

GRU 기반 초기 동작 인식*

김건우⁰¹ 장규삼² 임세민¹ 안인경² 박주영² 오형철²

¹고려대학교 병렬연산연구실

²고려대학교 과학기술대학

e-mail : {asdv1234,gsjang95,jaewoong819,ahnk,parkj,ohyeong}@korea.ac.kr

GRU-based Activity Recognition from Early-stage Motion

K.-W. Kim⁰¹ G.-S. Jang² S.-M. Lim¹ I.-K. Ahn² J. Park² and H.-C. Oh²

¹Parallel Computation Lab., Korea University

²College of Science and Tech., Korea University

Abstract

This paper proposes a system that recognizes a person's activity from the data, collected during a initial short interval of time, for the motion which he/she takes. Eight designs of RNNs (recurrent neural networks) are considered to identify which skills (stroke, drive, short, etc) that a table-tennis player takes. We show that the designs based on GRU (Gated Recurrent Units) require significantly less parameters but perform better than other designs.

I. 서론

최근 인공지능 분야에서, 사람의 움직임에 관한 정보를 분석하여 그의 행동을 인식하는 기술들이 활발하게 연구되고 있다[1-9]. 특히, 영상 데이터에서 직접 인식하거나, 휴대성을 추가한 스마트 보조기술에 적용하여, 일상생활에서의 사고(事故)나 동작, 또는 이동 경로를 인식하는 응용 분야가 주목을 받고 있다.

일상적으로 휴대하는 스마트폰에 내장된 센서를 이용하면 정보를 수집하기가 용이하지만, 재빨이나 스포

츠 훈련을 목표로 한, 세밀한 동작의 인식을 위하여 여러 개의 센서와 교신하는 별도의 휴대용 기기를 사용이 허용된 응용들도 관심을 받고 있다. [2]에서는 LSTM(Long Short-Term Memory) 순환신경망 (Recurrent Neural Network; RNN)을 사용하여, 운동자의 팔의 움직임으로부터 운동자의 동작이 시스템에 저장되어있는 동작 데이터 중에서 어느 동작에 해당하는가를 추론(Inference)하여 운동자의 훈련에 도움을 주는 운동 보조시스템을 제안하였으며, [3]에서는 보다 저렴한 하드웨어 구현이 가능한 추론 방안을 제안하였다. [5]에서는 상반신에 부착된 8개의 센서에서 수집되는 데이터 중 짧은 순간의 동작 정보만으로, LSTM 순환신경망을 이용하여, 동작을 추론하는 실험 결과를 발표하였다.

본 논문에서는 총 16개의 센서 모듈들을 상, 하반신에 각각 8개씩 부착하여 수집한 동작 데이터 중, 초기 동작에 관한 정보만을 이용하여 운동자가 행하는 탁구 기본 동작(드라이브, 커트, 스트로크 등)을 인식하고자 하였다. GRU(Gated Recurrent Unit)[10]를 사용하였으며, 2층(2-stacked) 혹은 3층의 구조 및 단방향(Uni-directional)과 양방향(Bidirectional) 구조 등을 고려하면서, RNN 및 LSTM 순환신경망을 사용한 설계와의 추론성능을 비교하였다.

* 본 연구는 한국연구재단의 중견과제 (과제 번호:2017R1E1A1A03070652)에서 지원하여 연구하였음.

II. 본론

2.1 LSTM 순환신경망

일반적인 순환신경망은, 음성이나 문자 등 순서를 갖는 데이터의 패턴을 인식하는데 적합한 모델로 많이 쓰이지만, 학습시키는 정보가 많아질수록 학습에 미치는 영향이 점점 감소하다가 결국 사라져버리는 기울기 값 소멸(vanishing gradient) 문제를 가지고 있다. LSTM 순환신경망은, 이 문제를 해결할 수 있는 신경망으로서, 단방향과 양방향 구조가 있다. 단방향 LSTM 순환신경망은 층(layer)들을 전방(forward)의 방향으로만 연결하는데, 양방향 LSTM 순환신경망은 보다 좋은 추론성능을 위하여, 후방향(backward) 연결도 추가하여 과거와 미래의 상태 모두 고려한다. 본 논문에서는, LSTM cell을 2개 혹은 3개의 층으로 쌓은 단방향 및 양방향 LSTM 순환신경망을 설계하였다. 그림 1과 그림 2는, $t-1, t, t+1$ 의 3개 time step에서의 단방향, 양방향 LSTM 순환신경망의 2층으로 설계하였을 때의 모습을 보인 것이다[4].

분류기(classifier)와 활성화 함수(activation function)는, [5]와 동일하게 ‘many-to-one’ 기법과 ReLU(Rectified Linear Unit)를 사용하였다.

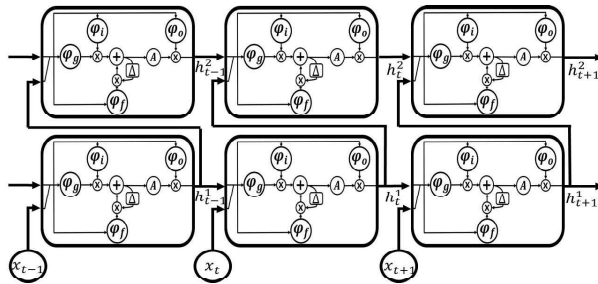


그림 1. 2-stacked 단방향 LSTM 순환신경망[4]

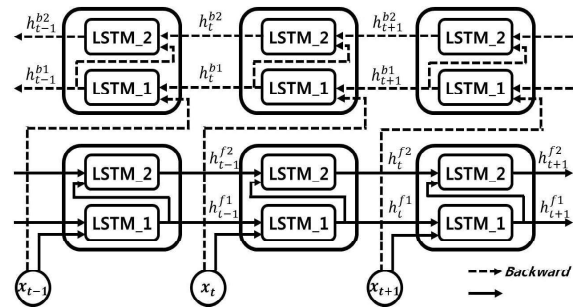


그림 2. 2-stacked 양방향 LSTM 순환신경망[4]

2.2 GRU

GRU[10]는 LSTM 순환신경망의 장점을 유지하면서,

연산량을 줄여 제안된 망이다. LSTM 순환신경망과 구조상 큰 차이가 없어 성능을 비교하여 보았을 때 차이점이 없다는 평가도 있지만, 사용하는 데이터 혹은 대상 응용에 따라 성능의 차이를 얻을 수 있다. GRU는 셀 상태를 저장하여 기억해주는 메모리가 없어 데이터양이 많아지면 LSTM 순환신경망보다 좋지 않은 결과를 보일 수 있지만, 본 논문에서와같이 적은 양의 시계열 데이터를 사용할 때에는 LSTM 순환신경망보다 수행하는 연산량이 적으므로 연산 시간과 추론 정확도 측면에서 좋은 결과를 기대할 수 있다. 그림 3은 2층으로 설계한 GRU 구조를 보인 것이다.

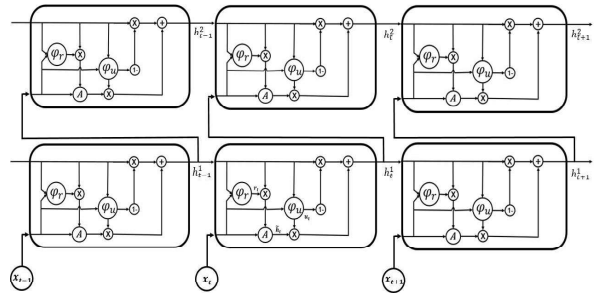


그림 3. 2-stacked GRU[9]

III. 구현

3.1 Data 수집

본 논문에서는, 그림 4와 같이, 16개의 IMU(Inertial Measurement Unit) 센서를 전신에 부착하고 라즈베리 파이와 연결하여 데이터를 수집하는 장치를 제작하여 사용하였다. 2인의 선수가 취하는 탁구의 기본 동작(드라이브, 커트, 스트로크 등)에 대해, 센서 모듈 16개에서 수신되는 각각의 x, y, z방향의 가속도 데이터(48 features = 3 axes x 16 sensors)를 수집하고, 한 동작을 표현하는 시간점(time series) 데이터 22개 중



그림 4. 데이터 수집 장치

초기 동작에 해당하는 3개(14%)만으로 구성하였다. 수집한 데이터의 70%는 훈련(train)-용으로, 30%는 테스트(test)-용으로 사용하였다.

3.2 실험 결과

본 논문의 실험은, Intel i5-8250U CPU와 NVIDIA GeForce GTX 1050 GPU 및 Ubuntu 18.04 환경에서 구글 오픈소스 라이브러리인 Tensorflow[11]를 사용하였다.

표 1은 신경망 별로 추론 정확도와 학습 소요시간을 보인 것이다. 표 1에 보인 바와 같이, 기본적인 RNN의 경우, 다른 신경망들에 비해 비교적 짧은 시간이 걸렸지만, 정확도 측면에서 다른 신경망들에 비하여 부족한 성능을 보였으며, 단방향 및 양방향 LSTM 순환신경망과 GRU는 90% 이상의 추론 정확도를 보였다. GRU는 기본적인 RNN보다는 많은 계산량을 요구하지만, 두 가지의 LSTM 순환신경망보다는 적은 양의 계산량을 요구하면서도 가장 좋은 추론성능을 보였다. 특히, 3-Stacked GRU의 경우 100%의 정확도를 보였으며, 단방향 및 양방향 LSTM 순환신경망과 비교하였을 때, 소요시간 측면에서도 우수함을 확인하였다.

층(Stacked)	신경망	초기 14% 데이터	
		추론 정확도	소요시간
2-Stacked	RNN	83.3%	6.43s
	단방향 LSTM	90.0%	10.41s
	양방향 LSTM	93.3%	13.50s
	GRU	96.7%	9.94s
3-Stacked	RNN	90.0%	7.70s
	단방향 LSTM	93.3%	13.42s
	양방향 LSTM	96.7%	15.27s
	GRU	100%	12.14s

표 1. 신경망 별 추론 정확도와 학습 소요시간

표 2는 각 신경망에서 사용된 매개변수(parameter)의 개수를 비교한 것이다. 기본적인 RNN은 매개변수의 양이 상대적으로 적지만, 다른 신경망들에 비해 정확도 측면에서 우수한 성능을 보여주지 못한다. 매개변수의 양과 정확도 면을 같이 고려하여 비교하였을 때, 비교적 적은 데이터를 활용하면서도 좋은 성능을 보인, GRU가 본 연구에서 목표로 하는 시스템에 좋은 선택이 될 수 있음을 보였다.

층(Stacked)	신경망	매개변수[개]
2-Stacked	RNN	26,752
	단방향 LSTM	107,008
	양방향 LSTM	214,016
	GRU	80,256
3-Stacked	RNN	40,128
	단방향 LSTM	160,512
	양방향 LSTM	321,024
	GRU	120,384

표 2. 매개변수의 양

IV. 결론 및 향후 연구 방향

본 논문에서는, 센서 모듈 16개를 사용하여 운동자의 전신에서 수집한, 탁구의 기본 동작의 초기 데이터만으로 운동자가 취하는 기술을 추론하는 시스템을 제안하였다. 제안된 시스템은 GRU를 사용하여 설계하였고, 기본적인 RNN, 단방향 및 양방향 LSTM 순환신경망들과 성능을 비교하였다. 같은 양의 초기 데이터를 사용할 때, 상체에 관한 정보만으로는 성능이 좋지 못하였으나, 전신에 대한 정보를 활용하면서 더 적은 데이터를 사용하여도 좋은 성능을 얻을 수 있었다. 제안하는 시스템은 센서에서 수신된 데이터만을 활용하여 운동자가 취한 동작이 어느 동작인가를 추론하였으나, 더 나아가 영상 혹은 이미지 등 다른 정보도 이용하고 운동자가 취하는 동작에 관한 다른 형태의 정보를 제공할 수 있으면 다양한 분야에서 활용될 수 있을 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 논문집, pp.960-962, 2017년.
- [3] 임세민 외, “저비용 탁구연습 보조시스템”, 한국정보과학회 KCC 논문집, pp.1042-1044, 2018년.
- [4] S.-M. Lim et al., “LSTM-Guided Coaching Assistant for Table Tennis Practice”, Sensors 2018, 18(12), 4112.
- [5] 김건우 외, “초기움직임에서의 동작인식”, 한국정보

과학회 KCC 논문집, pp.1035-1037, 2019년.

[6] A. Ullah, et al., "Action Recognition in Video Sequences using Deep Bi-Directional LSTM With CNN Features", IEEE Access, vol.6, pp.1155-1166, Nov. 2017.

[7] H. Xue et al., "Bi-Prediction: Pedestrian Trajectory Prediction Based on Bidirectional LSTM Classification", Int. Conf. on DICTA, 2017.

[8] M. Milenkoski et al., "Real Time Human Activity Recognition on Smartphones using LSTM Networks", 2018 MIPRO, pp.1126-1131, 2018.

[9] 임세민 외, "영상 데이터에서의 CRNN 기반 자세 인식", 한국정보과학회 KCC 논문집(예정), 2020년.

[10] K. Cho et al., "On the properties of neural machine translation: Encoder-decoder approaches", **arXiv preprint arXiv:1409.1259**, 2014.

[11] M. Abadi et al., Tensorflow: A System for Large-scale Machine Learning. In Proc. USENIX Conference on Operating Systems Design and Implementation, pp.265-283, 2016.

[12] R. Dey et al., "Gate-variants of Gated Recurrent Unit(GRU) neural networks", IEEE MWSCAS, Aug, 2017.