0	2	2	0	0	-2435	1962	-13899	642	-15831	-1900
2	2	2	0	2	2	0	0	-1857	4161	-13849	1309	-18076	-2568
2	2	2	0	2	2	0	0	-2269	4127	-14239	1117	-20511	-3976
2	2	2	0	2	2	0	0	-2338	3358	-14216	2418	-22371	-5021
2	2	2	0	0	2	0	0	-2545	3173	-16350	3944	-24469	-5637
2	2	2	0	0	2	0	0	-4623	3791	-17581	5357	-26034	-5781
2	2	2	0	0	1	0	0	-6692	4955	-25347	3620	-27582	-5721
2	2	2	0	0	1	0	0	-5625	4339	-32768	2196	-27510	-7194
2	2	1	0	0	1	0	0	920	5578	-29952	3918	-30988	-7115
2	1	1	0	0	1	0	0	-3971	4213	-25920	4263	-28428	-9126
1	1	1	0	0	0	0	0	5496	11174	-24353	6911	-30621	-9541

Fig. 8. Measure Gait Data (raw data)

Table 1. Number of data samples

Number of Total Samples	Walk	Up Hill	Down Hill	Up Stair	Down Stair	Running	Fast Walk
13,308	2,257	1,539	1,552	714	954	3,612	2,680

Table 2. Number of samples for each unit step

Unit Step(k)	1	2	3	4	5
Number of Samples	13,308	6,654	4,436	3,327	2,661

최종 성능으로 제시하였다. 네트워크의 학습은 역전파 (Back-propagation) 학습 알고리즘[35]을 이용하였다. 또한 ‘Internal Covariance Shift’ 현상[19]을 방지하고 학습의 안정성을 높이기 위해 배치 활성화 함수를 적용하고 배치 정규화[36]을 수행하였다[9]. 배치의 크기(batch size)는 64로 설정하였다. 네트워크의 가중치 업데이트 최적화를 위한 방법으로 아담 옵티마이저 (Adaptive Momentum Optimizer)[37]을 사용하였으며 학습률 (learning rate)은 0.0001로 설정하였다. 상기 하이퍼 파라미터(hyper-parameter)들은 실험적 경험을 바탕으로 선택하였다. CNN 및 RNN 기반 멀티 모달 분류 네트워크의 손실함수로는 ‘Categorical Cross Entropy’를 사용하였다.

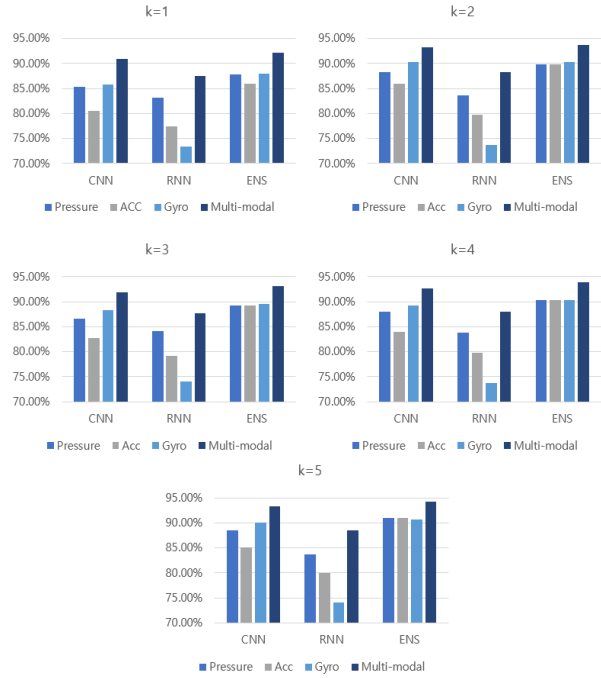


Fig. 9. Classification rates of CNN based-, LSTM based-, ensemble deep learning network

3. Experimental Results

본 실험에서는 먼저 단일 센서 데이터를 사용했을 때와 멀티 센서 데이터를 사용했을 때 CNN 기반 네트워크를 단독으로 사용했을 때[9]와 LSTM 기반 네트워크를 단독으로 사용했을 때의 분류 성능을 비교하고, 단일 센서 데이터에 대한 앙상블 딥러닝 네트워크의 성능과 멀티 센서 데이터에 대한 앙상블 네트워크의 성능을 평가해 보았다. 또한 k 를 늘려가며 분류 성능을 평가함으로써 앙상블 딥러닝 네트워크를 이용하여 효과적으로 보행 타입을 분류하는데 필요한 걸음 수를 살펴보았다.

Fig. 9는 단일 센서와 멀티 센서, 그리고 CNN 기반 네트워크와 LSTM 기반 네트워크, 앙상블 네트워크(ENS)에 대해 k 값(한 샘플에 포함된 걸음 수)을 다르게 했을 때의 분류 결과를 보여준다. Fig. 9에서 전반적으로 CNN 기반의 네트워크가 LSTM 기반의 네트워크보다 좋은 성능을 보여 주었다. 실제 보행은 발 뒤꿈치부터 지면에 닿은 다음에 앞꿈치가 닿는 순서의 연속적인 동작이고, 발바닥에 각각의 위치에 부착된 압력 센서 값들도 보행 주기와 관련이 있기 때문에 데이터들 간의 상관관계를 활용하는 CNN 기반 네트워크가 보다 효과적으로 보행 특징을 추출하는 것으로 판단된다.

또한 Fig. 9의 결과는 단일 센서 데이터를 사용한 것보다 여러 종류의 센서를 함께 사용한 멀티 모달 데이터가 ($x_i^{p,a,r}(k)$) 더 우수한 분류성능을 보여준다. 이는 센서의

종류가 많을수록 보행 움직임에 대한 정보가 많아지는 데 따른 결과이다. 전체적으로 볼 때, 멀티 센서 데이터에 대해 앙상블 네트워크를 사용했을 때 가장 좋은 성능을 보였는데, 이는 멀티 모달 데이터에 대해 각기 다른 속성을 갖는 네트워크를 통해 추출한 특징들이 상호 보완적인 역할을 함으로써 시너지 효과를 보인 것으로 판단된다.

한편, 데이터 샘플에 포함된 걸음수(k)에 따른 성능을 평가해 보았을 때, k 가 증가할수록 분류율이 증가하다가 '3~5 걸음수'에서 증가의 폭이 둔화되거나 소폭 감소하였다. 이는 네 걸음 정도만 확보되면 보행 패턴의 특징을 효과적으로 추출할 수 있음을 의미한다.

V. Conclusions

본 논문에서는 스마트 인솔에 장착된 멀티 센서들(압력, 가속도, 자이로)을 이용하여 측정된 데이터로부터 앙상블 딥러닝 네트워크를 이용하여 보행의 타입을 분류하는 방법을 제안하였다.

제안한 방법은 크게 데이터 전처리 부분과 딥러닝 네트워크에 의한 보행의 특징 추출 부분, 그리고 보행의 타입을 분류하는 부분으로 구성되어 있다. 보행의 연속 동작을 측정할 센서 데이터는 보행 주기에서의 SW 구간을 기준으로 단위 걸음으로 분할되어 네트워크의 입력으로 사용되었다. 특징 추출 부분은 서로 다른 성질을 갖는 CNN과 LSTM을 기반으로 하는 네트워크를 각각 독립적으로 학습하였으며, 각 네트워크의 분류 결과를 average softmax를 이용하여 최종 분류 결과를 도출하였다. 성인 14명에 대해 7종류의 보행을 실측한 데이터를 가지고 분류 실험을 수행한 결과, 한 종류의 센서를 사용하는 것보다 멀티 센서를 사용했을 때의 분류성능이 좋았으며, CNN 또는 LSTM 기반의 네트워크를 단독으로 사용했을 때보다 두 종류의 네트워크의 결과를 함께 사용했을 때의 가장 높은 분류성능을 보여주었다. 또한 한걸음에 대한 데이터보다 여러 걸음을 모아 데이터 샘플을 구성하였을 때 성능이 향상되는 것을 확인할 수 있었다.

추후, 시간과 공간에 대한 문맥 정보를 활용할 수 있는 다양한 종류의 네트워크를 설계함으로써 보행 데이터에 내포된 유용한 정보들을 보다 효과적으로 추출할 수 있는 방법에 대한 연구를 진행할 계획이다.

ACKNOWLEDGEMENT

This work was supported by the National Research Foundation of Korea(NRF) grand funded by the Korea government(MSIT) (No. 2021R1A2B5B0100141212) and Institute of Information & communications Technology Planning & Evaluation(IITP) grant funded by the Korea government(MSIT) (No.2022-0-00899 Development of a gait pattern analysis system using smart insoles equipped with multi-modal sensors).

REFERENCES

- [1] E. Smirni, and G. Ciardo, "Workload-Aware Load Balancing for Cluster Web Servers," IEEE Trans. on Parallel and Distributed Systems, Vol. 16, No. 3, pp. 219-232, March 2005. DOI: 10.1109/TPDS.2005.38
- [2] Kdhong, "An Efficient Dynamic Workload Balancing Strategy," Journal of The Korea Society of Computer and Information, Vol. 15, No. 1, pp. 1-10, Nov. 2010.
- [3] Kdhong, "C Programming Language" Korea-Press, pp.100-120, 1991.
- [4] Q. Deng, Y. Luo, and J. Ge, "Dual threshold based unsupervised face image clustering," Proceedings of the 2nd International Conference on Industrial Mechatronics and Automation, pp. 436-439, 2010.
- [5] SIMGRID Project, <http://simgrid.gforge.inria.fr>
- [6] Q. Deng, Y. Luo, and J. Ge, "Dual threshold based unsupervised face image clustering," Proceedings of the 2nd International Conference on Industrial Mechatronics and Automation, pp. 436-439, 2010.
- [7] Q. Deng, Y. Luo, and J. Ge, "Dual threshold based unsupervised face image clustering," Proceedings of the 2nd International Conference on Industrial Mechatronics and Automation, pp. 436-439, 2010.
- [8] Jucheol Moon, Nelson Hebert Minaya, Nhat Anh Le, Hee-Chan Park, Sang-Il Choi(Corresp. Author), "Can Ensemble Deep Learning Identify People by Their Gait Using Data Collected from Multi-Modal Sensors in Their Insole?", MDPI Sensors, (1424-8220, 2.457), 2020.
- [9] Sung-Sin Lee, Sang Tae Choi, Sang-Il Choi(Corresp. Author), "Classification of Gait Type Based on Deep Learning Using Various Sensors with Smart Insole", MDPI Sensor, (1424-8220, 2.457), 2019.
- [10] Wu, Jianning, Jue Wang, and Li Liu. "Feature extraction via KPCA for classification of gait patterns." Human movement

- science 26.3 (2007): 393-411.
- [11] Seo, Woo-Duk, et al. "Gait type classification using pressure sensor of smart insole." *Journal of The Korea Society of Computer and Information* 23.2 (2018): 17-26.
- [12] Yan, Lei, et al. "Walking Gait Phase Detection Based on Acceleration Signals Using Voting-Weighted Integrated Neural Network." *Complexity* 2020 (2020).
- [13] Bonnet, Stéphane, and Pierre Jallon. "Hidden markov models applied onto gait classification." 2010 18th European Signal Processing Conference. IEEE, 2010.
- [14] Y. Lecun, L. Bottou, Y. Bengio and P. Haffner, "Gradient-based learning applied to document recognition," in *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278-2324, Nov. 1998, doi: 10.1109/5.726791.
- [15] Taborri, Juri, et al. "A novel HMM distributed classifier for the detection of gait phases by means of a wearable inertial sensor network." *Sensors* 14.9 (2014): 16212-16234.
- [16] Davis III, Roy B., et al. "A gait analysis data collection and reduction technique." *Humanmovement science* 10.5 (1991): 575-587.
- [17] Hee Chan Park, "Classification of Gait Pattern Based on Ensemble Learning Using Smart Insole," M.S. Thesis, Dankook University, Korea, Feb. 2021.
- [18] Williams, Ronald J.; Hinton, Geoffrey E.; Rumelhart, David E. (October 1986). "Learning representations by back-propagating errors". *Nature*. 323 (6088): 533–536.
- [19] Hochreiter, S., & Schmidhuber, J"urgen. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780.
- [20] Karl Pearson F.R.S. . (1901). LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2(11), 559–572.
- [21] B. Schölkopf, A. Smola, and K-R. Müller. "Nonlinear component analysis as a kernel eigenvalue problem." *Neural Computation*, 10:1299-1319, 1998. Technical Report No. 44, 1996, Max Planck Institut für biologische Kybernetik, Tübingen.
- [22] Cevikalp, Hakan, et al. "Discriminative common vectors for face recognition." *IEEE Transactions on pattern analysis and machine intelligence* 27.1 (2005): 4-13.
- [23] L. Rabiner and B. Juang, "An introduction to hidden Markov models," in *IEEE ASSP Magazine*, vol. 3, no. 1, pp. 4-16, Jan 1986, doi: 10.1109/MASSP.1986.1165342.
- [24] Choi, Sang-Il, et al. "User identification from gait analysis using multi-modal sensors in smart insole." *Sensors* 19.17 (2019): 3785
- [25] Shetty, Sachin, and Y. S. Rao. "SVM based machine learning approach to identify Parkinson's disease using gait analysis." 2016 International Conference on Inventive Computation Technologies (ICICT). Vol. 2. IEEE, 2016
- [26] Murray, M.P.; Drought, A.B.; Kory, R.C. Walking patterns of normal men. *J. Bone Jt. Surg.* 1964, 46, 335–360.
- [27] Bridle, John. "Training stochastic model recognition algorithms as networks can lead to maximum mutual information estimation of parameters." *Advances in neural information processing systems* 2 (1989).
- [28] Charalambous, Charalambos P. "Walking patterns of normal men." *Classic Papers in Orthopaedics*. Springer, London, 2014. 393-395.
- [29] Schoenberg, Isaac J. (1946). Contributions to the Problem of Approximation of Equidistant Data by Analytic Functions: Part A.—On the Problem of Smoothing or Graduation. *A First Class of Analytic Approximation Formulae*. *Quarterly of Applied Mathematics*. 4 (2): 45–99.
- [30] E. Alpaydin, "Multiple networks for function learning," *IEEE International Conference on Neural Networks*, 1993, pp. 9-14 vol.1, doi: 10.1109/ICNN.1993.298539.
- [31] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. 2012. ImageNet classification with deep convolutional neural networks. In *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1 (NIPS'12)*. Curran Associates Inc., Red Hook, NY, USA, 1097–1105.
- [32] Hochreiter, Sepp. (1998). The Vanishing Gradient Problem During Learning Recurrent Neural Nets and Problem Solutions. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*. 6. 107-116. 10.1142/S0218488598000094.
- [33] Agarap, A. F. (2018). Deep learning using rectified linear units (relu). *ArXiv Preprint ArXiv:1803.08375*.
- [34] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. 2014. Dropout: a simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* 15, 1 (January 2014), 1929–1958.
- [35] Kelley, H. J. (1960). Gradient theory of optimal flight paths. *Ars Journal*, 30(10), 947–954
- [36] Ioffe, S. & Szegedy, C.. (2015). Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. *Proceedings of the 32nd International Conference on Machine Learning in Proceedings of Machine Learning Research*
- [37] Kingma, Diederik & Ba, Jimmy. (2014). Adam: A Method for Stochastic Optimization. *International Conference on Learning Representations*.

Authors



Hee-Chan Park received the B.S., degree in Applied Computer Engineering and M.S. degree in Computer Science from Dankook University, Korea, in 2019 and 2021, respectively.

He joined Alchera Inc as a researcher in 2021. His research interests include machine learning and pattern recognition and deep learning.



Young-Chan Choi received the B.S degrees in Applied Computer Science from Dankook University, Korea, in 2022. He is currently pursuing the M.S. degree in the Department of Data Science from Dankook University.

His research interests include machine learning and pattern recognition and deep learning.



Sang-II Choi received the B.S. degree in the Division of Electronic Engineering from Sogang University, Seoul, Korea, in 2005, and M.S, Ph.D degrees in Electrical Engineering and Computer Science from

Seoul National University, Seoul, Korea in 2007, 2010, respectively. He was a Postdoctoral Researcher in the BK21 Information Technology, Seoul National University, in 2010 and in the Institute for Robotics and Intelligent Systems of Computer Science Department, University of Southern California, Los Angeles, until August of 2011. He is currently a Professor with the Department of Computer Engineering, Dankook University, Korea. His research interests include pattern recognition, feature extraction and selection, machine learning, computer vision, and their applications.