A CNN-LSTM MODEL FOR IMU-BASED ENERGY EXPENDITURE ESTIMATION UNDER VARIOUS WALKING CONDITIONS

Chang June Lee[1] and Jung Keun Lee[2]
[1]Department of Integrated Systems Engineering, Hankyong National University, Anseong, South Korea
[2]School of ICT, Robotics & Mechanical Engineering, Hankyong National University, Anseong, South Korea

2024 IEEE International Symposium on Inertial Sensors and Systems (INERTIAL) | 979-8-3503-1669-8/24/$31.00 ©2024 IEEE | DOI: 10.1109/INERTIAL60399.2024.10502118

Energy expenditure (EE) is important quantities in evaluating the intensity of physical activity and health status. EE can be estimated based on signals from wearable sensors such as inertial measurement unit (IMU) and electromyography (EMG) sensor. Although many algorithms have been developed to estimate EE during daily life [1-4], only a few algorithms have evaluated the performance of EE according to walking conditions (speed and incline) [2-4]. This paper presents a combined model of convolutional neural network (CNN) and long short-term memory (LSTM) for an IMU-based EE estimation under various walking conditions. The results of this study suggest the optimal IMU location for estimating EE during level walkings/runnings and inclined walkings.

In this study, we developed a CNN-LSTM model for estimating steady-state EE during different walking from IMU signals. The model consists of an input layer feeding six-axis IMU signal (accelerometer and gyroscope) sequences of length 200 (two seconds data at 100Hz), a one-dimensional convolutional layer (with a kernel size of 3 and 64 filters), a maximum pooling layer (with a kernel size of 3 and 64 filter), two LSTM layers (with 64 hidden units), a flatten layer, and a fully connected layer that receives the output of the flatten layer concatenated with the body weight and height. The structure of the model is shown in Fig. 1. We conducted level walking and running experiments at four speeds (4 and 6KPH for walking and 7.5 and 9KPH for running) and inclined walking experiments at two speeds (4 and 6KPH) and two inclines (3 and 6%) on a treadmill. Six healthy adult males participated in the experiments (age: 24.7 ± 1.7years, height: 1.75 ± 0.05m, mass: 72.9 ± 6.9kg). Five IMUs (MTw, Xsens, Enschede, Netherlands) were placed on the chest, right wrist, right thigh, right shank, and right foot. A metabolic analysis system (K5, Cosmed, Rome, Italy) was used to obtain the truth reference of EE (in Watt) based on the volume of oxygen and carbon dioxide. The experimental setup is shown in Fig. 2. To compare the performance by sensor location, a total of five models were trained for each sensor location. Leave-one-subject-out cross-validation was performed to evaluate the performance of the model.

Table 1 shows the averaged normalized root mean squared error (NRMSE) related to body weight and mean absolute percent error (MAPE) for each of the three test sets. The first set includes level and inclined walking tests, the second set includes level walking and running tests, and the third set includes total tests. In Table 1, EE performance was shown to be excellent in the order of sensor locations: shank, foot, thigh, chest, and wrist. In particular, the RMSE and MAPE from most other models were higher than 0.9W/kg and 11%, respectively, while the shank attachment model showed RMSE and MAPE of less than 0.8W/kg and 11%, respectively, for the three test sets. This indicates that the shank is the most optimal sensor location for estimating EE during level/inclined walking and level running. Fig. 3 is an example result of EE estimation for the subject who produced the highest accuracy from the model based on the shank-attached IMU. The estimation results showed that the model accurately estimated EE for each walking speed and incline. In literature, recent studies [3,4] also developed data-driven models and evaluated their performances for various walking conditions. Particularly, [3] used EMG and vertical ground reaction force (GRF) data as model input and reported an RMSE of 0.83W/kg and a MAPE of 9.7% for inclined and loaded walking conditions. Also, [4] used the thigh and shank-attached IMU data and reported a MAPE of 13% for walking/running at various speeds. These studies used data segmented by each stride as model input, which requires detection of heel strike based on GRF or IMU signals. However, it is difficult to accurately detect heel strikes during walking at various inclines using IMU signals. This in deed affects the EE estimation performance. The results of this study show that the EE for different walking conditions can be estimated simply and reliably by using sensor signal sequences of a fixed-size from a single IMU without segmentation by stride.

References

[1] S. Paraschiakos, C. R. de Sá, J. Okai, P. E. Slagboom, M. Beekman, and A. Knobbe, *Data Min. Knowl. Discov.*, 36 (2022), pp. 477-512.

[2] J. M. Lopes, J. Figueiredo, P. Fonseca, J. J. Cerqueira, J. P. Vilas-Boas, and C. P. Santos, *Sensors*, 22 (2022), 7913.

[3] P. Slade, R. Troutman, M. J. Kochenderfer, S. H. Collins, and S. L. Delp, *J. Neuroeng. Rehabilitation.*, 16 (2019), pp. 1-10.

[4] P. Slade, M. J. Kochenderfer, S. L. Delp, and S. H. Collins, *Nat. Commun.*, 12(2021), 4312.

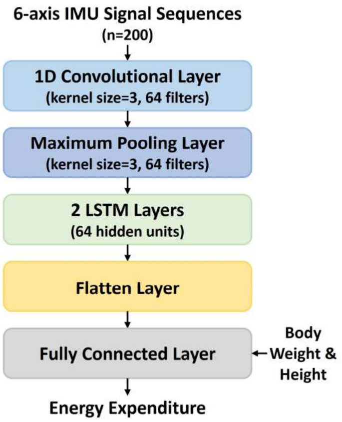

Figure 1. Structure of the model

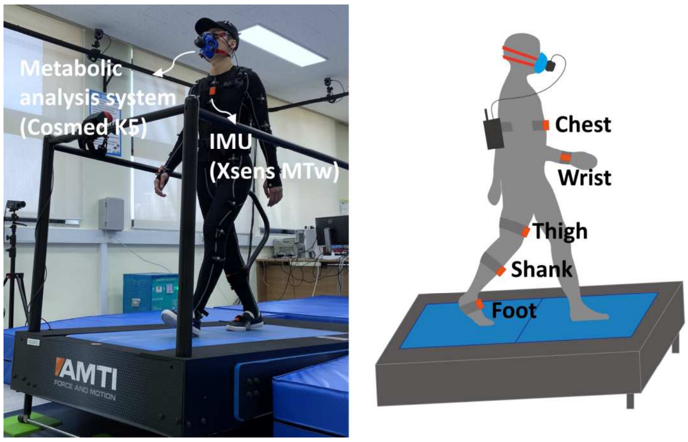

Figure 2. Experimental setup

Table 1. Averaged normalized root mean squared error in W/kg (with mean absolute percent error in %) of energy expenditure for each test sets.

Test sets	Chest	Wrist	Thigh	Shank	Foot
Level and inclined walkings	1.09 (14.03)	1.36 (17.03)	1.08 (15.63)	0.77 (11.00)	0.91 (13.28)
Level walking and running	0.96 (11.49)	1.17 (14.04)	0.92 (12.65)	0.76 (9.66)	0.87 (11.20)
Total tests	1.04 (12.30)	1.28 (14.75)	0.99 (13.14)	0.76 (9.66)	0.90 (11.65)

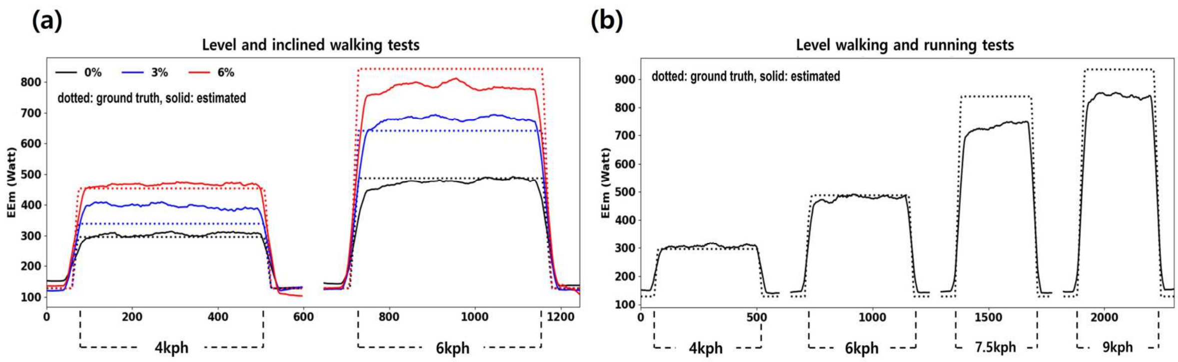

Figure 3. Estimation results of (a) level and inclined walking tests and (b) level walking and running tests from the shank-attached IMU-based model.