

LSTM 기법을 적용한 UTD 데이터 행동 분류

정겨운 · 안지민 · 신동인 · 원건 · 박종범

전자부품연구원

Classification of Behavior of UTD Data using LSTM Technique

Jeung Gyeo-wun · Ahn Ji-min · Shin Dong-in · Won Geon · Park Jong-bum

Korea Electronics Technology Institute

E-mail : wjdrudns92@keti.re.kr, jbpark@keti.re.kr

요 약

본 연구는 인공신경망의 한 종류인 LSTM(Long Short-Term Memory) 기법을 활용하기 위하여 진행하였다. UTD(University of Texas at Dallas)가 공개한 27종 동작 데이터 중 3축 가속도 및 각속도 데이터를 기본 LSTM 및 Deep Residual Bidir-LSTM 기법에 적용하여 행동을 분류해 보았다.

ABSTRACT

This study was carried out to utilize LSTM(Long Short-Term Memory) technique which is one kind of artificial neural network. Among the 27 types of motion data released by the UTD(University of Texas at Dallas), 3-axis acceleration and angular velocity data were applied to the basic LSTM and Deep Residual Bidir-LSTM techniques to classify the behavior.

키워드

LSTM(Long Short-Term Memory) Technique, Multimodal Human Action Dataset, acceleration, angular velocity, classification

I. 서 론

최근 영상과 신호를 사용하여 인식하고 분석하여 특정 상황이나 행동을 감지 및 분류하는 작업이 많이 진행되고 있다. 이러한 과정에서 사용되는 기계 학습(Machine Learning)은 스스로 학습시켜 자동으로 분류할 수 있게 한다. 기계 학습 방법론 중 하나인 RNN(Recurrent Neural Network)은 데이터베이스를 복잡한 구조로 학습시켜 음성인식, 언어인식 등에서 높은 성능을 보여준 사례들이 있다. 따라서 본 연구에서는 RNN을 변형시킨 LSTM 기법에 UTD가 공개한 가속도 및 각속도 데이터를 적용하여 27종 행동을 분류하고자 하였다.

II. LSTM 기법

RNN은 순차적인 데이터를 처리하는데 적합하며, 입력데이터가 주어지면 입력된 시간에 따라 처리하여 내부노드를 생성한 후 정보를 저장한다. 내

부노드에서 순환 형식으로 이전에 저장된 정보를 추가하여 상태값을 생성하고, 또 다른 가중치를 곱한 후 활성화함수를 적용해 출력값을 도출한다. 예상 결과와 도출한 출력값을 비교하여 오류를 계산하여 오류역전파 알고리즘을 적용해 가중치를 학습한다. 이 과정에서 가중치에 편도함수를 곱하는 과정이 필요한데, 연산의 수가 많아지면 학습속도가 매우 느려지거나 멈추게 될 수 있다. 이러한 문제를 해결하기 위해 RNN의 내부노드를 메모리 셀 구조로 바꾼 기법이 LSTM 기법이다.

LSTM[1]의 메모리 셀은 RNN의 상태값에 세 개의 게이트가 추가된 구조이다. 첫 번째 게이트는 어떤 정보를 반영할지 결정하는 Forget 게이트로 활성화함수를 적용한 값이 1에 가까울수록 정보를 많이 반영하고, 0에 가까울수록 적게 반영한다. 두 번째 게이트는 새로운 정보를 저장할지 결정하는 Input 게이트이며, Forget 게이트와 Input 게이트에서 출력된 값으로 셀 상태값을 계산한다. 마지막으로 최종 출력할 값을 결정하는 Output 게이트에서 셀 상태값을 필터링하여 활성화함수를 적용한다. 이

와 같이 세 개의 게이트가 추가된 LSTM 메모리 셀을 그림 1에 나타내었다.

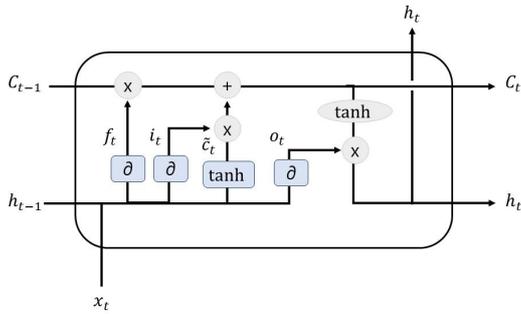


그림 1. LSTM 메모리 셀

본 연구에서는 기본적인 LSTM 기법과 기존의 연구에서 개발된 Deep Residual Bidir-LSTM 기법 [2]을 사용하여 데이터를 분류하였다. Deep Residual Bidir-LSTM 기법은 기본 LSTM 기법을 변형하여 셀을 양방향으로 학습하는 Bidirectional LSTM과 스택 형식의 Residual LSTM 을 합친 것이며, 이를 그림 2에 나타내었다.

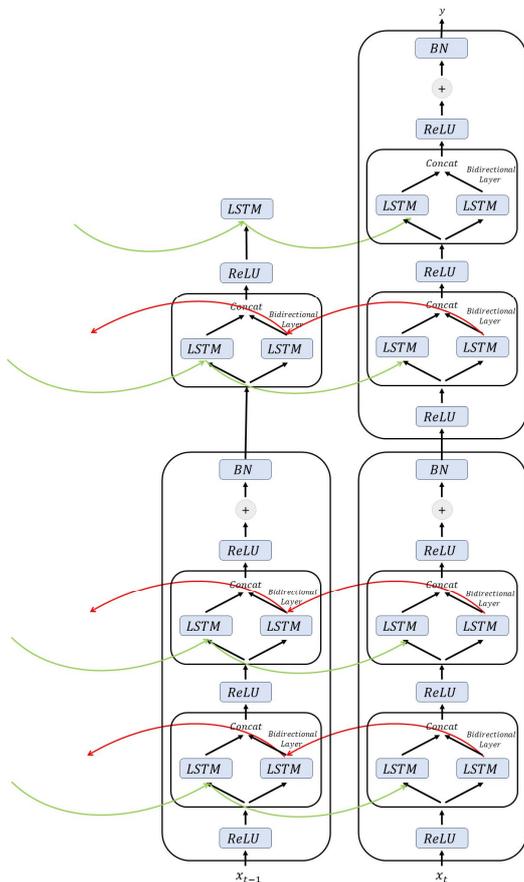


그림 2. Deep Residual Bidir-LSTM 구조

III. UTD 데이터베이스 및 입력값 설정

본 연구에서 사용한 데이터베이스는 UTD (University of Texas at Dalls)[3]가 공개했으며, 8명의 피험자에게 한 행동마다 약 4번씩 총 27종 행동을 취해 얻은 데이터이다(그림 3).



그림 3. UTD 27종 행동

데이터베이스에는 3축 가속도 및 각속도, 영상, 스켈레톤 데이터가 있으며 그 중 가속도 및 각속도 데이터를 사용하였다. 가속도 및 각속도 데이터는 팔 동작 위주인 21종 행동은 오른쪽 손목, 다리 동작 위주인 6종 행동은 오른쪽 허벅지에 센서를 착용하고 측정하였다. 각각의 데이터의 특징을 두드러지게 하기 위해 제공한 후, 리샘플링하였다. -1~1 사이의 값으로 정규화하여 기본 LSTM 및 Deep Residual Bidir-LSTM의 입력값으로 가속도 x, y, z, 각속도 x, y, z의 6가지 데이터를 설정하였다. 한 행동마다 약 32개의 데이터셋이 존재하고, 이 중에서 학습시킬 데이터셋의 개수를 10, 15개로 정하여 학습시켰다. 또한, Deep Residual Bidir-LSTM 구조의 LSTM 스택과 은닉층 스택의 개수를 조절하여 최고 정확도를 도출하였다.

IV. 행동 분류 결과

전체 861개의 데이터셋에서 각각 랜덤으로 320, 480개를 학습시킬 데이터셋으로 선정하여 학습하고, 이 외의 데이터로 테스트하였다.

표 1. 학습한 데이터셋 개수에 따른 정확도

데이터셋 개수	320개/	480개/
최고 정확도	541개	381개
LSTM	51.16%	57.29%
Deep Residual Bidir-LSTM	70.93%	78.17%

학습 데이터셋의 개수를 480개로 설정하였을 때 정확도가 더 높았다. Deep Residual Bidir-LSTM의 경우 LSTM 스택은 1개, 은닉층 스택은 5개로 설정하였을 때 더 높은 정확도를 보였다. 또한, 기본 LSTM과 Deep Residual Bidir-LSTM의 정확도를 비교하였을 때 후자가 더 높았으며, 결과 도출 시간도 현저히 감소했다.

V. 결 론

본 연구에서는 UTD가 공개한 27종 행동에 대한 3축 가속도 및 각속도 데이터를 기본 LSTM 및 Deep Residual Bidir-LSTM 기법을 사용하여 분류해 보았다. 두 기법 모두 학습하는 데이터셋이 많을수록 정확도가 증가했으며, 기본 LSTM 기법보다 Deep Residual Bidir-LSTM 기법을 사용했을 때 더 높은 정확도를 보였다. 추후에는 팔에 착용한 가속도 및 각속도 데이터를 Deep Residual Bidir-LSTM 기법에 적용하여 행동 분류하는 연구를 진행할 예정이다.

Acknowledgement

본 연구는 미래창조과학부 및 정보통신기술진흥센터의 정보통신·방송 연구 개발 사업의 일환으로 수행하였음. [2017-0-00162, 고령 사회에 대응하기 위한 실환경 휴먼케어 로봇 기술 개발]

References

- [1] Hochreiter, S. and J. Schmidhuber, "Long short-term memory," *Neural Computation*, Vol. 9, pp. 1735-1780, 1997.
- [2] Z. Yu, R. Yang, G. Chevalier, M. Gong, "Deep Residual Bidir-LSTM for Human Activity Recognition Using Wearable Sensors," arXiv: 1708.08989, 2017.
- [3] C. Chen, R. Jafari, and N. Kehtarnavazs, "UTD-MHAD: A Multimodal Dataset for Human Action Recognition Utilizing a Depth Camera and a Wearable Inertial Sensor," Proceedings of IEEE International Conference on Image Processing, Canada, September 2015.