Using Deep Learning for Energy Expenditure Estimation with Wearable Sensors

Jindan Zhu,[1] Amit Pande,[1] Prasant Mohapatra,[1] and Jay J. Han[2]

[1]Department of Computer Science,
University of California at Davis, Davis, CA 95616
{jdzhu, pande, pmohapatra}@ucdavis.edu

[2]Dept. Physical Medicine and Rehabilitation,
UC Davis School of Medicine,Sacramento, CA, USA
jay.han@ucdmc.ucdavis.edu

***Abstract*—Energy Expenditure (EE) Estimation is an important step in tracking personal activity and preventing chronic diseases such as obesity, diabetes and cardiovascular diseases. Accurate and online EE estimation using small wearable sensors is a difficult task, primarily because most existing schemes work offline or using heuristics. In this work, we focus on accurate EE estimation for tracking ambulatory activities (walking, standing, climbing upstairs or downstairs) of individuals wearing mobile sensors. We use Convolution Neural Networks (CNNs) to automatically detect important features from data collected from triaxial accelerometer and heart rate sensors. Using CNNs, we find a significant improvement in EE estimation compared to other state-of-the-art models. We compare our results against state-of-the-art Activity-Specific Linear Regression as well as Artificial Neural Networks (ANN) based models. Using a universal CNN model, we obtain an overall low Root Mean Square Error (RMSE) of 1.12 which is 30% and 35% lower than existing models. The results were calibrated against a COSMED K4b2 indirect calorimeter readings.**

I. Introduction

Overweight and obese population is becoming a serious health threat around the world and especially in developed countries, as surveys show that approximately two-third of adults in United States are overweight [1]. Alongside with healthy diet habits that regulate energy intake, increased physical activity levels are warranted in urban population. Accurate quantification and monitoring of energy expenditure (EE) in individuals using mobile wearable devices is useful to provide useful feedback to healthcare providers (physician) as well as the user himself, providing an awareness and motivation to improved physical activity.

EE can be most accurately estimated using a direct or indirect calorimeter, which measures the pulmonary gas exchange in terms of O_2 consumption and CO_2 production in the individual. However, the lack of portability and bulkiness of such apparatus makes it impractical to operate in daily lifestyles and the usage is limited to clinical studies. Increasingly ubiquitous and mature wearable devices provide a reasonable alternative. Wearable devices, such as smartwatch and smartphones, are usually equipped with accelerometer and heart rate sensor to continuously monitor physical activity of an individual, which can be used to infer the energy expenditure.

Numerous effort has been invested to model the relationship between sensor data and energy expenditure. Most of them follow a similar methodology, extended from traditional activity recognition studies. A number of hand-crafted features that summarize the raw sensor data are extracted from the noisy data. They will then go through an automatic or manual feature selection phase which excludes features with low correlation to the objective, in order to reduce interference as well as redundancy. Given the list of features, statistical regression tools are used to create various models, universal or activity-specific, that translate them into energy expenditure predictions. Linear regression based models [2], as well as conventional machine learning based models such as artificial neural networks (ANN) [3] and ensemble decision trees [4] have been used with varying accuracy.

As suggested in previous work, feature selection has a decisive impact on the performance of regression task that estimates energy expenditure, therefore needs to be conducted carefully. Since feature selection phase holds such great importance, it is natural to ask: *Are the hand-crafted features good enough to capture subtle patterns relevant to EE prediction, and can better features be discovered to improve the accuracy of EE prediction?*

One possible answer is *Deep Learning*, an emerging machine learning technique that has obtained substantial attention due to its superior performance achieved in many application domains, from image/speech recognition [5] to activity classification [6]. In general, a deep learning framework is built on multiple-layer neural networks that consists of connected nodes at each layer. Each node takes input from prior layer, transforms it through an parametric mapping function, and exposes to the next layer. One particular interesting property of deep learning techniques is that it can work on raw data and act as an automatic feature extractor. Noisy sample series is processed as input vector, and during each transformation a hidden representation of inputs from prior layer is generated to form a higher abstraction of the original data. One can train the network by adjusting the mapping parameters, in order to obtain finer abstractions,or features, of the data, and use them for classification or regression tasks.

In this work, we explore the feasibility of improving energy expenditure prediction by applying deep learning on sensing data collected from wearable devices. Specifically, we implement a Convolutional Neural Network (CNN) to

978-1-4673-8325-7/15/$31.00 ©2015 IEEE

perform deep learning on raw 3-axis accelerometer data. Features automatically learnt by the CNN, together with subject-specific anthropomorphic features, are used by an ANN (a Backpropogation MultiLayer Perceptron) on top of the CNN to produce the EE estimate. We also compare this method with two other methods that learn universal and activity specific models respectively, using manually extracted features. The results show promising improvement even without rigorous fine tuning of the convolutional neural networks. In summary, the main contributions of this paper are as follows:

- We propose and implement a new approach of predicting energy expenditure which utilizes convolutional neural networks for automatic feature extraction instead of hand-crafted features, and demonstrate its effectiveness.
- We conducted a data collection campaign involving 30 volunteers performing daily ambulatory activities. Data includes reference EE values (from O_2 exchange), accelerometer traces, and heart rate traces among other anthropomorphic measurements recorded by COSMED K4b2 calorimeter and five wearable devices.
- We perform a comparative study of the performance of EE prediction based on the dataset, between our method and models used in other state-of-the-art methods.

One possible limitation of CNNs is their large run-time. However, most of this run-time is used in training the model over data-points. This is helpful since the model can be trained in the cloud and only the trained model can be used for off-line prediction in smart device or wearables. The trained model is very simple and can easily be computed in smartphones and wearables.

The rest of the paper is organized as follows: In Section II we give an overview of related work in the area. Section III presents our proposed approach for estimating energy expenditure with deep learning. Then, we discuss the experiment settings for evaluating our method and the results in Section IV. Section V concludes the findings and discusses directions for future work.

II. Related Work

A. Estimating Energy Expenditure with Wearable Sensors

Considerable effort has been invested in estimating energy expenditure from data collected by accelerometer and heart rate monitor. Major body of the research work follows two methodologies:

Some authors applied machine learning techniques directly to accelerometer features for predicting EE, by creating a universal model agnostic to activity types. For instance in [7] and [8] the authors feed selected accelerometer features into an artificial neural networks and estimate EE directly from the output. Others follow an activity-specific approach by performing activity recognition first, then applying different EE prediction methods depending on the activity type detected. Activity-specific models (ASMs) [9], [2] have been observed to obtain better performance than corresponding universal model [10]. Nonetheless activity recognition is performed over a pre-defined set of activities, it is not clear about the treatment to activities not included in the set.

All the existing work use handcrafted features selected by respective methods. In this work we rely on deep learning framework to automatically extract complex features from sensor data. For evaluation, we implement an universal model in this paper, however it is important to emphasize that the same approach can also be applied to learning activity-specific models using Deep-learning. More specifically, we use a Convolution Neural Network (CNN).

B. Deep Learning

Deep learning [11] is an emerging machine learning technique based on well developed concept of neural networks [12], augmented with series of improvements [13] in the structure and training of such networks. It was originally designed for image classification and recognition, modeled after the working of animal visual cortex. Deep learning has been successfully applied to various application domains, such as image recognition [5], natural language processing [14] and activity classification [6]. More notably, deep learning has been practically used for image search and powers Bing and Google image search engines [15]. There are a few recent studies on activity recognition utilizing deep learning framework to process accelerometer data [16], [17]. However to the best of our knowledge, our work is the first one targeted at improving the estimation of energy expenditure with deep neural networks.

III. Methodology

In this section we explain our method of estimating EE from sensed data and anthropometric features using deep learning technique called as Convolutional Neural Networks (CNN). We first describe pre-processing that converts continuous sensor trace into input vectors acceptable to CNN. Then we introduce general architecture of CNN used to automatically extract features and perform EE estimation with regression. Finally, we specify some of the implementation details and parameter choices.

A. Data Pre-Processing

The sensor data used in this study consists of three dimensional acceleration traces continuously sampled by 3-axis accelerometer at a fixed frequency of 50Hz, which is comparable to previous studies on activity recognition [18] as well as EE estimation [2]. To create the input vector, we applied a half sliding window of size 256 samples to the trace, equivalent to approximately 5.12 seconds of acceleration data. As a result the acceleration traces are transformed into a sequence of input vectors. Each sample in the input vector corresponds to a node in the input layer of CNN. We use the term *input vector* and *windowed trace* interchangeably.

Energy expenditure and heart rate measures are obtain by COSMED K4b2 calorimeter at a much lower frequency of approximately 0.5Hz. To make them compatible with acceleration data, each EE and HR samples are synchronized to closest instances in acceleration trace. Every accelerometer samples between the instances are assigned an approximated EE and

HR value by interpolation. After applying the sliding window, average EE is calculated for each input vector as *ground truth EE*, while average HR is calculated as a feature associated with the input vector.

In addition to *heart rate*, we consider anthropometric features including *age, height, weight, Basal Metabolic Rate (BMR)*. BMR is calculated according to revised Harris-Benedict formula [19], which takes into account the gender of subject. Apart from groundtruth EE, each input vector is also assigned an activity label using annotations made during data collection. If a vector contains samples with distinct activity annotations, it is assigned a special label indicating a transition period.

In summary, after the pre-processing we have three sequences of input vector, and one matching sequence of anthropometric vector of five features. Each vector in the sequences is associated with a groundtruth EE and an activity label.

B. Multi-Channel Convolutional Neural Networks

We adopt a similar construction of convolutional neural networks introduced in [16]. An example of overall architecture is illustrated in Figure 1. The convolutional neural networks mainly consists of two parts: feature extractor and MultiLayer Perceptron (MLP) regression.

1) Feature Extractor: The feature extractor takes input vector from input layer, applies transformations, and produces *feature maps* between layers and as input to succeeding MLP layer. Feature extractor we use is a neural networks of two layers, where each layer performs transformation of three stages: filter, activation, and pooling.

- **Filter**: At this stage a number of linear filters are applied along the input vector x_i, in order to capture the temporal local correlation within input vector. The kth filter at a given layer is determined by weights W^k and bias b^k. Size of the filter f is much smaller than that of input vector i, and it decides the time window in which local correlation is extracted. Such correlation is calculated by convolution operation of input vector with the filter, as $W^k_{ij} * x_i + b^k_j$ where $*$ denotes convolution operator, i is the size of input, and $j = i - f + 1$ is the output size.
- **Activation**: At this stage an activation function is applied to respective output of filter stage, introducing non-linearity into the neural networks so that complex patterns can be discovered from input. Popular choices of activation function including $sigmoid()$ and $tanh()$ and we choose the later in this work. This stage can be conveniently merged with filter stage by outputing feature map $h^k_{ij} = tanh(W^k_{ij} * x_i + b^k_j)$.
- **Average Pooling**: At this stage a form of non-linear down-sampling called pooling is applied to the feature map. The feature map is partitioned into equal size non-overlapping subsequences, and outputs average or maximum value for each subsequence, depending on average-pooling or max-pooling is selected. As a result the feature map is down-sampled by specified factor, with the main purpose of reducing the dimensionality of intermediate representation as well as providing robustness to small variations.

At each layer, several feature maps are generated depending on the number of filters specified. By cascading the layers, a hierarchical feature map is learnt from the feature extractor. When processing data from 3-axis accelerometer, each axis constitutes a channel, and to which a dedicated feature extractor is applied.

2) MultiLayer Perceptron Regression: After feature maps are generated for each accelerometer channel by the multi-channel feature extractor, they are flattened, and combined into single feature vector together with the anthropometric vector. This feature vector can then be fed to MLP regression layer. Different from standard MLP classifier, we remove the sigmoid activation function at hidden layer of MLP as it effectively becomes a linear transformation layer. We found this modification produces better result than using sigmoid function. Furthermore, the output layer of MLP is replaced by a linear regression layer instead of logistic regression, so that a numerical prediction of EE is produced rather than likelihood of classes.

3) Parameter Learning with Gradient Descent: Trainable parameters in the CNN framework mainly consists of weights and bias of feature extractor filters, and those of the hidden layer and linear regression layer of Multilayer Perceptron (MLP). MLP is also called as Backpropagation Artificial Neural Network (ANN). Both ANN and MLP are used interchangeably in this paper.

Loss function for the model is Roor Mean Square Error (RMSE) between predicted EE values ($\overline{EE}$) and the ground truth (EE), where n is the number of windows in trace.

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(\overline{EE_i} - EE_i)^2}{n}} \tag{1}$$

To train the parameters, we use the typical stochastic gradient descent optimization that minimizes the loss function. Standard training process consists of three iterative stages: Feed-forward, Backpropagation, and Gradient Applied. At first all parameters in the feature extractor layers and the MLP hidden layer are initialized randomly, while those of the MLP linear regression layer initialized to zero. In the feed-forward stage input vectors are run pass the networks and produce an output of predicted EE value. At the backpropagation stage, error is calculated according to the cost function and propagated backwards along the networks, as contribution to the error calculated at each layer. Given the backpropagation errors, each layer update their parameters with gradient descent, thus minimizing the overall error in a greedy way.

Note that in practice an unsupervised "pre-training" phase is usually applied to bootstrap hidden nodes and initialize network with better parameters, which may observe significant improvement to the training performance [20]. However we limit our model learning to standard form in this work and leave further exploration to future study.

C. Implementation

We implement aforementioned framework with building blocks provided by deep learning library Theano [21]. The

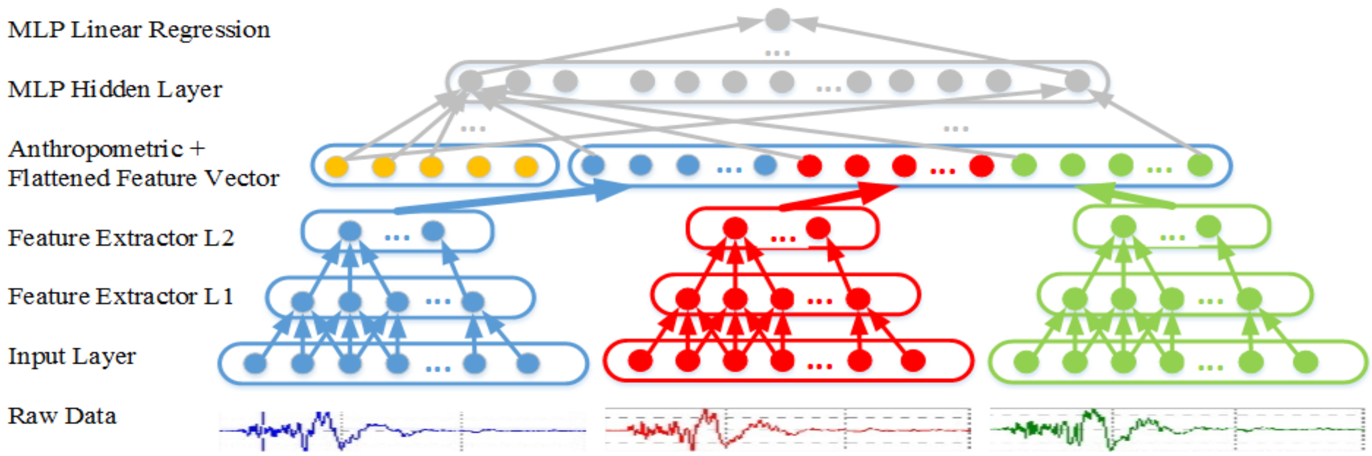

Fig. 1. Architecture of Multi-Channel Convolutional Neural Networks

TABLE I
PHYSICAL CHARACTERISTICS (MEAN±STDEV (MIN - MAX))

Number of Subjects	30 (20 males, 10 females)
Age(year)	$27.8 \pm 6.9 (19-45)$
Height(cm)	$168.9 \pm 10.8 (150-189)$
Weight(kg)	$66.9 \pm 15.1 (48-105)$
BMR($kcal/day$)	$1610.5 \pm 271.5 (1182-2214)$

implementation has 3 channel feature extractor, corresponding to accelerometer input vector at 3 axes. The feature extractor consists of two cascaded layers. The first layer employs 8 filters of size 5 and a pooling factor of 2, while the second layer has 4 filters of size 5 and same pooling factor. Number of nodes in the MLP hidden layer is 400.

IV. Experiments and Results

In this section we demonstrate the effectiveness of our method through a comparison study. First we describe the dataset used for evaluation. Then we introduce the evaluation metric and two benchmark methods we compare with. Quantitative results and analyses are shown next.

A. Dataset

1) Subjects and Data Collection: We recruited a total of 30 participants (10 females and 20 males) for the study. All participants are physically fit for performing ambulatory activities described below. Physical characteristics of all participants are summarized in Table I. The study is approved by Institutional Review Board at the University of California, Davis. Participants are recruited from the university campus and the medical center, and all have given oral consent prior to the experiment in accord with IRB regulation.

Before each test, the COSMED K4b2 components were calibrated according to the manufacturers instructions. Subject also carries a portable COSMED K4b2 indirect calorimeter for collecting EE groundtruth, as well as heart rate measurements, at a frequency of 0.5Hz. A smartphone with tri-axial accelerometer was placed in a waist pouch at a specific orientation. All sensor devices are time synchronized before collection. Cosmed samples are synchronized with sensor data following methods mentioned in previous section.

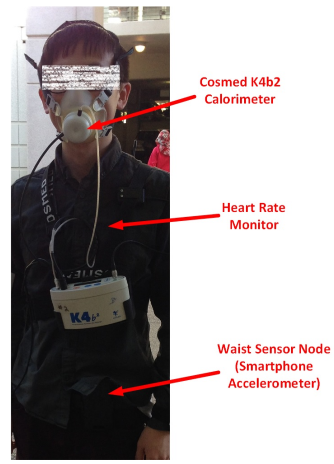

Fig. 2. Data collection platform of COSMED calorimeter, heart rate monitor and smartphone sensors

2) Protocol: We design a semi-controlled protocol for this data collection. Participants are asked to perform a sequence of pre-defined types of ambulatory activities, following specified routines. However the environment where the collection is conducted is inside a regular hospital building and the courtyard outside, with moderate traffic. In this work we study six major activities as described in Table II, which emulates ordinary behaviors of a person in ambulatory environment. The participants are asked to perform the activities as natural as possible, and no rest between activities in each part. Completion of the whole sequence will take each participant approximately 30 minutes. A supervisor accompanying the participant will manually time the instance when a activity transition happens. Annotation of the activities is performed later offline. The COSMED K4B2 indirect calorimeter reports averaged energy expenditure estimate over approximately 30-60 seconds interval. According to the manual, energy expenditure is calculated with abbreviated Weir equation:

$$RMR = (3.94 * VO_2 + 1.106 * VCO_2) * 1.44 \quad (2)$$

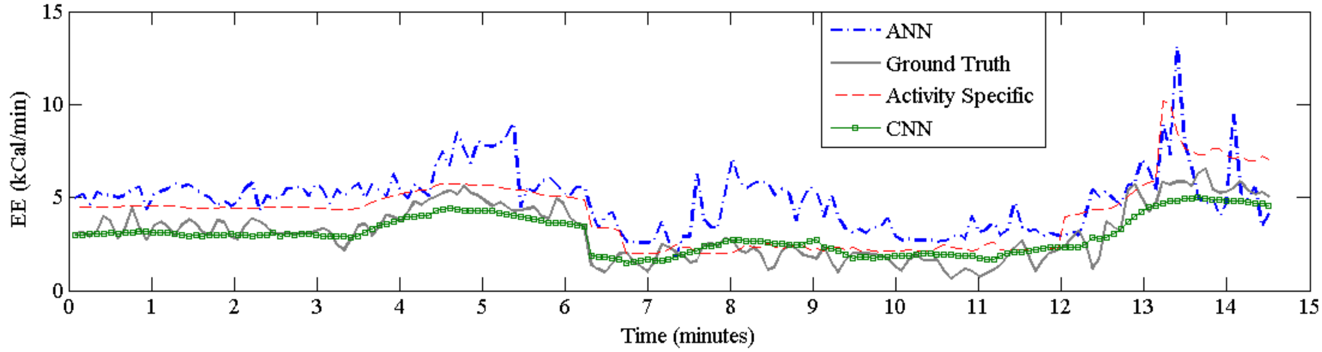

Fig. 3. Visual Trace of EE Prediction (Subject No.8): Comparing proposed method (CNN), existing methods (Activity Specific [2] and ANN [3] against ground truth (COSMED)

TABLE II
AMBULATORY ACTIVITIES IN THE STUDY

Activity Label	Activity Cluster
Walking on Ground	Walking
Stairs Climbing	Walking
Running	Running
Static(Standing/Sitting)	LWBM
Riding Elevator	LWBM

, where $RMR(Kcal/min)$ is the resting energy expenditure, $VO_2(ml/min)$ is the oxygen uptake, and $VCO_2(ml/min)$ is Carbon Dioxide production.

B. Evaluation Settings and Metric

To evaluate the effectiveness of proposed model, we performs leave-one-out cross validation on the collected dataset. At each iteration one subject's data is used as testing set, while combined data from the rest of users is used for training the model. It iterates through all subjects, and the average testing error of all iterations are reported. We use Root Mean Squared Error (RMSE) as a standard evaluation metric.

In order to assess the potential improvement provided by the convolutional neural networks, over conventional models using handcrafted features, we also implement two state-of-the-art methods for estimating energy expenditure: an activity-specific linear regression model proposed in [2], and a universal model based on ANNS proposed in [3].

Same data preprocessing are used for all methods. To be compatible with activity-specific models, we adopt the same activity clusters classification as in [2] and partition the dataset according to different activity types in Table II. Further, we assume perfectly correct activity classification and use information from activity-labels to classify activity. To train and test an activity-specific model, only data from targeted activity cluster is used. In addition, the same feature sets selected by the author [2] for respective activity clusters are used. For the universal model, we extract all feature vector specified in [3], and use Artificial Neural Networks(ANN) to perform the regression. The methods for comparison are implemented in JAVA for feature extraction, after which it invokes Weka API [22] for training and testing. Same leave-one-out cross validations are performed, and average RMSE reported. LWBM refers to low whole body motion activities such as static standing/ sitting/ riding elevator.

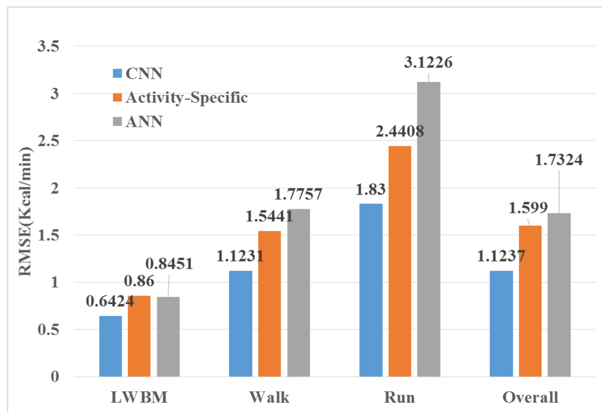

Fig. 4. Overall and By-Cluster Performance of EE Prediction

C. Evaluation Results

Figure 3 gives a plot with representative performance of different algorithms for a subject. The x axis shows time (in minutes) while y-axis shows the EE values in kCal per minute. The ground truth is obtained using Cosmed k4b2 calorimeter. It can be seen that CNN matches most closely to the ground truth while ANN diverges widely (for some time-periods). In this specific individual CNN gives 23.44% error to ground truth while Activity-specific linear model gives 44.44% error and ANN gives 85.41% error (These values are for this subject only).

The average RMSE of CNN (over all participants) is also lower than other models. Figure 4 shows the performance of the three methods in terms of RMSE. Overall, CNN gives a RMSE of 1.12 which is less than Activity-Specific [2] model (1.59) and ANN [3] (1.73) by 30% and 35% respectively. The overall improvements are similar across all activity clusters - LWBM (25%,23%), Walk (27%, 36%) and running (25%,40%) respectively. Running is one of the most intense activity and the estimation error is also significantly high in this case. However, CNN is able to reduce the RMSE to 1.8.

A further comparison of CNN with Activity-Specific model is performed in Figure 5. It shows results for walking on ground as well as stairs. It can be seen that CNN consistently outperforms Activity-Specific model. Results of low whole body motion activities are shown in Figure 6.

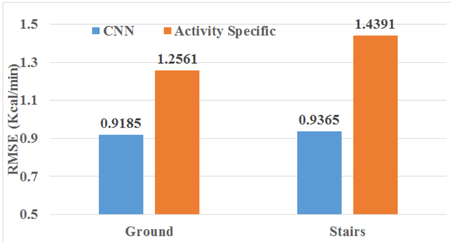

Fig. 5. Performance of EE Prediction in Walking Cluster

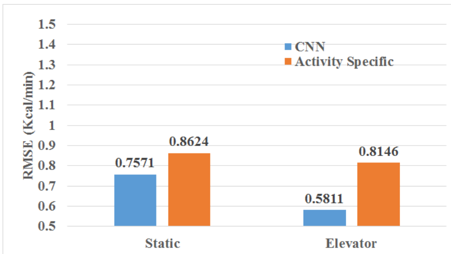

Fig. 6. Performance of EE Prediction in LWBM Cluster

V. Conclusion and Future Work

In this work, we proposed the use of deep-learning technique based on CNN for EE estimation in individuals using wearable sensors. Our evaluation gives low RMSE (1.12) in EE estimation using CNN with significant improvements over state-of-the-art methods including ANN (35%) and Activity-specific models (30%), suggesting a promising new approach of applying data mining technique in improving energy expenditure estimation.

Motivated by these results, we plan to evaluate our model on a larger data corpus, as well as implement a real application which can be used in smart-watches to reliably estimate EE. Model training with CNN requires tremendous computational resources at current stage, thus the goal of this work is to construct minimal number of models that can be applied to a vast population without compromising estimation accuracy too much for individuals. However we are also working on personalized model learning that customize the model at individual level based on the general model from CNN, which we expect to further improve efficiency as well as accuracy. One main limitation of our work is that is works based on fixed location of accelerometer sensor i.e. at a waist pouch. Although we have also collected data from arm, wrist, thing, ankle, and used them for CNN learning, the preliminary result didn't show much difference from waist results reported in this work. We will continue to look into potential ways to integrate additional data so that substantial positive impact can be achieved.

References

[1] Y. Wang, M. A. Beydoun, L. Liang, B. Caballero, and S. K. Kumanyika, "Will all americans become overweight or obese? estimating the progression and cost of the us obesity epidemic," *Obesity*, vol. 16, no. 10, pp. 2323–2330, 2008.

[2] M. Altini, J. Penders, and O. Amft, "Energy expenditure estimation using wearable sensors: A new methodology for activity-specific models," in *Proceedings of the Conference on Wireless Health*, ser. WH '12. New York, NY, USA: ACM, 2012, pp. 1:1–1:8.

[3] A. Pande, Y. Zeng, A. K. Das, P. Mohapatra, S. Miyamoto, E. Seto, E. K. Henricson, and J. J. Han, "Energy expenditure estimation with smartphone body sensors," in *Proceedings of the 8th International Conference on Body Area Networks*, ser. BodyNets '13. ICST, 2013, pp. 8–14.

[4] A. Pande, G. Casazza, A. Nicorici, E. Seto, S. Miyamoto, M. Lange, T. Abresch, P. Mohapatra, and J. Han, "Energy expenditure estimation in boys with duchene muscular dystrophy using accelerometer and heart rate sensors," in *Healthcare Innovation Conference (HIC), 2014 IEEE*, Oct 2014, pp. 26–29.

[5] D. Ciresan, U. Meier, and J. Schmidhuber, "Multi-column deep neural networks for image classification," in *in proceedings of CVPR 2012*, 2012, pp. 3642–3649.

[6] S. Ji, W. Xu, M. Yang, and K. Yu, "3d convolutional neural networks for human action recognition," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 35, no. 1, pp. 221–231, Jan 2013.

[7] P. S. Freedson, K. Lyden, S. Kozey-Keadle, and J. Staudenmayer, "Evaluation of artificial neural network algorithms for predicting mets and activity type from accelerometer data: validation on an independent sample," *J Appl Physiol (1985)*, vol. 111, no. 6, pp. 1804–12, 2011.

[8] M. P. Rothney, M. Neumann, A. Beziat, and K. Y. Chen, "An artificial neural network model of energy expenditure using nonintegrated acceleration signals," *J Appl Physiol (1985)*, vol. 103, no. 4, pp. 1419–27, 2007.

[9] F. Albinali, S. Intille, W. Haskell, and M. Rosenberger, "Using wearable activity type detection to improve physical activity energy expenditure estimation," in *Proceedings of the 12th ACM International Conference on Ubiquitous Computing*, ser. UbiComp '10. New York, NY, USA: ACM, 2010, pp. 311–320.

[10] A. G. Bonomi and G. Plasqui, ""divide and conquer": assessing energy expenditure following physical activity type classification," *J Appl Physiol (1985)*, vol. 112, no. 5, p. 932; author reply 933, 2012.

[11] L. Deng and D. Yu, "Deep learning: Methods and applications," *Found. Trends Signal Process.*, vol. 7, no. 3–4, pp. 197–387, Jun. 2014.

[12] W. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *The bulletin of mathematical biophysics*, vol. 5, no. 4, pp. 115–133, 1943.

[13] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Comput.*, vol. 18, no. 7, pp. 1527–1554, Jul. 2006.

[14] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," in *Advances in Neural Information Processing Systems*, 2014, pp. 3104–3112.

[15] Q. V. Le, R. Monga, M. Devin, K. Chen, G. S. Corrado, J. Dean, and A. Y. Ng, "Building high-level features using large scale unsupervised learning," in *In International Conference on Machine Learning, 2012. 103*, 2012.

[16] Y. Zheng, Q. Liu, E. Chen, Y. Ge, and J. Zhao, "Time series classification using multi-channels deep convolutional neural networks," in *Web-Age Information Management*, ser. Lecture Notes in Computer Science. Springer International Publishing, 2014, vol. 8485, pp. 298–310.

[17] M. Zeng, L. T. Nguyen, B. Yu, O. J. Mengshoel, J. Zhu, P. Wu, and J. Zhang, "Convolutional neural networks for human activity recognition using mobile sensors," 2014.

[18] L. Bao and S. Intille, "Activity recognition from user-annotated acceleration data," in *Pervasive Computing*, ser. Lecture Notes in Computer Science, A. Ferscha and F. Mattern, Eds., vol. 3001. Springer Berlin Heidelberg, 2004, pp. 1–17. [Online]. Available: http://dx.doi.org/10.1007/978-3-540-24646-6_1

[19] A. M. Roza and H. M. Shizgal, "The harris benedict equation reevaluated: resting energy requirements and the body cell mass." *The American Journal of Clinical Nutrition*, vol. 40, no. 1, pp. 168–82, 1984.

[20] Y. Bengio, P. Lamblin, D. Popovici, H. Larochelle, U. D. Montral, and M. Qubec, "Greedy layer-wise training of deep networks," in *In NIPS*. MIT Press, 2007.

[21] F. Bastien, P. Lamblin, R. Pascanu, J. Bergstra, I. J. Goodfellow, A. Bergeron, N. Bouchard, and Y. Bengio, "Theano: new features and speed improvements," Deep Learning and Unsupervised Feature Learning NIPS 2012 Workshop, 2012.

[22] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The weka data mining software: an update," *ACM SIGKDD Explorations Newsletter*, vol. 11, no. 1, pp. 10–18, 2009.