

Shanakht-Net: Person Re-identification using Inertial Sensors data generated by Smart-wearables from Daily Activities

Hamza Ali Imran^{1*}

¹ Design Engineer, Embedded Systems, aXelerated Software, Emumba Private Limited, Islamabad Pakistan

Manuscript received

Abstract—An important component of an automated surveillance environment is person re-identification. The problem is often addressed using data received from vision sensors using appearance-based features, which are heavily reliant on visual cues such as colour, texture, and so on, limiting the reliability of re-identification of an individual. Much research has been performed to solve the problem of re-identification utilising human gait using inertial measurement units (IMU) data, which is thought to be unique and offer a distinct biometric signature that is especially ideal for re-ID in uncontrolled conditions. The locomotive activity of walking was the primary emphasis. The current study utilised not only locomotive activities but also non-locomotive activities of daily living. The data was obtained from the WISDM lab. The data is collected while engaging in six distinct everyday activities. The dataset was originally gathered for the purpose of Human Activity Recognition. Nonetheless, each person is given a unique ID. This information was utilised to re-identify the individual. The dataset consists of data of 36 volunteers. Shanakht-Net, a novel convolutional neural network, is introduced. The F1-score obtained is 93%. Precision, recall, and accuracy are assessed and reported as well.

Index Terms—Person re-identification, Non-locomotive motion, Convolutional neural networks, deep learning

I. INTRODUCTION

By virtue of rapid development in hardware technologies, most modern digital devices such as smartphones, smartwatches, wearable fitness bands etc. are equipped with a wide range of on-board sensors including inertial sensors. An inertial sensor usually in-houses tri-axial accelerometer, tri-axial gyroscope, and tri-axial magnetometer and measures accelerations, angular velocities, and earth's magnetic north respectively. These modern gadgets are easy to mount on the human body and can record low level kinematics at higher frequencies. Such low level inertial data have been successfully used in a number of applications including human activity recognition (HAR) [1], sports activity recognition [2], fall detection [3], identifying human emotions [4], estimation of human soft biometrics [5] and person re-identification [6], etc.

Person re-identification (re-ID) is an important and ongoing research topic with applications in automated surveillance and monitoring, robotics, human-computer interaction, and digital forensics. The purpose of the study is to re-identify a person by establishing a relationship between attributes of the same individual obtained at different locations and periods. Vision sensors are mostly utilised for human re-identification (re-ID) applications. The disadvantage of appearance-based approaches is that they make re-ID systems very dependent on visual cues such as colour, texture, and so on, limiting the ability for just short-term re-ID. Similarly, in practise, these image- or features-based re-ID techniques are only appropriate for use in controlled circumstances where the change in appearance (e.g., genuine change in clothes, etc.) has a major influence on and negatively affects their performance. Furthermore, difficulties such as picture blurring, optical motion, gender effects, fluctuating lighting, and so on provide additional obstacles in extracting critical

characteristics for unique representation, resulting in lower re-ID accuracy. Very recently some initial attempts has being made to re-identify a person based on inertial sensors data generated by wearable sensors and smart-wearable [6], [7]. But the main focus of these studies is on the use of locomotive activities i.e. human gait. The main contribution of this article is that we are attempting to make use of both locomotive and non-locomotive activities for person re-identification. A novel convolutional neural network named "Shanakht-Net" is presented. The word "Shanakht" is an Urdu team that mean "Identity or to Identify". The presented model has "95,640" trainable and "128" non-trainable parameters. Summarizing the key contributions of this letter as follows:

- We presented a convolutional neural network, which is named as Shanakht-Net. It works on raw inertial sensors data Daily-Life Activities which includes both locomotive and non-locomotive activities for person re-identification.
- The presented model has being evaluated WISDM 2011 dataset. Which was originally collected for Human Activity Recognition [8].

The remaining of the article is such, section II shares details on proposed approach which includes details of dataset used, presented neural network and segmentation methodology. Section III presents details of the results of experiments conducted. Section ?? concludes the article with future directions.

II. PROPOSED APPROACH

Smart-wearables are becoming increasingly popular, and they include a number of embedded sensors that may be used for a variety of applications. The Inertial Measurement Unit is one of the most valuable sensors (IMU). When we go about our normal lives, they can detect very low-level inertial data. This study was motivated by the idea that how we do these tasks reflects our identity. It has previously been shown that humans do this when conducting locomotive tasks like as walking, but the question was if we have a

Corresponding author: H. A. Imran (e-mail:himran.mscs18seecs@seecs.edu.pk)

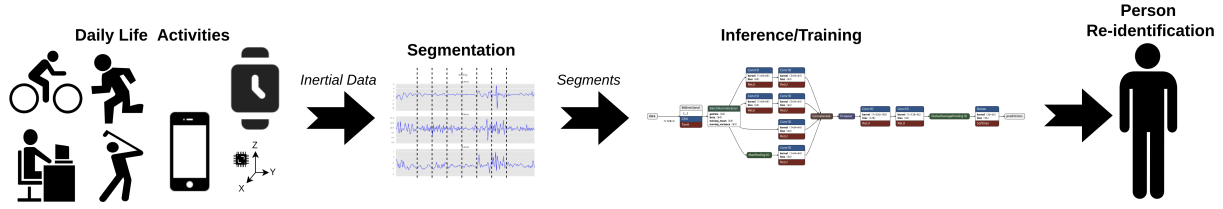


Fig. 1: End-to-end pipeline of the entire project



Fig. 2: The performance report for model classification

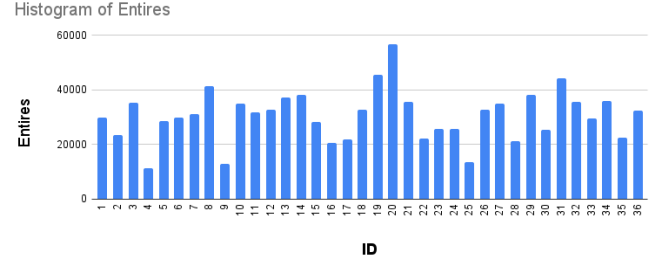


Fig. 3: Data Distribution according to each individual

distinctive style of completing any sort of everyday activity such as eating, running, dancing, and so on. The overall end-to-end pipeline of the project is depicted in 1. The input data was the accelerometer data of the smartphone while performing 6 different activities.

The input signal must be segmented since deep neural networks have a fixed input size. The size of the segmentation window is also significant; we used empirical research to determine that a size of 128 with a step size of 32 functioned best. The selected segmented were used for training or inference. There was no preprocessing or feature engineering was done. The motivation was to keep the entire solution entirely a deep learning one.

A. Dataset Used

We have used a dataset that was originally collected with the motive of human activity recognition using a smartphone's accelerometer [8]. The dataset is commonly known as WISDM Activity Prediction or WISDM 2011 dataset and has data on 6 different activities. The dataset has It had a unique ID for each volunteer. We have used this ID as a class and used the inertial data for different activities for training and inference by a deep learning model. The dataset has data of 36 volunteers. Figure 3 shows the number of dataset samples for each individual. It can be seen from the data distribution that the dataset is not balanced. Therefore we have reported Recall, Precision and F1-score along with Classification Accuracy.

B. Model's Motivation

The presented model is a Bi-directional Gated Recurrent Unit - Convolutional Neural Network kind of deep learning model. The motivation was that the GRU part of the model will extract the feature set which will be representing the relation between two segments of the time-series dataset. Then that feature set will be used by the CNN part of the model to further enhance the feature set before applying the softmax or classification layer. The model dose not have any fully-connected or dense layers, in fact we have used 1x1 convolutions at the ends of the model. Before applying the software,

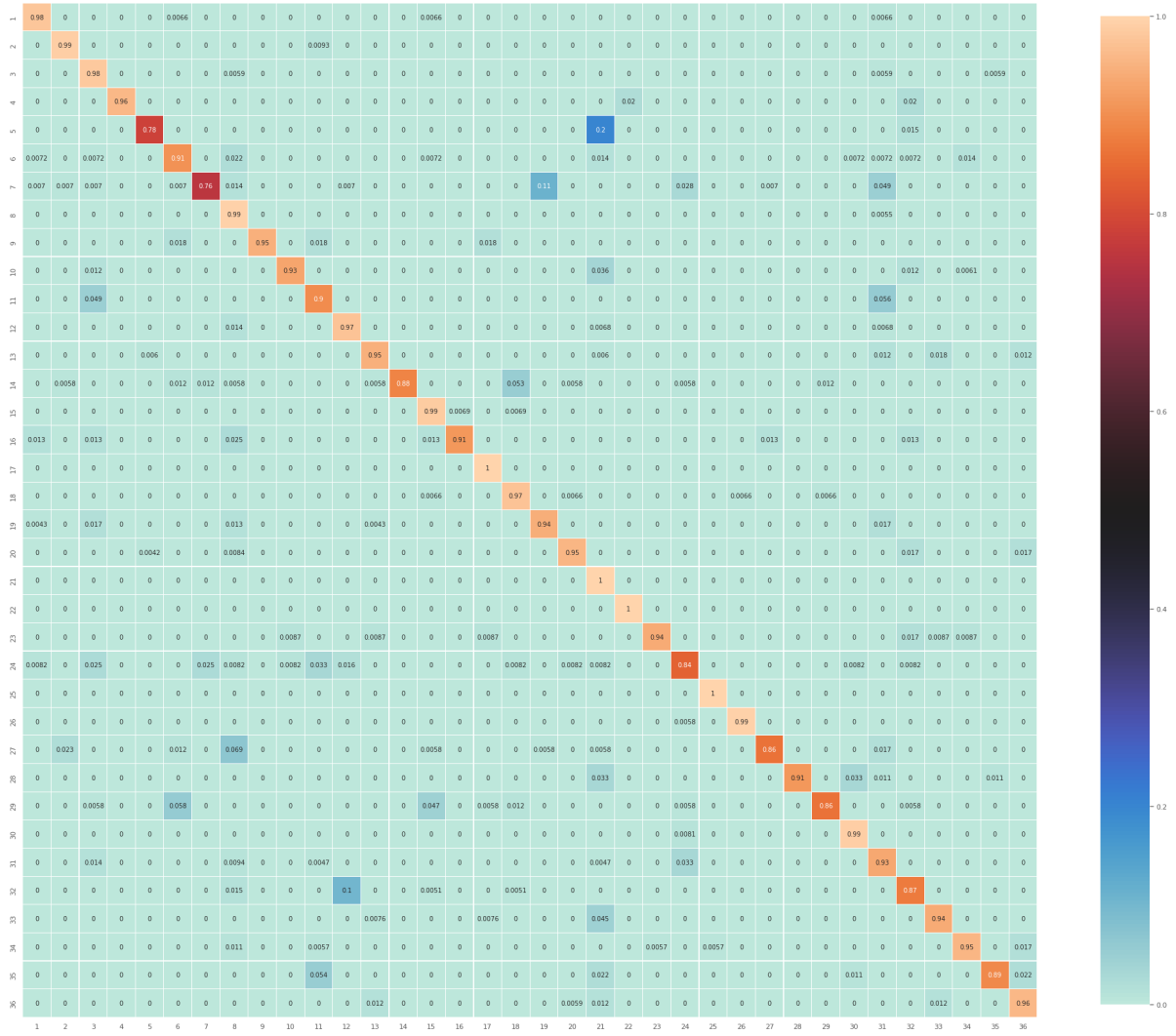


Fig. 4: Confusion matrix

the number of the last 1x1 convolutional filters used was set to 36 which is the number of classes. The idea was taken from [1].

The presented model has a total parameters of 95,768 out of which trainable are 95,640 and non-trainable are 128. The input feature map of (128,3) is passed to the Bidirectional-GRU layer which has 32 units. Batch Normalization is applied to the generated feature map. After that 64, 1x3 and 1x5 kernels are applied. Before that 64, 1x1 kernels were applied. Along with these, 64 separate 1x1 kernels are also applied. Max-pooling is also done and after max-pool 64, 1x1 kernels are further applied. The feature maps generated by all of these are concatenated. A dropout of 0.5 was applied to these concatenated features and than 128, 1x1 kernels were applied. After those 36 kernels were applied. 36 is the actual number of classes we have in our dataset. Than Global Average Pooling was applied before applied the softmax.

III. RESULTS

The epoch number used was 40 with batch size of 128. These figures were selected based on empirical analysis. Figure 5 shows the accuracy plot. 'Adam' optimizer was used. The achieved cross-validation accuracy was 93.84% and the test accuracy was 93.35%.

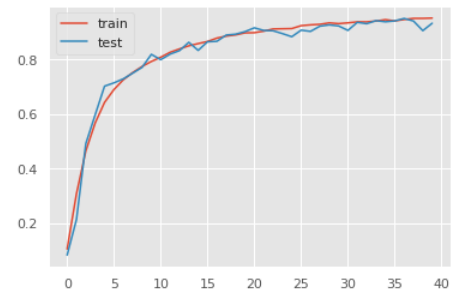


Fig. 5: Training and testing accuracy plot

Figure 2 shows the entire performance report. It show the average F1 score of 93%, the average precision is 94% and recall is 93%. The dataset is not balanced therefore, we have report F1, Recall and precision as well.

IV. CONCLUSIONS

We attempted to re-identify the individual in this article by analysing inertial data collected while undertaking non-locomotive and locomotive everyday life activities. To the best of our knowledge,

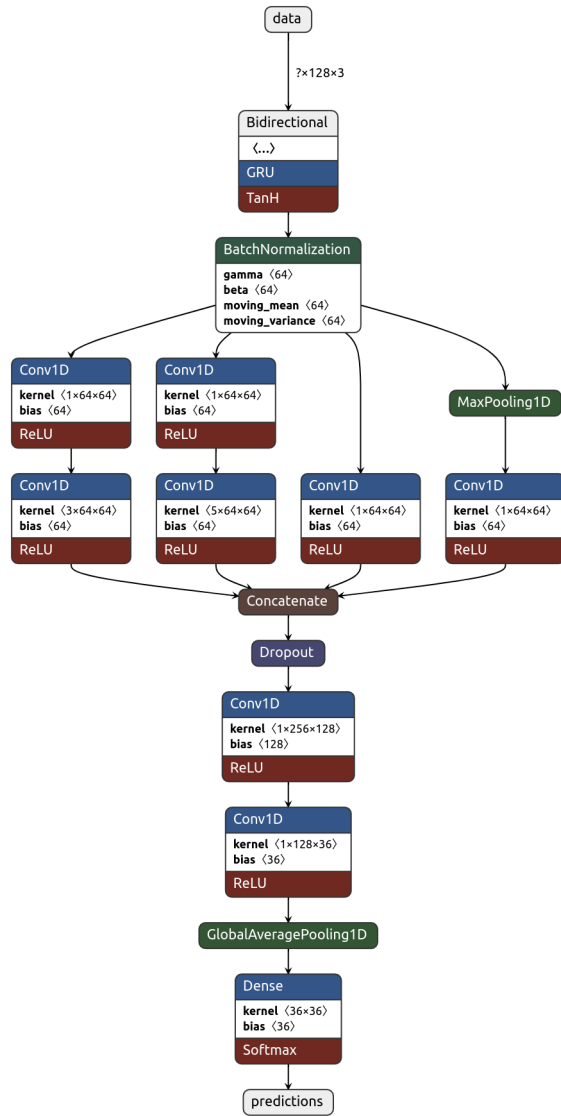


Fig. 6: Shanakht-Net: The presented model which is based on k mode

it is one of the first attempts to address the re-identification problem using both sorts of activities. Methodologies based on gait (locomotive activity) already exist in the literature. Using the presented deep neural network, it has a classification accuracy of 93.84%.

A. Future Experimentation and Directions

Following are the future exploration required:

- Train the model separately for locomotion and non-locomotion activities to determine how much performance is gained by non-locomotion activities.
- Train the model on more datasets to reach comprehensive conclusions. WISDM 2019 [9] is another suitable dataset.

REFERENCES

- [1] H. A. Imran, "Ultanet: A antithesis neural network for recognizing human activity using inertial sensors signals," *IEEE Sensors Letters*, pp. 1–1, 2022.

- [2] —, "Khail-net: A shallow convolutional neural network for recognizing sports activities using wearable inertial sensors," *IEEE Sensors Letters*, 2022.
- [3] X. Chen, S. Jiang, and B. Lo, "Subject-independent slow fall detection with wearable sensors via deep learning," in *2020 IEEE Sensors*. IEEE, 2020, pp. 1–4.
- [4] M. A. Hashmi, Q. Riaz, M. Zeeshan, M. Shahzad, and M. M. Fraz, "Motion reveal emotions: Identifying emotions from human walk using chest mounted smartphone," *IEEE Sensors Journal*, vol. 20, no. 22, pp. 13 511–13 522, 2020.
- [5] Q. Riaz, A. Vögele, B. Krüger, and A. Weber, "One small step for a man: Estimation of gender, age and height from recordings of one step by a single inertial sensor," *Sensors*, vol. 15, no. 12, pp. 31 999–32 019, 2015.
- [6] I. Gohar, Q. Riaz, M. Shahzad, M. Zeeshan Ul Hasnain Hashmi, H. Tahir, and M. Ehsan Ul Haq, "Person re-identification using deep modeling of temporally correlated inertial motion patterns," *Sensors*, vol. 20, no. 3, p. 949, 2020.
- [7] T. T. Ngo, Y. Makihara, H. Nagahara, Y. Mukaigawa, and Y. Yagi, "The largest inertial sensor-based gait database and performance evaluation of gait-based personal authentication," *Pattern Recognition*, vol. 47, no. 1, pp. 228–237, 2014.
- [8] 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 *Proceedings of the Fifth International Workshop on Knowledge Discovery from Sensor Data*, 2011, pp. 25–33.
- [9] G. M. Weiss, K. Yoneda, and T. Hayajneh, "Smartphone and smartwatch-based biometrics using activities of daily living," *IEEE Access*, vol. 7, pp. 133 190–133 202, 2019.