IEEE Sensors Letters VOL. 6, NO. 9, SEPTEMBER 2022

IEEE Sensors Council

Khail-Net: A Shallow Convolutional Neural Network for Recognizing Sports Activities Using Wearable Inertial Sensors

Hamza Ali Imran

Design Engineer, Embedded Systems, aXelerated Software, Emumba Private Limited, Islamabad 44000, Pakistan

Manuscript received 7 July 2022; accepted 5 August 2022. Date of publication 9 August 2022; date of current version 23 August 2022.

Abstract—Time-series data can be generated by wearable sensors such as accelerometers, gyroscopes, and magnetometers. This data may be used to classify various everyday life activities using machine learning or deep learning models. Athletics, education, child monitoring, ambient assisted living, and other applications benefit from human activity recognition. Human activity recognition includes sports activity recognition. A typical sports activity is any action that is often employed in a variety of sports. Walking, jogging, sprinting, and leaping are basic sports motions. A unique sports action exists exclusively in one sport, such as a badminton smash. We proposed a shallow convolution neural network for sports activity recognition. It has just 1251 trainable parameters. The test accuracy attained is 98.387%. The average F1 score, recall, and precision are 98.0, 98.7, and 98.0%, respectively. We have also trained the model on a benchmark human activity recognition dataset from WISDM lab for performance evaluation and comparison of the model.

Index Terms—Sensor signal processing, convolutional neural network, edge computing, embedded sensors, human activity recognition (HAR), inertial sensors, inertial sensors signal processing, shallow network, smart-wearable.

I. INTRODUCTION

Human activity recognition (HAR) based on wearable sensors has emerged as a new research area in the fields of artificial intelligence and pattern recognition and numerous applications benefit from HAR, including sports activity detection, smart homes, and health support, among others. Sensors such as accelerometers, gyroscopes, and magnetometers, can produce time-series data for HAR. Traditional approaches, which were established to aid in the identification of human action, are under the scope of supervised learning. For example, previous approaches such as SVM and random forest need handmade characteristics to be extracted as classifier inputs. Deep learning, namely, convolutional neural networks (CNNs), has now become widely applied in the field of HAR [1], [2].

Sports activity is classified into two types: 1) general activity and 2) particular activity. Everyday sports activity includes any action that is often employed in a variety of sports. Walking, jogging, sprinting, and leaping are basic athletic movements. While particular sports activity is described as an activity that exists specifically in one form of sport, such as a smash in badminton, a slice in tennis, a dribble in hockey, and other types of activity for a different kind of sport, with the rising studies of human activity detection utilizing inertial sensors in recent years, many researchers have begun to pay attention to athletic activity recognition [3], [4]. The benefits of athletics activity recognition might be manifold. For example, it can help players and coaches improve their performance, physiotherapists avoid player injuries, and journalists obtain data such as the number of shots on goal.

In this letter, we have proposed a shallow convolution neural network that has only 1251 trainable parameters only. We have named it "Khail-Net," where the word "Khail" is an Urdu team which means "game or sports." We have experimented with it on a sports activities dataset named IM-SportingBehaviors dataset collected and made public by Air University Pakistan [5]. The achieved test accuracy is 98.387%. The achieved average F1-score, recall, and precision are 98%. The details of the dataset are presented in Section III-C.

The main contributions of this letter are as follows:

1) We have Edge-friendly shallow CNN, which is named as Khail-Net. It works on raw inertial sensors data for sport activity recognition which has only 1251 trainable parameters.
2) The presented model has being evaluated IM-SportingBehaviors dataset [5].
3) Model has been evaluated on a human activity recognition dataset WISDM, i.e., 2011/WISDM activity prediction [6] for performance comparison of classifying between classes based on inertial sensors data.

The rest of the letter is organized as follows: Section II shares details of related work. Section III talks about details of the presented model, the dataset used, segmentation strategy, and entire end-to-end pipeline of the project. Section IV shares details of the results obtained. Finally, Section V concludes this letter.

II. RELATED WORK

The use of inertial sensors for HAR is becoming prevalent these days. A couple of studies are published in recent times that make use of wearable inertial measurement unit (IMU) sensor data for HAR. Imran and Latif [7] proposed a novel model which has inception kind of modules with dense links in between them for HAR using smartphone IMU sensors data. Imran [2] proposed an antithesis kind of model which has bidirectional GRU units before multitype of convolutional units. The model was named as UltaNet. This model is state-of-the-art for HAR with smartphone IMU data. Mehmood et al. proposed a model which has dense links between convolutional kernels [8]. It also made use of smartphone IMU data for HAR. Zhang et al. [9] attempted to use an attention mechanism for HAR based on IMU data of smartphones. Deep human activity and location recognition is the building of a monolithic two-stream CNN for predicting both human behaviors and different sensor locations by Lawal et al. [10]. The device localization

Corresponding author: Hamza Ali Imran (e-mail: himran.mscs18seecs@seecs.edu.pk).
Associate Editor: W. Pu.
Digital Object Identifier 10.1109/LSENS.2022.3197396

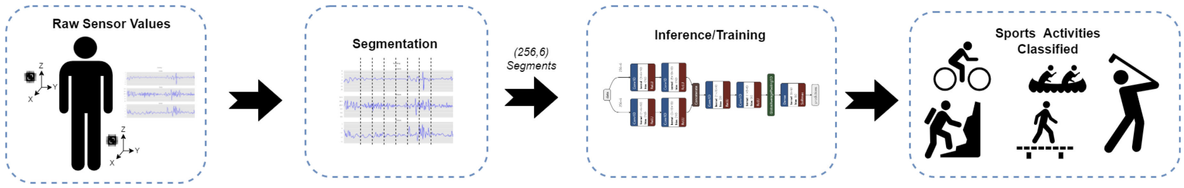

Fig. 1. End-to-end pipeline of the entire project.

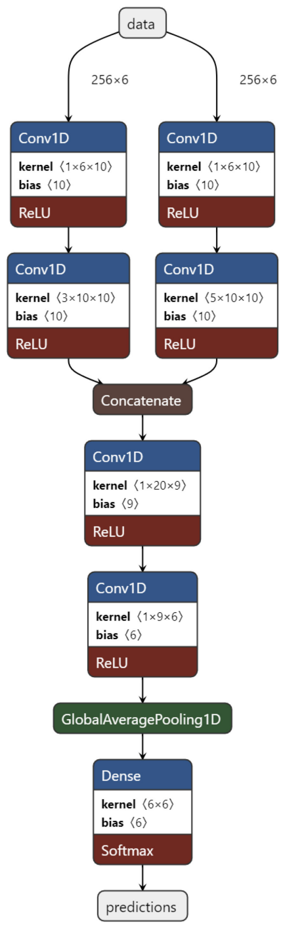

Fig. 2. Khail-Net: Proposed neural network.

permits the examination of the influence of location information on the accuracy of activity recognition. Hsu et al. [4] tried athletic activity recognition; each participant was encouraged to put the wearable inertial sensor network on their wrist and ankle and perform the eleven categories of sports activities in a laboratory setting. For classification, traditional machine learning methods were applied. 99.55% accuracy. There has been a little published study on athletic or sports activity recognition. In general, significant contributions are made to normal daily activities. Our study is one of the first to present a fully deep learning approach and a model that is shallow in nature and hence edge-friendly in the field of athletic or sports activity recognition.

III. PROPOSED APPROACH

Fig. 1 shows the ends-to-end pipeline of the project. The use of smart-wearable is becoming predominant these days. We have made an attempt to classify between different sports activities using the data from the inertial sensors placed on two body parts including the wrist and knee of the athlete. The dataset used has the data of three body parts but there are some missing entries for the sensors placed on the neck so we have not incorporated that sensors data in our study. The target of this study was to come up with a less complex model so that it can be deployed on the Edge devices. The model present has only 1251 trainable parameters. The 6-D input data which includes the 3-D accelerations data form the IMU attached to the wrist and the 3-D accelerations data from the IMU attached to the knee is given to two different sets of kernels. The first set has 10 (1×1) kernels followed by 10 (1×3) kernels and the second set has 10 (1×1) kernels followed by 10 (1×5) kernels. The feature set generated by the two sets are then concatenated and we apply 9 (1×1) kernels to them followed by 6 (1×1) kernels. The point to be noted is that at the end of the mode we have reduced the numbers of the kernels to exact number of classes which is 6. After applying 6 (1×1) kernels we have applied global average pooling to further reduce the size of the feature-set. After that softmax is applied for classification. Fig. 2 presents the model diagram.

A. Model's Motivation

The model's major motivation was to provide an edge-friendly solution that could be deployed on the wearable devices that generated the data. Because recent literature supports the use of CNNs for classification of human activities, we aimed to avoid any recurrent neural units in the model to keep it less complex. The number of kernels and the type of kernels were chosen empirically. The model was most inspired by HHARNet a model by Imran and Latif [7]. This model recognizes human activity using inertial sensor data from a smartphone and does not include any recurrent units. The model also used global average pooling to reduce the number of kernels to the exact number of classes.

B. Segmentation Size and Step Size

Because a DNN has a fixed input size, segmentation of the input signal is required. The size of the segmentation window is also significant [2]. Empirical investigation revealed that a size of 256 with a step size of 16 performed best. These parameters were tuned once the model hunt was completed. Initially, we set these values as 128 for window size with step size of 32.

C. Dataset Used

A team from Air University Pakistan acquired and made available a sporting behaviors dataset (IM-SportingBehaviors) [5] that captured important aspects of human mobility using triaxial accelerometers attached to the subject's wrist, knee, and below neck region. The dataset includes motion data taken during six sports: badminton, basketball, cycling, football, skipping, and table tennis. Professional and amateur athletes aged 20–30 with weights ranging from 60–100 kgs were among those who participated. Due to the multisensor environment and interbehavioral similarities, the dataset itself presents a considerable

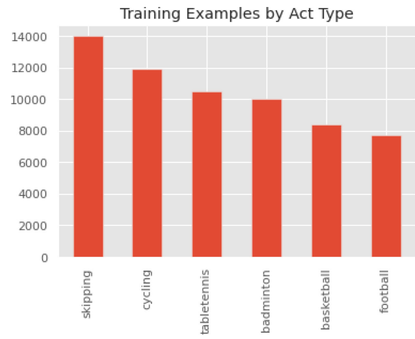

Fig. 3. Class distribution of the used dataset.

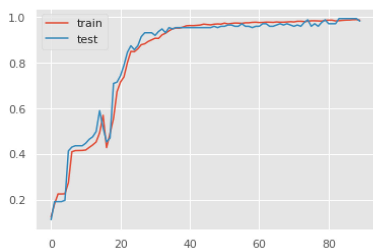

Fig. 4. Accuracy plot for test and training data.

amount of challenges. We have divided the dataset into three sets, which consists of 70% training, 15% for testing, and the remaining 15% for cross-validation. Fig. 3 shows the class distribution.

IV. RESULTS

A. Sports Activity Recognition

The number of epochs was set to 90 and the batch size selected was 120. These values were selected empirically by observing the accuracy and loss plots for cross-validation and training data. Fig. 4 presents the accuracy plot. The optimizer we used was "Adam." The achieved cross-validation accuracy is 98.39% and the test accuracy is 98.30%. The entire performance report is shown in Fig. 7. It can be observed that the average F1 score, recall, and precision are 98%. The performance parameters for each class are also shown in the same figure. It can be seen that the minimum recall and F1 score are for the "basketball" class. A total of nine samples of this class are confused with "badminton." As reported in the confusion matrix shown in Fig. 5. Although the confusions are very low, but still, the hypothesis is that the kind of movements involved in the two joints where sensors are attached is similar in the case of these two games. This makes it difficult for the model to classify between them. Considering the remaining confusion matrix we can see there is negligible confusion between classes which results in good overall accuracy and F1 scores. Fig. 6 represent the entire performance report.

B. Human Activity Recognition

Very limited work is done on sports activity recognition based on wearable or inertial sensors and therefore, to the best of our knowledge no other dataset for sports activity recognition is available openly. Considering this situation for comparison of the performance of the model, we trained our model on one of the most famous benchmark datasets from the WISDM lab for human activity recognition, namely, the WISDM 2011/WISDM Activity Prediction dataset [6].

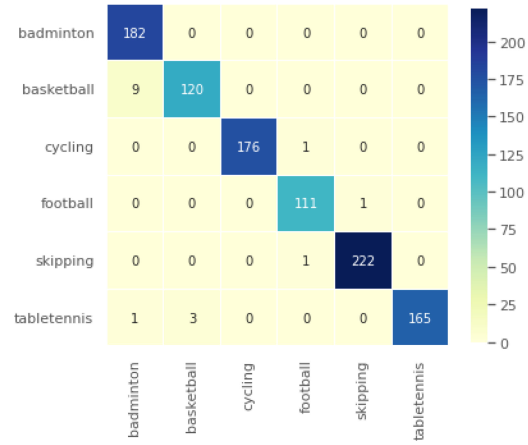

Fig. 5. Confusion matrix.

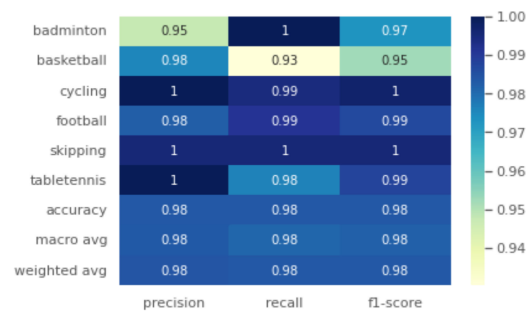

Fig. 6. Performance report for model classification.

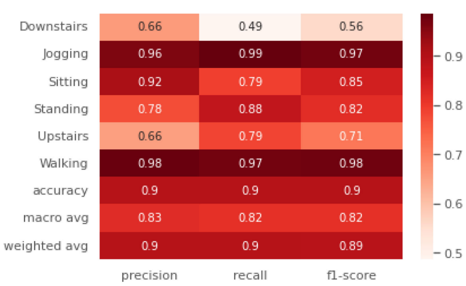

Fig. 7. Performance report for model classification for WISDM 2011 HAR dataset.

Table 1. Performance Comparison

Type & Reference	Accuracy
[12] CNN	95.26%
[13] LSTM-CNN	95.85%
[9] CNN	94.65 %
[14] CNN	93.32%
[15] CNN with an attention mechanism	96.4%
[8] BiGRU-CNN	97.2%
Proposed Model	89.60%

This dataset has data from a smartphone accelerometer for six different daily activities. The smartphone was attached to the volunteer's chest for the collection of the data. All other settings are kept the same including the sampling window size, number of epochs, optimizers, etc. The details of the dataset are well elaborated in [13]. Table 1 does comparison of this model for the WISDM 2011 dataset with recently published studies. It can be seen that the accuracy archived is less than the state-of-the-art model, i.e., *UltaNet* but the complexity of the model is relatively very high. The model *UltaNet* has a total of "35 492" parameters, of which "35 236" are trainable and "256" are nontrainable. While presented being Edge-friendly has only 1251 trainable parameters with no nontrainable parameter. This makes the presented model 28.37 times less complex. Moreover, it was tuned for sports activities and for the sake of comparison was trained for the HAR dataset with no modifications or tunning in the training process or model itself.

V. CONCLUSION

Human activity recognition is an active area of search these days due to its applications in various domains. In this study, we have focused on sports activity recognition and have proposed a neural network that is edge-friendly and dose classification based on sensors placed on wrest and knee. A total of six activities are classified with an accuracy of 98.387%.

REFERENCES

[1] K. Wang, J. He, and L. Zhang, "Attention-based convolutional neural network for weakly labeled human activities' recognition with wearable sensors," *IEEE Sensors J.*, vol. 19, no. 17, pp. 7598–7604, Sep. 2019.

[2] H. A. Imran, "UltaNet: A antithesis neural network for recognizing human activity using inertial sensors signals," *IEEE Sens. Lett.*, vol. 6, no. 1, Jan. 2022, Art no. 7000304.

[3] E. Mitchell, D. Monaghan, and N. E. O'Connor, "Classification of sporting activities using smartphone accelerometers," *Sensors*, vol. 13, no. 4, pp. 5317–5337, 2013.

[4] Y.-L. Hsu, S.-C. Yang, H.-C. Chang, and H.-C. Lai, "Human daily and sport activity recognition using a wearable inertial sensor network," *IEEE Access*, vol. 6, pp. 31715–31728, 2018.

[5] A. Jalal, M. A. K. Quaid, and A. S. Hasan, "Wearable sensor-based human behavior understanding and recognition in daily life for smart environments," in *Proc. IEEE Int. Conf. Front. Inf. Technol.*, 2018, pp. 105–110.

[6] J. W. Lockhart, G. M. Weiss, J. C. Xue, S. T. Gallagher, A. B. Grosner, and T. T. Pulickal, "Design considerations for the WISDM smart phone-based sensor mining architecture," in *Proc. 5th Int. Workshop Knowl. Discov. Sensor Data*, 2011, pp. 25–33.

[7] H. A. Imran and U. Latif, "HharNet: Taking inspiration from inception and dense networks for human activity recognition using inertial sensors," in *Proc. IEEE 17th Int. Conf. Smart Communities: Improving Qual. Life Using ICT*, 2020, pp. 24–27.

[8] K. Mehmood, H. A. Imran, and U. Latif, "HardenseNet: A 1D densenet inspired convolutional neural network for human activity recognition with inertial sensors," in *Proc. IEEE 23 rd Int. Multitopic Conf.*, 2020, pp. 1–6.

[9] H. Zhang, Z. Xiao, J. Wang, F. Li, and E. Szczerbicki, "A novel IoT-perceptive human activity recognition (HAR) approach using multihead convolutional attention," *IEEE Internet Things J.*, vol. 7, no. 2, pp. 1072–1080, Feb. 2020.

[10] I. A. Lawal and S. Bano, "Deep human activity recognition with localisation of wearable sensors," *IEEE Access*, vol. 8, pp. 155060–155070, 2020.

[11] K. Xia, J. Huang, and H. Wang, "LSTM-CNN architecture for human activity recognition," *IEEE Access*, vol. 8, pp. 56855–56866, 2020.

[12] A. Ignatov, "Real-time human activity recognition from accelerometer data using convolutional neural networks," *Appl. Soft Comput.*, vol. 62, pp. 915–922, 2018.

[13] H. A. Imran, K. Hamza, and Z. Mehmood, "Harresnext: An efficient ResNext inspired network for human activity recognition with inertial sensors," in *Proc. 2nd Int. Conf. Digit. Futures Transformative Technol.*, 2022, pp. 1–4.