

## 초기움직임에서의 동작인식<sup>\*</sup>

김건우<sup>01</sup> 임세민<sup>2</sup> 박주영<sup>3</sup> 안인경<sup>3</sup>, 오형철<sup>3</sup>

<sup>1</sup>고려대학교 대학원 전자정보공학과

<sup>2</sup>고려대학교 병렬연산연구실

<sup>3</sup>고려대학교 과학기술대학

{asdv1234,jaewoong819,parkj,ahnk,ohyeong}@korea.ac.kr

## Activity Recognition from Early-stage Motion

K.-W. Kim<sup>01</sup> S.-M. Lim<sup>2</sup> J. Park<sup>3</sup> I.-K. Ahn<sup>3</sup> and H.-C. Oh<sup>3</sup>

<sup>1</sup>Dept of Electron. & Info. Eng., Graduate School, Korea University

<sup>2</sup>Parallel Computation Lab., Korea University

<sup>3</sup>College of Science and Tech., Korea University

### 요 약

탁구 선수의 초기움직임으로부터 선수가 사용할 기술을 인식하였다. 2인의 선수의 머리와 상체에 부착된 8개의 가속도 센서로부터, 탁구의 기본동작(드라이브, 커트, 스트로크 등)의 정보를 수집한 후, 초기 (23~32%) 데이터 부분만을 사용하여 LSTM 순환신경망(단방향 및 양방향)을 학습시켰다. 추론성능을 비교한 결과, 신체의 여러 부분의 움직임을 이용함으로써 초기 (32%) 움직임 정보만으로도 90%이상의 추론성능을 얻을 수 있음을 관찰하였다.

### 1. 서 론

최근 인공지능 이론의 발전으로 스마트 보조기술 분야에서 사람의 움직임에 관한 정보를 분석하여 행동을 인식하는 기술의 적용이 활발하게 연구되고 있다[1,2,4,5,6,7]. 논문[2,4]에서는 한쪽 팔에 부착된 3개의 센서 모듈을 사용하여 움직임에 대한 정보를 얻어내고, 얻은 정보 데이터와 딥 러닝 기술을 활용하여 사용자가 취한 동작이 시스템에 저장되어 있는 기준이 되는 탁구 동작 중 어느 동작에 해당하는가를 추론(Inference)하여 사용자의 훈련에 도움을 주는 시스템을 제안한 바 있다. 또한, 논문[3]에서는 딥 러닝 기술에 비하여 비용 측면에서 장점을 보이는 저비용 가속기도 제안하였다. 이 밖에도 걸음 걸이 등의 정보에 대한 사진을 수집하거나 휴대폰에 있는 센서를 이용하여 정보를 수집하고 행동을 인식하는 다양한 연구 결과들이 발표되어 있다[5,6,7].

본 논문에서는 상체에 부착한 8개의 센서 모듈에서 수신되는 정보와 딥 러닝 기술 중 LSTM(Long Short-

Term Memory) 순환신경망(Recurrent Neural Network: RNN)을 사용하여 탁구선수가 사용하는 기술을 인식하는 연구를 수행한다. 특히, 모든 동작이 취해지기 전, 동작의 초기 동작에 관한 정보만으로 어느 동작인지 추론할 수 있는 시스템을 제안한다. LSTM 순환신경망의 설계에는 단방향(unidirectional)과 양방향(bidirectional) 순환신경망을 고려하였다.

### 2. 순환신경망

#### 2.1 단방향(Unidirectional) LSTM 순환신경망

단방향 LSTM 순환신경망은 가장 간단한 LSTM 순환신경망이다. 보통의 순환신경망은 음성 신호, 문자, 주가 등 배열(sequence)의 형태를 갖는 데이터의 패턴을 인식하는데 적합한 모델로 많이 쓰이고 있으나, 정보가 많아질수록 훈련된 정보가 학습에 미치는 영향이 점점 감소하다가 결국 사라져버리는 문제점을 가지고 있다. LSTM 순환신경망은, 이와 같은 기울기값 소멸(vanishing gradient)의 문제를 해결하기 위한 모델인데, 단방향

\* 본 논문은 한국연구재단 중견과제(과제번호: 2017R1E1A1A03070652)의 지원으로 수행되었습니다.

LSTM 순환신경망은 층(layer)들을 전방(forward)의 방향으로만 연결한다. 본 논문에서는, 논문 [4]와 같이, LSTM cell을 2개 층으로 쌓은 단방향 LSTM 순환신경망을 설계하였다. 그림 1은,  $t-1, t, t+1$  의 3개 time step에서의 단방향 LSTM 순환신경망의 동작을 보인 것인데, 각 cell은 시간  $t$ 에서 다음과 같은 연산을 수행한다[4].

$$f_t^{d,l} = \phi_f(W_{xf}^{d,l}x_t, W_{hf}^{d,l}h_{t-1}, b_f^{d,l}) \quad (1)$$

$$i_t^{d,l} = \phi_i(W_{xi}^{d,l}x_t, W_{hi}^{d,l}h_{t-1}, b_i^{d,l}) \quad (2)$$

$$o_t^{d,l} = \phi_o(W_{xo}^{d,l}x_t, W_{ho}^{d,l}h_{t-1}, b_o^{d,l}) \quad (3)$$

$$g_t^{d,l} = \phi_g(W_{xg}^{d,l}x_t, W_{hg}^{d,l}h_{t-1}, b_g^{d,l}) \quad (4)$$

$$c_t^{d,l} = f_t^{d,l} \otimes c_{t-1}^{d,l} + g_t^{d,l} \otimes i_t^{d,l} \quad (5)$$

$$h_t^{d,l} = o_t^{d,l} \otimes A(c_t^{d,l}) \quad (6)$$

식(1)-(6)에서,  $x_t$ 는 시간  $t$ 일 때 입력되는 데이터이고,  $d$ 는 방향,  $l$ 은 네트워크의 레벨이며,  $\phi_f, \phi_i, \phi_o, \phi_g$ 는 각각 망각 게이트(forget gate), 입력 게이트(input gate), 출력 게이트(output gate), 입력 변조 게이트(input modulation gate)라고 하는데 식(7)과 같이 정의된다[4].

$$\phi_k(W_{xk}, x, W_{hk}, h, b_k) = A(W_{xk}x + W_{hk}h + b_k) \quad (7)$$

식(7)에서  $A$ 는 활성화 함수(activation function)이다.

분류기(classifier)는 ‘many-to-one’ 기법을 사용하여 많은 양의 동작 데이터들이 입력되면, 미리 설정해준 레이블(label)들을 출력하여 서로를 분별할 수 있도록 구현하였는데, 사용되는 데이터를 고려하여 드롭아웃(dropout) 계층을 추가하여, 생길 수 있는 과적합(overfitting)의 문제를 해결하였다. 활성화 함수로는 ReLU(Rectified Linear Unit) 함수를 사용하였다.

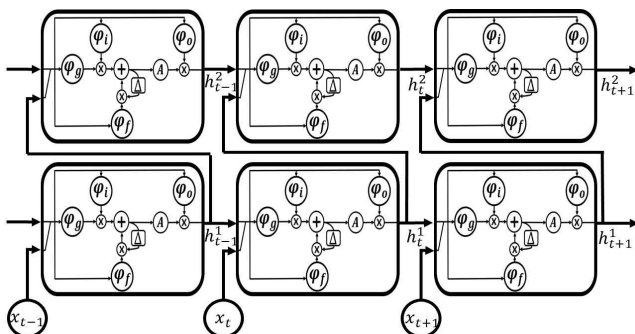


그림 1. 2-stacked 단방향 LSTM 순환신경망[4]

## 2.2 양방향(Bidirectional) LSTM 순환신경망

양방향 LSTM 순환신경망은 숨은 층(hidden layer)에 전방향 뿐만 아니라, 후방향(backward) 접속을 추가하여 양방향의 정보에 의존한 출력 값을 도출할 수 있다. 단방향 LSTM 순환신경망과 비교하였을 때, 양방향 LSTM 순환신경망은 과거의 상태뿐만 아니라, 미래의 상태까지 고려하므로, 시계열 데이터를 사용하는 본 실험에서 좋은 결과가 예상되는 모델이다. 논문 [4]에서와 같이, 본 논문에서는 단방향 LSTM 순환신경망 모델과 동일하게 2개의 LSTM cell을 2개 층으로 사용한 양방향 LSTM 순환신경망을 구현하였는데 (그림 2), 단방향 LSTM 순환신경망과 달리 후방향이 추가되었기 때문에, 각 LSTM cell의 출력 부분의 가중치를 2배로 늘려서 설계하였다. 가중치 이외의 조건들은 정확한 비교를 위해 동일하게 구성하였다.

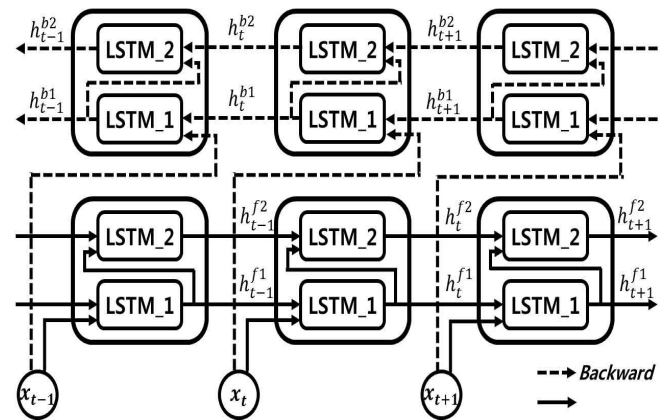


그림 2. 2-stacked 양방향 LSTM 순환신경망[4]

## 3. 실험

### 3.1 Data 수집

본 논문에서는, 그림 3과 같이, 총 8개의 IMU(Inertial Measurement Unit) 센서를 머리에 2개, 팔과 가슴에 각각 4개, 2개를 부착하고 라즈베리파이 보드에 연결하여 데이터를 수집하는 장치(프로토타입)를 사용하였다. 센서 모듈 8개에서 수신되는 각각의 x, y, z방향의 가속도 데이터 (24 features = 3 axes x 8 sensors)를 수집하고, 한 동작을 표현하는 22개의 시간점(time series)으로 구성하였다. 수집한 데이터의 70%는 훈련(train) 용으로, 30%는 테스트(test)용으로 구성하였다.

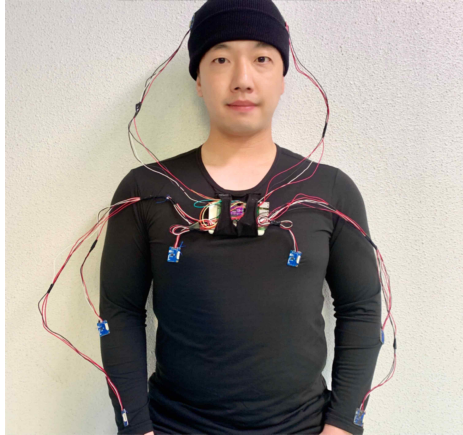


그림 3. Data 수집 장치 (Prototype)

### 3.2 실험 결과

본 논문의 실험에는, Intel i5-8250U CPU 와 NVIDIA GeForce GTX 1050 상에서 구글 오픈소스 라이브러리인 Tensorflow가 사용되었다[8].

표 1에서는 본 논문의 센서모듈 8개를 사용하여 추론한 결과를, 논문[2]에서와 같이 센서모듈 3개를 사용한 결과와 비교하였다. 또한, 동작의 22개 시간점 데이터 모두를 사용하여 학습하고 추론한 결과와, 초기의 5개(23%) 또는 7개(32%)의 시간점 데이터만을 사용하여 학습시키고 추론한 결과를, 추론의 정확도 (Accuracy) 면에서 비교하였다.

표 1에 보인 바와 같이, 모든 데이터를 사용하였을 때 양방향 LSTM 순환신경망은, 센서모듈 3개를 사용하였을 때와 8개를 사용하였을 때 모두 높은 정확도(93.3% 이상)를 보였다. 또한 센서모듈 8개를 사용하면, 동작의 초기 32%에 해당하는 데이터만 사용하여도, 단방향 및 양방향 LSTM 순환신경망 모두 90% 이상의 높은 추론 정확도를 얻을 수 있었다.

순환신경망	데이터사용	센서모듈 3개	센서모듈 8개
단방향	초기 23%	80.0%	83.3%
	초기 32%	83.3%	90.0%
	모두 사용	86.7%	93.3%
양방향	초기 23%	83.3%	86.7%
	초기 32%	86.7%	93.3%
	모두 사용	93.3%	96.7%

표 1. 센서 개수에 따른 정확도 비교

### 4. 결 론

본 논문에서는 센서 모듈에서 수신되는 탁구의 기본 동작에 관한 정보에서 초기 동작의 정보만으로 선수가 사용하는 기술을 추론하는 시스템을 제안하였다. 제안된 시스템은 단방향 및 양방향 LSTM 순환신경망을 사용하여 설계하였으며, 신경망의 구조와 센서 모듈의 개수에 따른 신경망의 추론성능을 실험을 통하여 비교하였다. 본 논문에서 제안하는 시스템은 단순히 운동 보조시스템만이 아닌, 장애인의 재활 훈련 등 다양한 의료 진단 및 훈련에 도움이 되는 시스템으로 발전해 나아갈 것으로 기대된다.

### 참고 문헌

- [1] M. Kranz, et al., The Mobile fitness coach: Towards individualized skill assessment using personalized mobile devices. *Pervasive and Mobile Computing*. 9(2), pp.203-215, 2013.
- [2] 임세민 외, “딥 러닝을 이용한 탁구연습 보조시스템”, 한국정보과학회 KSC 2017 논문집, pp.960-962, 2017년 12월.
- [3] 임세민 외, “저비용 탁구연습 보조시스템”, 한국정보과학회 KSC 2018 논문집, pp.1042-1044, 2018년 6월.
- [4] Lim. S.-M., et al., “LSTM-Guided Coaching Assistant for Table Tennis Practice”, *Sensors* 2018, 18(12), 4112.
- [5] A. Ullah, et al., “Action Recognition in Video Sequences using Deep Bi-Directional LSTM With CNN Features”, *IEEE Access.*, vol. 6, pp1155-1166, Nov. 2017.
- [6] Hao Xue, Du Q. Huynh, Mark Reynolds, “Bi-Prediction: Pedestrian Trajectory Prediction Based on Bidirectional LSTM Classification”, *International Conference on DICTA 2017*, Dec. 2017.
- [7] M. Milenkoski, et al., “Real Time Human Activity Recognition on Smartphones using LSTM Networks”, *MIPRO 2018*, pp1126-1131, May. 2018.
- [8] Abadi, M.; Barham, P.; Chen, J.; Chen, Z.; Davis, A.; Dean, J.; Devin, M.; Ghemawat, S.; Irving, G.; Isard, M.; et al. Tensorflow: A System for Large-scale Machine Learning. In *Proceedings of the 12<sup>th</sup> USENIX Conference on Operating Systems Design and Implementation*, Savannah, GA, USA, 2-4 November 2016; pp.265-283.