Time-Series Deep-Learning Classifier for Human Activity Recognition Based On Smartphone Built-in Sensors

Ghada Qanbar Ali, Dr. Hilal Al-Libawy

Department of Electrical Engineering, University of Babylon, Iraq

Abstract. Human Activity Recognition (HAR) is gaining more interest in recent years due to its growing role in many human-related sectors such as the health sector especially with elderly people and motion restricted patients. In recent years, there has been great progress in identifying human activity using various machine learning approaches. However, traditional methods of feature extraction are the most challenging in the feature selection process. Deep learning is a promising approach in the human activity recognition research area and has overcome the feature selection problem. However, several challenges are still open to research issues such as classification performance. This paper describes how to identify specific types of human physical activities using the accelerator and gyroscope data generated by the smartphone user. A deep convolutional neural network architecture has been proposed to perform HAR efficiently and effectively the system has been trained and tested over a dataset generated with the aid of 50 volunteers with four activities (walking, running, walking up-down stairs finally sitting-standing on the chair) events in real-world conditions. We chose four classes, each of which performs well, get to know our range of activities achieving 99% for validation and 99.8% for testing overall accuracy

Keywords. Human Activity Recognition (HAR); Deep Learning (DL); Convolutional Neural Network (CNN); continuous wavelet and smartphone sensors

1. Introduction

Knowledge of human activities is an important field of research in the omnipresent computer sciences and analysis of human behavior and computer interaction. In particular, learning daily activities are crucial to maintaining healthy lifestyles, patient rehabilitation, and elderly activity, which helps detect and diagnose critical conditions [1] HAR algorithms still face many challenges, including many applications :

- the diversity and complexity of daily activities
- variation of the same activity intra-subject and cross-subject
- the trade-off between specificity and performance
Mobile and computational expertise
- data hard to explain. [2]

Therefore, a mechanism for discovering postural and motor activities and body movements is provided by the framework for recognized human activities. Previous studies can be classified according to various devices, sensor patterns and data used to detect information on the activity in a large way. HAR can be classified into two main types [3]: video and sensor-based on HAR, the first type analyzes images or videos that contain human movements from the camera which are among the surrounding sensors, are environmental sensors or video cameras placed in the environment at specific points [4] while sensor-based HAR is focused on smart sensor motion data are integrated sensors that are integrated with clothing or other medical devices or integrated into personal devices such as smartphones or smartwatches such as accelerometer, gyroscope, Bluetooth, sound sensors, etc.

The massive development of smartphones integrated with multiple devices - systems Easy sensor for human physiologic signal collection researchers to monitor daily life activity, motion sensors (accelerometers, gyroscopic and magnetometers) provide important information that makes it easier to recognize and monitor the movement of users like walking, standing or running. Deep learning methods provide high-resolution results for large activities data systems HAR research has seen the development of profound deep learning methods (DL). Classic machine learning (CML) models in many HAR applications may be more appropriate due to the smaller size of the data set, fewer input data dimensions, availability, and expertise in the problem of the formulation [8]. The research [9] focused on the statement of published papers and the average accuracy of activity recognition in the past years from 2015 to 2019 for deep learning and classic machine learning models, our main focus in this paper is on HAR based on sensors (accelerometer and gyroscope) then data processing and analysis using a convolutional neural network. Which extracts the time series' local dependence and properties of fixed measures. The results showed that the CNN-based approach outperforms the new proposed technological approaches by extracting these properties

2. Related work

Many types of research have been proposed in the field of human activity despite the difference in sensors, the number of activities, and volunteers. the Accelerometer And GPS have been suggested [10] with four activities walking, standing, sitting, and Jogging using machine learning, each with a total accuracy of 95%. Table 1 [11] Comparison between state - of- artworks. A new independent procedure was proposed to define Parameters Adjustment Corresponding to the Position of the smartphone (PACP) [12]. Some studies use steady-position smartphones, like pockets of trousers, during recognition, which limits the behavior of users. Human activity using smartphones at various locations was taking raw data from sensors such as accelerometer and gyroscope data to recognize the position of the smartphone first and the accelerometer corresponding to the position was then changed; the activities were finally acknowledged by a Support Vector Machine model (SVM) trained. Experimental findings show that the PACP is more efficient than previous methods with 91% accuracy. Table 3 in the research [3] indicated that there are some amount of HAR datasets that most researchers used. Some of the common ones are summarized available to the public for different research purposes. In addition to accelerometer and gyroscope data, the convolutional neural networking (ConvNet) was finally used in work to identify gestures which found that CNN is more efficient than other techniques, such as Dynamic Time Warping (DTW) method for gestures. [13] Both[14][15] ConvNets were implemented in HAR with sensor signals, but the former generally evaluated the problem with time series, while the latter only used a single-layer network, which ignores the high potential advantage of hierarchical feature extraction. Table (1) illustrates some state of the art of this area However, Yang et al. Convnets to HAR have been implemented in a hierarchical model to ensure superiority in many standard problems[16] in [17] CNN conducted time-series image data analyzes using wavelet transformation and then image analysis using CNN data set. Their analyzes were conducted using LeNet 5, a typical CNN algorithm. For precision of the classification,
the results showed that the original data set did not differ from the data set used for the transformation of the wavelength. However, the analysis can be completed with wavelet transformation compared to the original dataset in a very short time. We can extract features directly from the input data by using deep learning like CNNs to create more general learning methods. Current methods, however, do this usually with additional classification layers and nodes that add to the computational complexity. For instance, in the described CNN method [18], further layers of pooling are applied when the characteristics from the first input are detected to produce variable-size features. To incorporate multiple channel features, 1024 hidden layers of neurons are then introduced and an additional softmax layer is employed in generating classification results. Yang et al [16], Chen, and Xue [19] proposed CNN's with multiple warps and reduction layer iterations or warp and stack layers used to extract features.

Table 1. list of research papers using a smartphone to identify recognition of human activities

| References and dataset | Sensor type | Activities | Subjects | Accuracy |
|------------------------|-------------|------------|----------|----------|
| [20] UCI smartphone (2015) | accelerometer and gyroscope (smartphone) | upstairs walking, walking downstairs, Sitting, Standing walking, and laying (6) | 30 | 94.79% |
| [21] UCI-HAR (2013) | accelerometer and gyroscope (Samsung Galaxy S II) | walking, upstairs, downstairs, sitting standing and laying (6) | 30 | 96% |
| [22] Their own dataset (2016) | accelerometer, gyroscope (smartphone) | upstairs walking, walking downstairs, walking, sitting, standing and laying (6) | 30 | 94.79% |
| [23] Their own dataset (2018) | accelerometer, gyroscope (smartphone) | walking, sitting, standing, and jogging (4) | 2 | 99.55% |
| [24] UCI-HAR (2018) | accelerometer, gyroscope (smartphone) | climbing up the stairs, climbing down the stairs, walking, sitting, standing, and laying (6) | 9 | 99.4% |
3. Convolution Neural Network (CNN)

For many years, automatic learning methods are heavily based on handmade features ranges from basic statistics to public techniques to reduce dimensions into different scales. Although the methods of generating features have been successfully used in many places, usually include an arduous design phase, it requires the engineer to create knowledge about types of patterns in data taken by relevant information. This has led to specialized educational systems that are only able to fill narrow and unable roles, but important, deep learning tends to overcome these restrictions. Instead of designing the deep learning features automatically via the network. Also, the profound neural network can extract deep class high-level representation, making them more convenient to learn about complex tasks. A deep neural artificial (ConvNets or CNN) network usually has several lays of the neural network, each having more than one layer of neurons. CNN is specialized in the extraction of signal features and has demonstrated promising image classification, voice recognition, and text analysis outcomes far superior to other classifiers [3]. The results can be further developed exactly by increasing numbering convolution layers and hidden neurons [25] many important network layers have different roles in CNN’s like convolutional, pooling, and fully connected layers. We use time-series data in experiments in this paper the convolutional layer in extracting data and in extracting functions from subsequent convolutions is more complex and efficient. The number of features can be reduced by using a pooling layer and the maximum pooling is used in the experiment. The intermediate layer of the neural network acts as each convolutional layer and single pooling layer. Before the output, the layer is placed in the final fully-connected layer (Fc), which collects the previous layer results for the score calculation for each last class. The role of the output layer produces classification outcomes based on the output of fully connected layers. [26]. A performance for convolution is calculated as:

\[
C_i^{l,j} = \sigma(b_j^l + \sum_{m=1}^{M} W_{m}^{l,j} x_{i+m-1}^{l-1,j})
\]

Where \(\sigma\) function of activation, the term \(b\) pointing to feature map bias, the symbol \(l\) indicates an index of layer, \(M\) is the size of kernel/filter, and \(w\) refers to feature map weight. The overall structure of CNN’s shows in figure (1).

![Convolutional network structure (CNN)](image)

Figure 1. Convolutional network structure (CNN)
• Input Layer: the raw input data are stored in the input layer. It's a 3D input of width and height of the image with depth of the color channels, usually 3 for RGB.
• Convolutional layer: the basis of the CNN is the layer of convolution. Uses smaller filters or kernels than the image input. The wrapping is performed with a portion of the input and the core, which finally covers the whole input of the image. A feature map or activation map in this process is the output; the activation map is stacked to produce a 3D tensor. As filters are trained, they will be able to recognize edges and patterns more deeply on the network. The function ReLU is widely employed CNNs. All negative inputs are taken from this layer, set to zero. The ReLU layer does not contain hyper parameters, that is, the designer's parameters specified.
• Batch Normalization: it is common to normalize data before data is entered into NN. Batch normalizes the activation of the layer's average output to close to 0. The theory is typically to accelerate the process of CNN training.
• Pooling Layer: a pooling layer is always accompanied by a convolutional layer; max-pooling is popular. Since its output pooling layer reduces data size, its pooling layer takes as many small features as possible.
• Fully Connected Layer: after several convolutional and max-pooling layers, CNN has one or more fully connected layers on top of it that perform the classification. For different classes, the classifier outputs the probabilities.
• Soft-max layer: the final output is finally transferred to the soft-max layer for the calculation of the probability distribution over the predicted classes.

4. Continuous wavelet

A wave is a perturbation of periodic oscillations that propagates across space and time, usually as energy is transmitted. The word wavelet means "small wave" and is given in the early 1980s by Morlet and Grossmann. Or little because the wavelet functions have a finite length. When time localization of the spectral components is required, a transformation is required that gives the temporal frequency representation of the signal. Wavelet transformations have advantages over traditional Fourier transforms to represent functions with sharp discontinuities and peaks, and to precisely deconstruct and reconstruct finite, Signals which are non-periodic and/or stationary. In the last decade, several solutions were developed to solve this problem and better represent a signal in the domain of time and frequency. This deficiency can be overcome by wavelet function representation, wavelet transfer. The transmission of the wavelet at high frequencies provides good time resolution and poor frequency precision while at low frequencies; the transformation of the wave-length results in good frequency precision and poor time resolution. Therefore, for signal analysis, we can gather all information, wavelet transfers are a mathematical approach [27] widely used for signal processing applications. It can decompose special patterns hidden in the data block. Regarding the issue of prediction by time series and neural networks, the CNN route has been proposed along with the use of WT to extract frequency features. The proposed model is trained with the proportion of the data; Moreover, it solves the problem of overfitting. The wavelet transform can display functions simultaneously and demonstrate their local properties in the time-frequency domain. Therefore, it is more effective to replace the Short-Term Fourier Transform (STFT) with a Wavelet Transform (WT), which is a more powerful tool for extracting features in the frequency domain. Replacing STFT with WT addresses the non-stationary nature of a signal. Therefore, to solve the resolution problem and extract accurate frequency information in this paper, we used Continuous Waveform Transfer (CWT). The CWT is defined in Equation (2) as follows:

\[
C(s,t) = \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} f(t) \psi^\prime \left( \frac{t - \tau}{s} \right) dt
\]
Where \((t)\) denotes the mother wavelet, \(s\) pointing to scale, is the translation, and \(*\) represents the operation of complex conjugation. The CWT results are many C-wavelet coefficients that are scale and position functions. Therefore in this study, use Wavelet Transform; we have achieved much better accuracy in extracting time-related frequency information. Due to Heisenberg's Uncertainty Principle, one cannot know the exact frequency at a particular point in time but we can know the range of frequencies present in a specific period. We explain this with some examples when converting time series to images using CWT and as shown in figure (2), the use of CWT helped us to increase the accuracy of our system.

![Walking Gyroscope](image1.png) ![Image for Walking Gyroscope](image2.png)

**Figure 2.** Example of converting time series to 2D images: (a) Walking Gyroscope, (b) Image for Walking Gyroscope

5. **Proposed human activity recognition (HAR) algorithm**

The framework for defining typical activity follows mostly stages including as shown in figure (3): data collection, data organization, and classification. Multiple sensors collect data regarding human use and behavior. Server-side classification techniques are applied to capture activity. Wearable sensors and smartphone devices are used to collect data. The choice of sensors and measuring characteristics that play an important part in the performance of an activity recognition system is one of the most important issues in data collection. If sensors were not identified, the performance of recognition could have an adverse effect. For activity recognition, inertial sensors such as accelerometers and gyroscopes are used. The accelerometer measures an intelligent device's non-gravity speed whilst the gyroscope detects a change in direction or angular velocity.
5.1. **Experiment setup and data collection**

Data information was collected from the accelerometer in the smartphone and the gyroscope sensors, where 50 volunteers were in the age group from (19 to 49 years) (number of males 31 and number of females 19) with weights ranging between (53 - 90) the range of physical activities (running, walking, walking up-down stairs, standing-sitting on chairs). Table(2) provides an overview of the data collection, and the sensor time-series information from the phone and gyroscope accelerometer is shown as described in figure (4), in addition to the activity recognition models as it was previously mentioned that acceleration and gyroscope sensors obtained from downloading the Kinetic Sensor pro application[28] were used on the smartphone, and then the data was collected by the volunteers by
holding the smartphone in the hand in the position that the person used to hold his phone or in the way he/she feels comfortable.

Table 2. Summary of data collection

| Number of subjects | 50 |
|--------------------|----|
| Number of male/female | 31 male/19 female |
| Age/weight/length | 19-49 years/53-90 kg /140-190 |
| Number of activity | 4(running, walking, walking up-down stairs, standing-sitting on chairs) |
| Sample rate | 30 Hz |
| smartphone | Huawei Mate 20,model/HMA-L29, Android version/9, screen resolution/1080*2244 |
| sensor | Accelerometer and Gyroscope |
| laptop | Processor/ Intel(R) Core(TM) i7-3740QM CPU @ 2.70GHz 2.70 GHz, RAM/ 16 GB |

Figure 4. (a) Accelerometer, (b) Gyroscope screenshot of smartphone application [28] (walking activity for example)
5.2. Data organization

The raw time-series sensor data is recorded by the accelerometer and gyroscope of the smartphone at a rate of 500 Hz, each consisting of three axes x, y, z. The data collected in a format CSV was transferred from the smartphone to the laptop, with the specialization of sharing in files, sharing via private e-mail, and downloading files using Google spreadsheet finally, download the file in Excel format to the laptop in an Excel format. Sensor data is stored in separate subdirectories. Each volunteer is arranged with its file, and the accelerometer and gyroscope data for each device are stored in two subdirectories. Within each sub-directory there is a file for each topic and all volunteers were collected in one file after each volunteer was numbered with his numbering, and the accelerometer and gyroscope files were collected in one file according to the volunteers’ order, so there was a file with 400 Excel files containing all the activities (walking, running, walking up-down the stairs and standing-sitting on the chair for 50 volunteers).

5.3. Converting time series to 2d images using CWT

The initial sensory data may contain noises or abnormal heights on the smartphone due to unintended change in the sensor direction, a certain change in position or a device drop, noise from human motion, etc. Filtering systems remove signal noise from the accelerometer and preserve the medium frequency signal components. An important process for validating and normalizing the filtered database is data transformation. The square root of the value is another nonlinear process used for statistical analysis. Signal Vector Magnitude (SVMag) is applied to convert data from three-dimension (the accelerometer and gyroscope sensors consist of three-dimension( Ax, Ay, Az) ( Gx, Gy, Gz) respectively) into one-dimension shown in figure(4), this facilitates the process of converting the Time-Series a chronologically indexed series of data points (listed or graphed). Especially frequently, a time series is a sequence that is captured at consecutive equidistant time points to images. applying the following equation [29]

\[ \text{SVMag for Accelerometer} = \sqrt{A_x^2 + A_y^2 + A_z^2} \]  

\[ \text{SVMag for Gyroscope} = \sqrt{G_x^2 + G_y^2 + G_z^2} \]

As shown in figure (5)
(a) All activity for Accelerometer (three-dimension)

(b) All activity for Gyroscope (three-dimension)
Figure 5. (a) All activity for Accelerometer (three-dimension); (b) All activity for Gyroscope (three-dimension); (c) All activity for Accelerometer (one-dimension); (d) All activity for a Gyroscope (one-dimension)
5.4. CNN architecture

Data Preparation Conventional algorithms particularly deep learning models on neural networks of convolution (CNNs) can be taught to "learn" patterns and be adapted to a wide variety of problems and tasks, such as denoising super-resolution, and segmentation. However, whether deep or not model training relies heavily on data. As shown in the following flowchart (figure (6))

![CNN workflow diagram](image)

- **Original data**
  - (Time series X, Y, Z)

- **Data preparation**
  - (Number of image=4000
    Image size= 227*227*3)

- **Training sets**
  - (Ratio=0.6)

- **Validation sets**
  - (Ratio=0.1)

- **Testing sets**
  - (Ratio=0.3)

- **Train the model**
  - (Number of layers=23
    Learning rate=0.01
    Mini batch size=15
    Validation frequency=30
    Max epochs=4)

- **Evaluate the model**

- **Predictive model**
  - (Accuracy=99% for validation and 99.8% for testing)

- **Find tune the model**

*Figure 6. CNN workflow*
three different data sets the training set (0.6), the validation set (0.1), and the testing set (0.3) is used for training, fine-tuning a model, and testing. In this paper, CNN consists of one layer of input, one layer of output and five convolution layers (c1, c2, c3, c4, and c5), using (ReLU) Rectified linear Unit as an activation function which in its input arguments looks for positivity. If the input is positive, the value is returned, if not, zero is returned as the final output value, where zero is less than zero (0).

ReLU is the following expression:

$$y = x^+ = \max(0, x)$$  \hspace{1cm} (5)

Four layers (p1, p2, p3, and p4) for pooling and one layer with full connection (Fc). Data and pre-process data are received by the input layer. The CNN number is adjusted (23 layers), the numbering of convolution layers, the filter size, and the pooling size (in the same order).

6. Experimental Results

From 1,200 test cases, after 4 reigns. The results are very good for a simple model with less CPU training and less time-consuming training (approx. 10 min). While some images are difficult to define, our model can correctly classify them. The test accuracy of 99.8% indicates that the model is well trained for prediction. As the amount of data increases, the size of the workout affects increasing accuracy. The more data in the training set, the less the effect of training errors and test errors, the better the accuracy. When evaluating the activity performance recognition system we can achieve high levels of accuracy. Therefore, if correctly identified, the activity can be classified as true positional or true negative (TP), false positive (FP), or false negative (FN) when classified incorrectly. Table 3 shows the evaluation metrics for the classifier, other performance measures are derived from true positives or true negatives. As evident in figure (7) calculation of the confusion matrix.

(a)Confusion matrix for validation \hspace{1cm} (b) Confusion matrix for testing

Figure 7. (a) Confusion matrix for validation, (b) Confusion matrix for testing
Table 3. Evaluation metrics of proposed algorithm: (A-Running, B-Standing-Sitting, C-Walking, D-Walking up-down stairs)

| Measure               | Formula                                  | A    | B    | C       | D       |
|-----------------------|------------------------------------------|------|------|---------|---------|
| sensitivity, recall   | $\frac{TP}{TP + FN}$                     | 100% | 100% | 99.7%   | 99.7%   |
| precision             | $\frac{TP}{TP + FP}$                     | 99.7%| 100% | 99.7%   | 100%    |
| F-score               | $\frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$ | 99.8%| 100% | 99.7%   | 99.8%   |
| recognition rate or accuracy | $\frac{TP + TN}{TP + TN + FP + FN}$ |      |      |         | 99.8%   |

Finally, the accuracy of the comprehensive recognition of the view is compared with some related work as shown in table 4.

Table 4. Comparison of our method with some state-of-art studies

| Method                          | Subjects | Overall Accuracy % |
|---------------------------------|----------|--------------------|
| [21] UCI smartphone in (2015)   | 30       | 94.79              |
| [22] UCI-HAR in (2013)          | 30       | 96                 |
| [23] Their own method in (2016)| 30       | 94.79              |
| [24] Their own method in (2018)| 2        | 99.55              |
| [25] UCI-HAR in (2018)          | 9        | 99.4               |
| Our Proposal method             | 50       | 99.8               |
In Table 4, we also noted that our results in comparison with other research are the best since experience has been applied in real-world conditions, efforts have to be made to make the system acceptable to many users, because the determination of human activity is very close to the everyday life of a human, and therefore the system has to be useful to enable it to respond immediately to mobile devices. The issue of calculation costs should therefore be dealt with. We, therefore, see that in this paper the proposed method exceeds all of the methods of comprehensive accuracy of recognition.

7. Conclusion

Understanding the topic of research in pattern recognition is important for human activity. In this paper, we explore recent advances in the field of deep learning to recognize the sensor-based activity. In comparison with traditional models, deep learning reduces dependence on the extraction of human-made features and improves performance by automatically acquiring higher-level sensor data representations. In three key categories, we highlight recent advances: sensor method, deep model, and application. Accordingly, this study aimed to demonstrate the basic theory of concepts and the potential to benefit from the application of high CNN networks in the classification of images. Time-series signals are first converted into images using Wavelet and then processed by the CNN model. Moreover, we need further study to automatically analyze the extracted features. Although the dominant technique for HAR may be deep neural networks, further study of the method properties and the use of a larger data set should be conducted.

References

[1] Nweke, H. F., Teh, Y. W., Al-garadi Mohammed Ali, & Alo, U. R. Deep learning algorithms for human activity recognition using mobile and wearable sensor networks: State of the art and research challenges. Expert Systems with Applications, 105(2018), 23

[2] Lara, Oscar D., and Miguel A. Labrador. "A survey on human activity recognition using wearable sensors." IEEE communications surveys & tutorials 15.3 (2012): 1192-1209.

[3] Wang, Jindong, et al. "Deep learning for sensor-based activity recognition: A survey." Pattern Recognition Letters 119 (2019): 3-11.

[4] Liu, Ye, et al. "From action to activity: sensor-based activity recognition." Neurocomputing 181 (2016): 108-115.

[5] Kanjo, Eiman, Eman MG Younis, and Chee Siang Ang. "Deep learning analysis of mobile physiological, environmental and location sensor data for emotion detection." Information Fusion 49 (2019): 46-56.

[6] Liu, L., Peng, Y., Liu, M., and Huang, Z. Sensor-based human activity recognition system with a multilayered model using time series shapelets. Knowledge-Based Systems 90 (2015), 138–152.

[7] Neverova, N., Wolf, C., Lacey, G., Fridman, L., Chandra, D., Barbello, B., and Taylor, G. Learning human identity from motion patterns. IEEE Access 4 (2016), 1810–1820.

[8] LeCun, Y., Bengio, Y., and Hinton, G. Deep learning. Nature 521, 7553 (2015), 436–444.

[9] Demrozi, Florenc, et al. "Human activity recognition using inertial, physiological and environmental sensors: a comprehensive survey." IEEE Access (2020)
[10] Lee, Kangjae, and Mei-Po Kwan. "Physical activity classification in free-living conditions using smartphone accelerometer data and exploration of predicted results." Computers, Environment and Urban Systems 67 (2018): 124-131.

[11] Saha, Jayita, et al. "An ensemble of condition-based classifiers for device-independent detailed human activity recognition using smartphones." Information 9.4 (2018): 94.

[12] Yang, Rong, and Baowei Wang. "PACP: a position-independent activity recognition method using smartphone sensors." Information 7.4 (2016): 72.

[13] Duffner, Stefan, et al. "3D gesture classification with convolutional neural networks." 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2014.

[14] Zeng, Ming, et al. "Convolutional neural networks for human activity recognition using mobile sensors." 6th International Conference on Mobile Computing, Applications and Services. IEEE, 2014.

[15] Zheng, Yi, et al. "Time series classification using multi-channels deep convolutional neural networks." International conference on web-age information management. Springer, Cham, 2014.

[16] Yang, Jianbo, et al. "Deep convolutional neural networks on multichannel time series for human activity recognition." Ijcai. Vol. 15. 2015.

[17] Levinskas, A. "Convolutional neural network feature reduction using wavelet transform." Elektronika ir Elektrotechnika 19.3 (2013): 61-64.

[18] Zeng, Ming, et al. "Convolutional neural networks for human activity recognition using mobile sensors." 6th International Conference on Mobile Computing, Applications and Services. IEEE, 2014.

[19] Chen, Yuqing, and Yang Xue. "A deep learning approach to human activity recognition based on single accelerometer." 2015 IEEE international conference on systems, man, and cybernetics. IEEE, 2015.

[20] Ronao, Charissa Ann, and Sung-Bae Cho. "Deep convolutional neural networks for human activity recognition with smartphone sensors." International Conference on Neural Information Processing. Springer, Cham, 2015.

[21] Anguita, Davide, et al. "A public domain dataset for human activity recognition using smartphones." Esann. Vol. 3. 2013.

[22] Ronao, Charissa Ann, and Sung-Bae Cho. "Human activity recognition with smartphone sensors using deep learning neural networks." Expert systems with applications 59 (2016): 235-244.

[23] Lee, Kangjae, and Mei-Po Kwan. "Physical activity classification in free-living conditions using smartphone accelerometer data and exploration of predicted results." Computers, Environment and Urban Systems 67 (2018): 124-131.

[24] Bulbul, Erhan, Aydin Cetin, and Ibrahim Alper Dogru. "Human activity recognition using smartphones." 2018 2nd international symposium on multidisciplinary studies and innovative technologies (ismsit). IEEE, 2018.

[25] Hossain, Md Anwar, and Md Shahriar Alam Sajib. "Classification of image using convolutional neural network (CNN)." Global Journal of Computer Science and Technology (2019).
[26] Wan, Shaohua, et al. "Deep learning models for real-time human activity recognition with smartphones." Mobile Networks and Applications 25.2 (2020): 743-755.

[27] Furuya, Tomohiro, Hiromi Kondo, and Fumiyo N. Kondo. "Convolutional Neural Networks on Time Series for Smartphone Application Activations Using Wavelet Transform." 2019 8th International Congress on Advanced Applied Informatics (IIAI-AAI). IEEE, 2019.

[28] Sensor kinetics pro. Available online: https://play.google.com/store/apps/details?id=com.innoventions.sensorkineticspro&hl=ar&gl=US

[29] Saha, Jayita, et al. "An ensemble of condition based classifiers for device independent detailed human activity recognition using smartphones." Information 9.4 (2018): 94.