

Smart Devices Based Multisensory Approach for Complex Human Activity Recognition

Muhammad Atif Hanif¹, Tallha Akram¹, Aamir Shahzad², Muhammad Attique Khan³, Usman Tariq⁴, Jung-In Choi⁵, Yunyoung Nam^{6,*} and Zanib Zulfiqar⁷

¹Department of Electrical and Computer Engineering, COMSATS University Islamabad, Wah Campus, 47040, Pakistan

²Department of Electrical and Computer Engineering, COMSATS University Islamabad, Abbottabad Campus, 22040, Pakistan

³Department of Computer Science, HITEC University Taxila, Taxila, Pakistan

⁴College of Computer Engineering and Sciences, Prince Sattam Bin Abdulaziz University, Al-Khraj, Saudi Arabia

⁵Department of Applied Artificial Intelligence, Ajou University, Suwon, Korea

⁶Department of Computer Science and Engineering, Soonchunhyang University, Asan, Korea

⁷Department of Electrical and Computer Engineering, CUI, Sahiwal, Pakistan

*Corresponding Author: Yunyoung Nam. Email: ynam@sch.ac.kr

Received: 26 April 2021; Accepted: 17 June 2021

Abstract: Sensors based Human Activity Recognition (HAR) have numerous applications in eHealth, sports, fitness assessments, ambient assisted living (AAL), human-computer interaction and many more. The human physical activity can be monitored by using wearable sensors or external devices. The usage of external devices has disadvantages in terms of cost, hardware installation, storage, computational time and lighting conditions dependencies. Therefore, most of the researchers used smart devices like smart phones, smart bands and watches which contain various sensors like accelerometer, gyroscope, GPS etc., and adequate processing capabilities. For the task of recognition, human activities can be broadly categorized as basic and complex human activities. Recognition of complex activities have received very less attention of researchers due to difficulty of problem by using either smart phones or smart watches. Other reasons include lack of sensor-based labeled dataset having several complex human daily life activities. Some of the researchers have worked on the smart phone's inertial sensors to perform human activity recognition, whereas a few of them used both pocket and wrist positions. In this research, we have proposed a novel framework which is capable to recognize both basic and complex human activities using built-in-sensors of smart phone and smart watch. We have considered 25 physical activities, including 20 complex ones, using smart device's built-in sensors. To the best of our knowledge, the existing literature consider only up to 15 activities of daily life.

Keywords: Complex human activities; human daily life activities; features extraction; data fusion; multi-sensory; smartwatch; smartphone



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1 Introduction

Sensor based Human Activity Recognition (HAR) is an emerging field of machine learning, having several advantages as compared to the vision based human activity recognition [1]. Vision based HAR is less preferable due to the installation of hardware setups, cost, data storage requirements and computational time [2,3]. The constraints due to variable light conditions also affects the system's compatibility which leads to limited portability [4]. On the contrary, wearable sensor-based human activity recognition systems do not contain the aforementioned problems. Initially, the wearable sensors used for human activity recognition were not portable due to their bulky structure, cost, and additional power setup. Also, these sensors were impractical round the clock due to resource constraints. These days, majority of people use smart phones, watches and fitness bands that contain various build-in energy efficient smart sensors [5]. Generally, the sensors like accelerometer, gyroscope, magnetometer, global positioning system (GPS), temperature sensor, proximity sensor, barometer and others are present in smart devices. Due to the rich features and easy availability of these sensors, the sensor based human activity recognition systems become more popular [6]. These sensors are used in various manners. In order to observe less repetitive activities (i.e., hand/arm movements) the smart watch sensors are more effective. Similarly, to observe more repetitive activities (i.e., overall human body movements) the smart phone in pocket gives better results. This implies that the sensor position also plays an important role in activity recognition. Recent studies showed that when the collective usage of smart phone and smart watch sensors produced better results [7]. The basic human activities (e.g., sitting, walking, running, stair up and down etc.) are recognized by using smart phone (pocket position). But the complex human activities (Eating, Smoking, and talking etc.) is easily recognizable using wrist wearable sensors as hand movements are involved in these activities. In this research, the complex activities are overlapped with the basic activities e.g., smoking while walking or sitting, similarly eating ice-cream while walking etc. These activities are recognized by using multiple sensors. The combination of wrist wearable sensors and the smart phone sensors efficiently recognize the type of activity being performed by a person. Reliable recognition of complex human activities gives a new direction of HAR applications, including tracking bad habits and providing coaching to individuals [8]. According to the World Health Organization (WHO) report, smoking, alcohol usage, lack of physical activity and poor nutrition are the main causes of early deaths. The main purpose of this research is to identify these bad activities by using multiple smart sensors. Furthermore, the recognition and tracking of these activities can be useful for awareness feedback or fitness assessment [9]. The major contribution is to incorporate the applications of both smart watches and smart phones to recognize those complex human activities which are less repetitive in nature efficiently. The multi-sensor data fusion enables accurate recognition performance for human activity [10].

The proposed sensor based human physical activity recognition is divided into four main steps. (1) data acquisition, (2) pre-processing, (3) feature extraction (4) classification of various activities. Firstly, we have formulated a dataset for more than 20 complex human physical activities using built-in sensors of smart phone and smart watch. The sensor data fusion is followed by the pre-processing stage which removes noise from raw data and divides it into windows or segments. Further the dataset is divided into test and train data. After this, feature extraction is performed by using the proposed feature extraction method. At the final stage, classification of various activities is done on the basis of feature extraction. The main contribution of this work is listed below:

- The effect of two devices i.e., wrist wearable device and pocket position smart phone is studied. The results from the data fusion from these devices are presented for better understanding of basic and complex human physical activities which is not possible by using single device sensors.
- A dataset for several overlapped human basic and complex physical activities is formulated. These activities are recognized by using the proposed methodology.
- A machine learning approach is presented that provide better performance for the recognition of more than 25 human physical activities.

2 Related Work

Human activity recognition is an important research area in computer vision [11,12]. Many techniques are introduced in the literature for action recognition using classical techniques [13,14] and deep learning based techniques [15]. The major application of activity recognition is video surveillance [16] and biometrics [17,18]. The sensor based human physical activity recognition using smartphone and smartwatch (wrist wearable devices) has been widely studied from last few years due to its various application in daily life specially in healthcare of elderly, disable or dependent peoples, sport coaching, fitness assessments, ambient assisted living (AAL), human-computer interaction, bad habits monitoring, exercise tracking, quality of life monitoring, entertainment etc. In previous studies mostly researcher use wearable motion sensors for few basic human activities recognition. When number of human physical activities increased, especially similar type (overlapped) of human physical activities, the requirement of additional sensors increased. Furthermore for several complex human activities recognition, it is also required that the sensors are placed at more than one position of human body. To overcome wearable sensor limitations, the researcher also explored the smartphone-built sensors. With the growing trends of smart phone and smartwatch usage, the complex human activities can be recognized using both devices built-in sensors and data fusion followed by efficient feature extraction and machine learning methods. In [19], the authors used smartphone accelerometer data to classify fast vs slow walking, aerobic dancing activities that were not studied previously. In this study, authors have used combination of classifiers (MLP, Logit Boost, SVM) to classify six human activities (slow walking, fast walking, running, stair-up, stair-down, dancing). An accuracy of 91.15% has been achieved. In [20], authors used on body accelerometer and recognized seven human activities (sitting, walking, stair up, stair down, standing, walking, lying, running, cycling) using hidden Markov model (HMMs). In [21], authors used smart phone and smart watch built-in sensors for recognition of nine human activities (standing, sitting, walking, running, stair up, stair down, cycling, elevator up and elevator down) using five different classifiers. However, in this study the smartphone and smart watch sensors data were used independently. In [22], the authors used two motion sensors, one at wrist and other at hip independently to recognized five human physical activities.. However, they also treat both sensors independently. They used logistic regression classifier and detected activities like standing, sitting, walking, lying, and running. In [23] a wrist worn motion sensor was used to detect eight human activities including additional activity that is working on computer. In [24] the authors recognized a complex activity that is eating, or drinking using wrist worn accelerometer and gyroscope. Authors used HMM for classifying the activities and achieve 84.3% accuracy. They detect eating activity by splitting into sub activities like eating, drinking, resting etc. In [25] the authors detect the same as in [24] eating activity using wrist worn accelerometer and gyro scope. In this work they identified non-eating and eating periods and they achieve 81% accuracy. In [26], similar work as in [25] was done where authors recognized eating and non-eating activities using a smartwatch that have built-in accelerometer and gyro scope achieving 70%

accuracy. In [27], the author used accelerometers at wrist and foot position and they recognized smoking activity independently with other activities like running, walking etc., reporting 70% accuracy. In [24] some complex activities including smoking and eating are detected by using three accelerometers and gyroscope sensors. In this study, authors have used naive bays, decision tree, k-nearest neighbor classifier to classify activities. The sensors were placed at pocket and wrists, which is a more practical approach rather than using foot or other body parts. The authors treated the complex activities separately with some basic activities for example smoking while running or standing, eating while talking etc. In real world scenario complex activities do overlap with some basic activities, the aspect ignored in this study. In [28], the authors proposed a framework for sensor based human physical activity recognition. They used a central server-based data collection from multiple devices like smartphone, smart watch etc. via bluetooth and transfer data using advanced messaging queuing protocol (AMQP). In this work authors used three different types of fusion of multiple features and classified using SVM early fusion, late fusion and dynamic fusion of features method. Fusion is an important research domain and many techniques are introduced in the literature [29,30]. In this study authors recognized human physical activities such as lying, walking, standing, sitting, bedding, getting up, lying down, standing up, standing down, putting hand back, stretching a hand. They reported 87.4% accuracy. Furthermore, they investigate effectiveness of smartphone and smartwatch sensor on different activities. In this paper, we have proposed a system that considers the fusion of both smart phone and smart watch sensors data. Moreover, we classify the overlapped basic and complex activities, based on statistical features analysis.

3 Proposed Methodology

Complex human activities overlapped with basic human activities leads towards the complexity of daily life activity recognition system. In order to overcome these complications, we have designed a mathematical model for multisensory learning and multilocational placement human activity recognition. The designed approach efficiently uses the data from smart phone and smart watch. This system provides advantages in terms of less computational time, reduction in installation cost, power efficiency, also the system is more appealing and contemporary than other conventional methods. The system effectively monitors the human activities for 24-h interval. The main objective of proposed system is to classify various complex human activities using built-in sensor of smartphone and smartwatch with better accuracy and less computational requirements. Fig. 1 depicts the proposed methodology of multisensory learning approach for complex human activity recognition. The overall performance of complex human activity recognition system depends on the following components.

- (i) Sensing of human activities
- (ii) Pre-processing of sensor raw data
- (iii) Features extraction and selection
- (iv) Classification Algorithms.

3.1 Sensing of Human Activities

For complex human activity recognition, selection and positioning of a sensor is a most challenging step. The appropriate placement of sensor on human body leads toward the better performance of human activity recognition system. The basic daily life activities e.g., walking, sleeping, running, sitting, stair down, stair up, dancing etc. are recognized by using smartphone built-in sensors (i.e., pocket position) only. Whereas the complex human activities like eating, drinking coffee, typing, playing, smoking and talking etc. are not easily identified by using

smartphone sensors due to the involvement of hand movements. Therefore, wrist worn sensor and smart watches are used for better recognition of these activities. In this work, we have used a multisensory and multilocational approach which provides better results for the identification of several complex human activities. It involves the overlap positioning, having smart phone in pocket/hand and a smart watch at wrist position. After sensing the human activity, the data fusion of cross pending sensors of both smart devices provides more useful information against different human physical activities. We have used following major components for better sensing of more complex human activities.

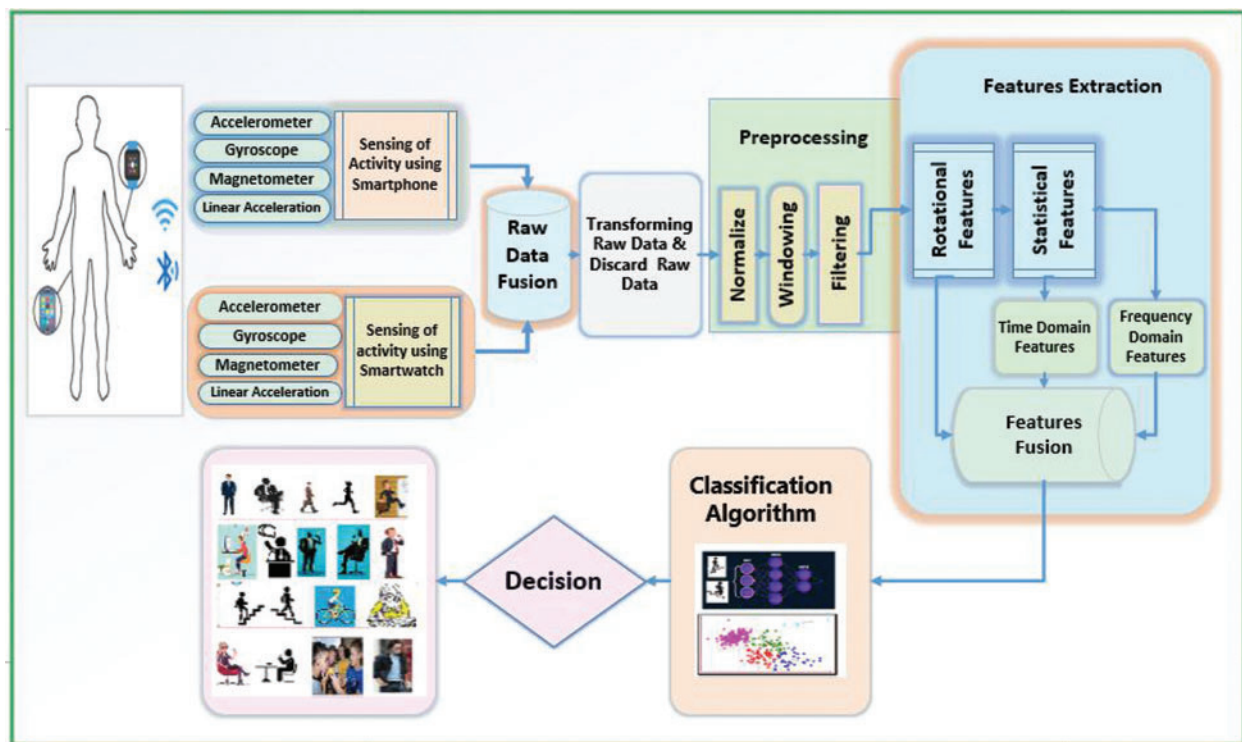


Figure 1: Proposed method of a multisensory learning approach for complex human activity recognition

- (i) Multisensory approach
- (ii) Multilocational sensor placement
- (iii) Sensor data fusion

3.1.1 Multisensory Data Collection

Data for 13 basic human activities from [23,26] is integrated with the publicly available dataset for 13 complex activities and a complete data set of 26 basic and complex activities is formed. An android application named “Linear data collector V2” [20] was used for sensor data logging against each human activity performed by multiple individuals of different age. Data is collected through accelerometer, gyroscope, linear acceleration sensor, and magnetometer, at the rate of 50 samples per second.

The application easily interfaces with the sensor for the recording and collection of data. Before recording sensor data, the application required the name and ID of participants in order to keep track of different participants. During the activity, the android device interacting with the application is placed at specific body position. The data has been collected as comma-separated values (CSV) file on the smart device which can be processed later. The system has also been tested for 14 overlapped basic and complex activities. Hence, final dataset comprises of more than 25 human activities. Each physical activity has been performed by multiple individuals for more than 60 s with known labels and timestamp. Fifty samples (50 Hz of each sensor output) have been collected for different human physical activities performed by multiple individuals. Thus, for each device that consist of four triaxial sensor, total ($4 * 3 = 12$) dimensional data is collected at a single instant. Particularly, for one second time period against a single human activity, 50 sample of each 12-dimensional data is collected from each device. We have collected hundreds of samples against each human activity performed by single participant and finally series concatenate the samples of all participants. A big dataset of similar complex human activity leads to complexity and misclassification of recognition system. To overcome these barriers and forming an accurate activity recognition system, preprocessing and feature extraction techniques are used.

3.1.2 Sensor Data Fusion

12-dimensional data has been collected from each device (smart phone, smart watch) with frequency of 50 Hz, where each sample of data is labeled and time stamp. For better understanding of complex human activities, we have collected the data from smart phone and smart watch at the same instant. Afterwards, concatenate the data from both devices in parallel. Hence, our sample space dimensions become 24 against each each human activity.

3.2 Preprocessing and Feature Extraction

The human physical activity recognition system has not been classified directly using the raw data of sensor. The classification task has been performed by using structural data representation (features vector) obtained from several pre-processing and feature extraction techniques. Each sensor generates three time series, along x-axis, y-axis and z-axis. After preprocessing, the tri-dimensional (x, y, z) raw data from four sensors of each smart device contain 12-dimensional vector. Subsequently, the data fusion of both devices with corresponding sensor makes the data 24 dimensional. Afterwards, feature extraction increases the feature vector size up to 72 dimensional features, which means 9 features from each sensor of each device. The features vectors comprise of rotational, time domain and frequency domain features.

The instantaneous rotational feature derived from orientation of device, like pitch (θ) feature is rotation over x-axis, roll(ϕ) feature is rotation over z axis and module of acceleration vector (α). These rotational features are calculated as;

$$Pitch(\theta_i) = \tan^{-1} \left[\frac{sy_i}{\sqrt{sx_i^2 - sz_i^2}} \right] \quad (1)$$

$$Roll(\phi_i) = \tan^{-1} \left[-\frac{sx_i}{sz_i} \right] \quad (2)$$

$$Module\ of\ acceleration\ vector\ (\alpha_i) = \sqrt{sx_i^2 + sy_i + sz_i^2} \quad (3)$$

Let $sx_i; sy_i; sz_i$ are x; y; z sensors reading for all four sensors $i = 1; 2; 3; 4$ of each device i.e., smartphone, smartwatch. These features are normalized using the `mat2gray()` command in Matlab. Furthermore, we have derived some statistical features over a defined period. Six features of each sensor have been extracted using windowing method. For a given input data X_t , a window of size k calculated as; $X_t = [x(t); x(t - 1); x(t - 2); \dots x(t - k)]$. The window is filtered using an average filter of size q , using the `fspecial()` and `imfilter()` functions in Matlab, where; $q = \text{floor}(k^{\frac{2}{3}})$. Once the averaged window for particular timeperiod(t) has been obtained, the next six features for each sensor are computed. Features 4–6 are the variances of the pitch, roll, and acceleration magnitude over the window, calculated by unbiased variance as follows:

$$\sigma^2 \beta_i = \frac{1}{n-1} \sum_{n=1}^k (x \beta n_i - \bar{x} \beta_i), \quad (4)$$

where β can be, θ, ϕ, α or \bar{u} , the mean of the input feature over window, and n are number of sensors. Features 7–9 are components of the fast Fourier transform (FFT) over the window. The preceding steps provide us all the nine features of each sensor, hence, a set of 72 features for each time sample has been obtained. Feature space information has been given below. To find the time series components, the data from four sensors of smartphone $s_1; s_2; s_3; s_4$ is collected, where; s_1 is accelerometer sensor data, s_2 is linear acceleration sensor data, s_3 is gyroscope sensor data and s_4 is magnetometer data. Each sensor contain axis (x; y; z) of time series (t) data, so sensor s_1 data represented as $s_{1x}; s_{1y}; s_{1z}$ that show x-axis, y-axis and z-axis data of 1st sensor. Similarly, s_2, s_3 and s_4 data of smartphone is denoted by $(s_{2x}; s_{2y}; s_{2z}), (s_{3x}; s_{3y}; s_{3z}), (s_{4x}; s_{4y}; s_{4z})$ respectively. Overall data is represented [Tab. 1](#).

$$X_1(t) = [s_1(x, y, z) \dots s_4(x, y, z)] \quad (5)$$

The vector dimension of $X_1(t)$ is 12 in every sample of each human activity. Similarly, smartwatch sensors time series data is represented as;

$$X_2(t) = [s_1(x, y, z) \dots s_4(x, y, z)] \quad (6)$$

After combining the data of sensors using both devices in time series, we obtain $X(t)$.

$$X(t) = X_1(t) : X_2(t) \quad (7)$$

where $X(t)$ is 24-dimensional raw data in every sample of each human activity. Further, we have calculated rotational feature $V(t)$ i.e., Pitch (θ), Roll (ϕ) and acceleration magnitude (α) of raw data.

$$V_1(t) = [s_1(\theta, \phi, \alpha) \dots s_8(\theta, \phi, \alpha)] \quad (8)$$

Next 48 statistical features are obtained using windowing method over a small defined period we calculate variance in pitch $\sigma^2 \theta_i$, roll $\sigma^2 \phi_i$ and acceleration magnitude $\sigma^2 \alpha_i$

$$V_2(t) = [\sigma^2(\theta_1, \phi_1, \alpha_1) \dots \sigma^2(\theta_8, \phi_8, \alpha_8)] \quad (9)$$

In next step we have calculated 24 frequency domain features V_f using fast Fourier transform (FFT) over the window having size (3×3) . It can be given as;

$$V_f = [f(\theta_1, \phi_1, \alpha_1) \dots f(\theta_8, \phi_8, \alpha_8)] \quad (10)$$

Table 1: Contribution of sensors in feature extraction

Smartphone sensor	smart watch sensor	Smartphone sensor	smart watch sensor
Accelerometer (x, y, z)	Accelerometer (x, y, z)	Accelerometer (x, y, z)	Accelerometer (x, y, z)
Linear acceleration sensor (x, y, z)	Linear acceleration sensor (x, y, z)	Linear acceleration sensor (x, y, z)	Linear acceleration sensor (x, y, z)
Gyroscope (x, y, z)	Gyroscope (x, y, z)	Gyroscope (x, y, z)	Gyroscope (x, y, z)
Magneto meter (x, y, z)	Magneto meter (x, y, z)	Magneto meter (x, y, z)	Magneto meter (x, y, z)
4 Triaxial sensor of smartphone	4 Triaxial sensor of smartwatch	4 Triaxial sensor of smartphone	4 Triaxial sensor of smartwatch
1 Sample: $4 \times 3 = 12$ -dimensional Data	1 Sample: $4 \times 3 = 12$ -dimensional Data	1 Sample: $4 \times 3 = 12$ -dimensional Data	1 Sample: $4 \times 3 = 12$ -dimensional Data
· Both devices sensor data fusion: $12 \times 2 = 24$ -dimensional (column) data			
Add Label and Time Stamp: $24 + 2 = 26$ column in one sample			
—Total number of human physical activity classes: 26			
· Total number of samples: 23, 75, 713			
· Total sample space dimension is: $[23, 75, 713 \times 26]$			
24 Rotational features:			
· Calculate three orientation or rotational feature (Pitch (θ), Roll(ϕ), Magnitude of acceleration(α)) from each sensor			
· Number of triaxial sensor = 4 (smartphone) + 4 (smartwatch) = $8(s_1, s_2, s_3, \dots, s_8)$			
· 8 (Sensor) * 3 (features(θ, ϕ, α)) = 24 rotational features			
24 Time domain statistical features			
· Calculate variance over Pitch ($\sigma^2\theta$), Roll ($\sigma^2\phi$), Magnitude of acceleration ($\sigma^2\alpha$) data against each sensor			
· Number of triaxial sensor = 4(smartphone) + 4(smartwatch) = $8(s_1, s_2, s_3, \dots, s_8)$			
· 8 (Sensor) \times 3(variance features ($\sigma^2\theta, \sigma^2\phi, \sigma^2\alpha$)) = 24 Time domain statistical features			
24 Frequency domain statistical features			
· Calculate absolute Dc component of FFT over (Pitch ($f\theta$), Roll ($f\phi$), Magnitude of acceleration ($f\alpha$)) data against each sensor			
· Number of triaxial sensor = 4(smartphone) + 4(smartwatch) = $8(s_1, s_2, s_3, \dots, s_8)$			
· 8 (Sensor) \times 3 (basic component of FFT features ($f\theta, f\phi, f\alpha$)) = 24 Frequency domain statistical features			
Total number of features obtained $24 + 24 + 24 = 72$			
Total feature space is $[23, 75, 713 \times 72]$			

Finally, the whole sensor data has been fused using Eqs. (10)–(12) in order to obtain final feature space. Let Z be the total number of features.

$$Z = |V_1 : V_2 : V_f| \quad (11)$$

$$Z = |Z_1 : Z_2 \dots Z_8| \quad (12)$$

$$Z_1 = |\theta_1, \phi_1, \alpha_1 : \sigma^2\theta_1, \sigma^2\phi_1, \sigma^2\alpha_1, f\theta_1, f\phi_1, f\alpha_1| \quad (13)$$

Hence, the whole dataset has been formed which is then utilized for classification process by splitting the data for training and testing process with ratio 8:2. Finally, different classification algorithms applied which classify the several human physical activities. Fig. 2 summarizes the proposed methodology of a multisensory learning approach of complex human activity recognition.

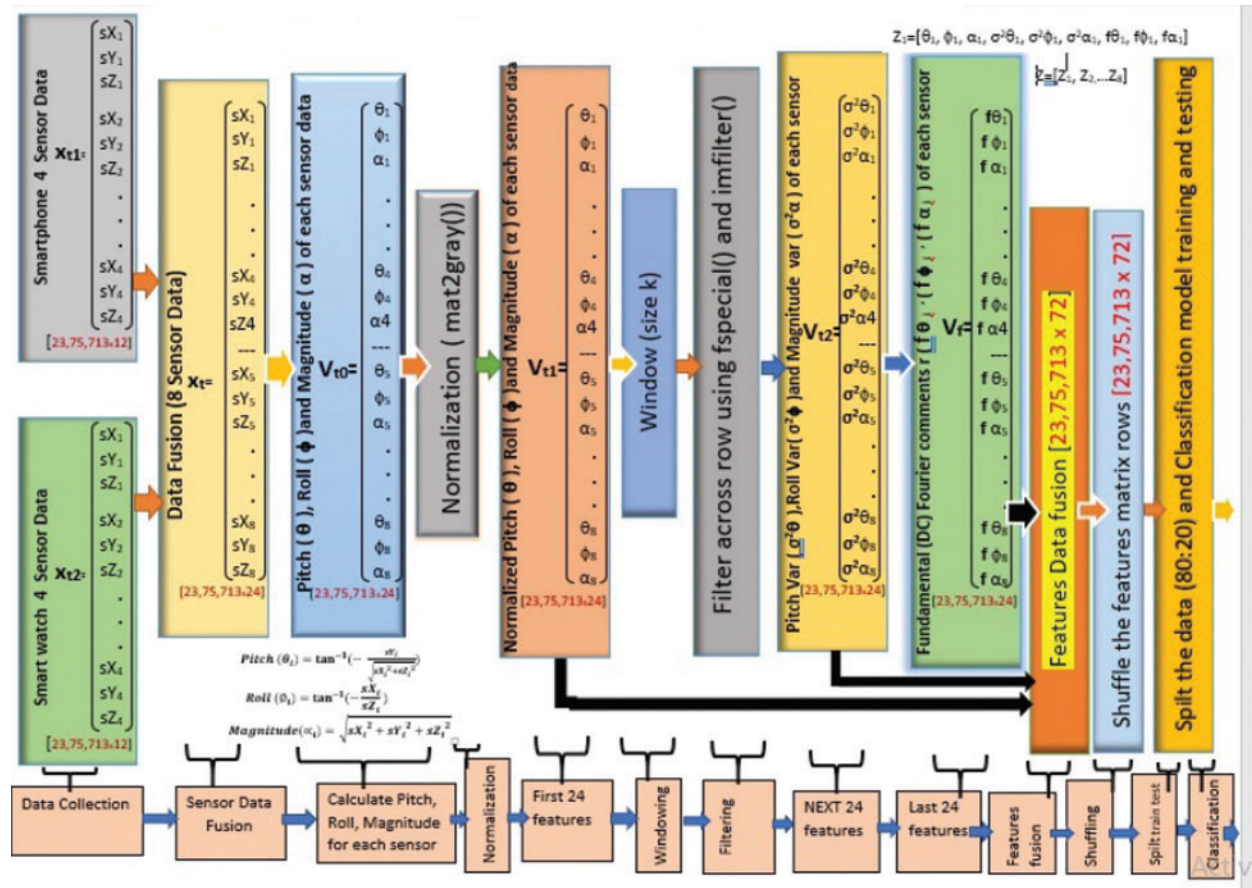


Figure 2: Proposed methodology of a multisensory learning approach of complex human activity recognition

4 Results and Analysis

In literature, researchers have used different approaches to classify the human activities using sensor-based data. In this work, we have used Naive Bayes (NB), K-Nearest Neighbors (KNN) and Neural Network (NN) for classification.

4.1 Experimental Configuration

To perform Naive-Bayes Classification, we split the data into two groups: 80% of the data is used for training of the classifier and remaining 20% is reserved as testing data. Let P be the sensor readings in a single class (where J be total no of classes), and x_p denotes the p th sensor reading from the training dataset. The elements of x_p are 12 sensor readings (3-axis for 4 sensor).

We collected the whole training data and convert the data into a single vector. Then, we count the frequency of each value and use it to construct the probability of each value in the class. This gives us $p(x_{ij} = C_j)$ for each value x_i and each class C_j . Then we have determined whether a sensor reading belongs to class a or b , by taking all the elements of the new reading x_{testi} , and calculated it as;

$$\prod_{i=1}^n p(x_{testi}/y = C_a) \quad (14)$$

$$\prod_{i=1}^n p(x_{testi}/y = C_b) \quad (15)$$

If (14) > (15), we can assume that new reading belongs to class a otherwise new reading belongs to class b . The classification is readily extended to comparison of multiple classes by taking the maximum.

$$\operatorname{argmax} x_j \prod_{i=1}^n p(x_{testi}/y = C_j) \quad (16)$$

This gives overall classes being compared with C_j . For human activity recognition using KNN, number of neighbor K is assigned to each number of human physical activities. In this experiment, we have assigned 26 activities as “ k number neighbor.” For several human activity recognition system KNN is used as multiclass classifier. We have calculated the Euclidean distance between the data points. The distance between two points is Euclidean distance, it has been calculated as:

$$\text{Euclidean distance between } A_1 \text{ and } B_2 = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (17)$$

Similarly, for multipoint, we have repeated the above procedure. We have calculated KNN Euclidean distance from mean value of each neighbor group. After re-assigning each datapoint to new class of minimum distance, we have calculated centroid of these neighbor groups. We have repeated the procedure until datapoints left with only few fixed numbers. The parameters of neural network (NN) i.e., number of neuron and number of hidden layers have been selected experimentally, until reached to minimum value of cross-entropy error of testing data. The simulation has been implemented using neural network toolkit of MATLAB with custom coding to overcome the limitations of the MATLAB GUI. We have used following configurations for artificial neural networks: Inputs features = 72, output classes = 26, Two hidden layers [110, 75 neurons], Sigmoid Activation function, Conjugate gradient training function, Error Back propagation, 500 epochs. We have analyzed our proposed methodology using multiple classifiers including Neural Network-Nearest Neighbor, Naive Bayes, Ensemble method AdaBoost Decision Tree. For ensembles purpose, we have used MATLAB learner GUI application. The classification performance of these class er has been given in the next sections.

4.2 Simulation Results

As we have explained earlier, four sensors of each device i.e., smartphone and smartwatch are used for the sensing of human physical activities. Nine features are derived from each sensor raw data. Although various sensors have been used for the identification of human activities, however,

each feature derived from a sensor plays an independent role in sensing and recognizing of specific activity. The classification algorithm recognizes 26 basic and complex activities listed previously. Two types of datasets i.e., pocket position of smart phone and wrist position of smart watch have been considered which are lately classified by various classification methods. These methods include Naive Bayes, K-nearest neighbor, and Neural Network (NN).

Neural network having 2 hidden layers is implemented for activity recognition. The MATLAB pattern recognition (PR) neural network (NN) tool has been used. The configuration set as the number of given inputs and classified outputs are 72 and 26 respectively. The total of 72 inputs passed through first hidden layer of 75 neurons and second hidden layer of 50 neurons and lastly classify the 26 human activities. For classification purposes, Sigmoid activation function has been used which led to conjugate gradient training function and back propagation for error minimization.

The NN gather results for 500 epochs and it has been proven that NN with 2 hidden layer and KNN perform better than other algorithms. Accuracy of classifier in term of percentage is calculated as: Percentage Correct Classification = $100 * (1 - C_r)$ Percentage Incorrect Classification = $100 * C_r$. Where, C_r is a confusion value that shows declassification rate and obtained from MATLAB function confusion which returns false positive, false negative, true positive and true negative rate information. The Neural Network provides the accuracy of 99.340162% which is the highest accuracy among all the algorithms and make it more efficient. The confusion matrices of Naive Bayes, KNN and NN are given in Figs. 3–5 respectively.

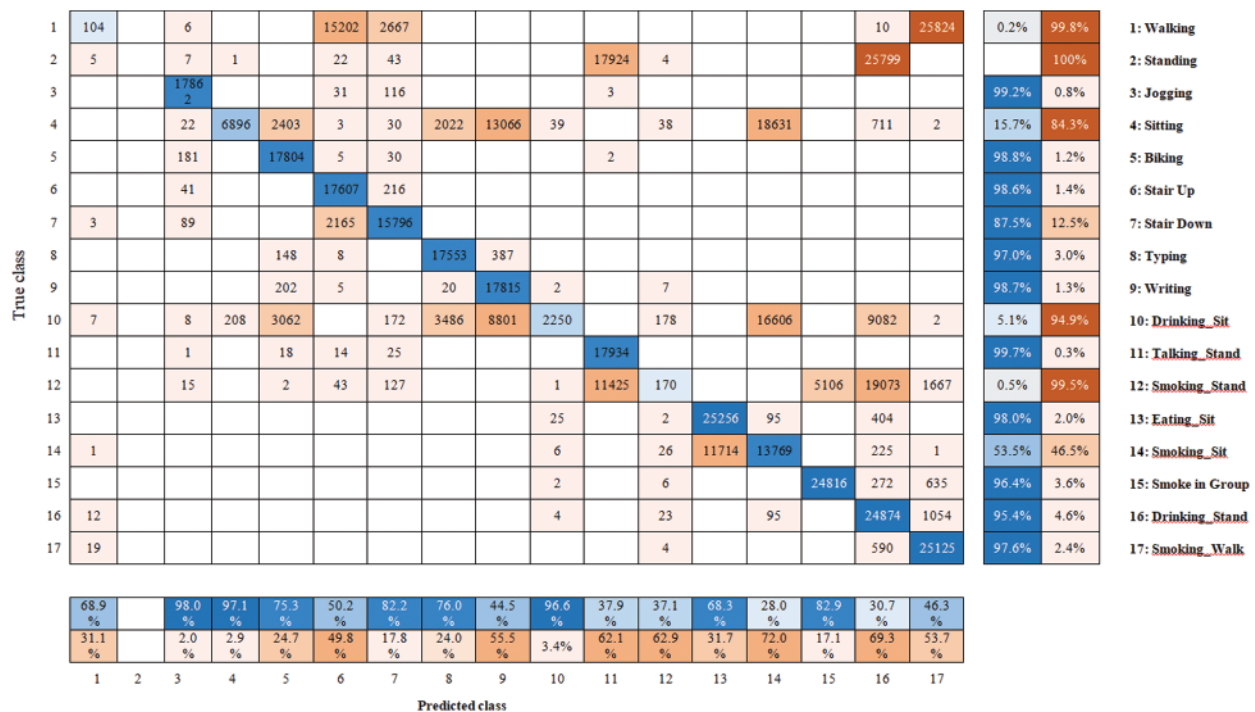


Figure 3: Confusion matrix of Naive Bayes classifier

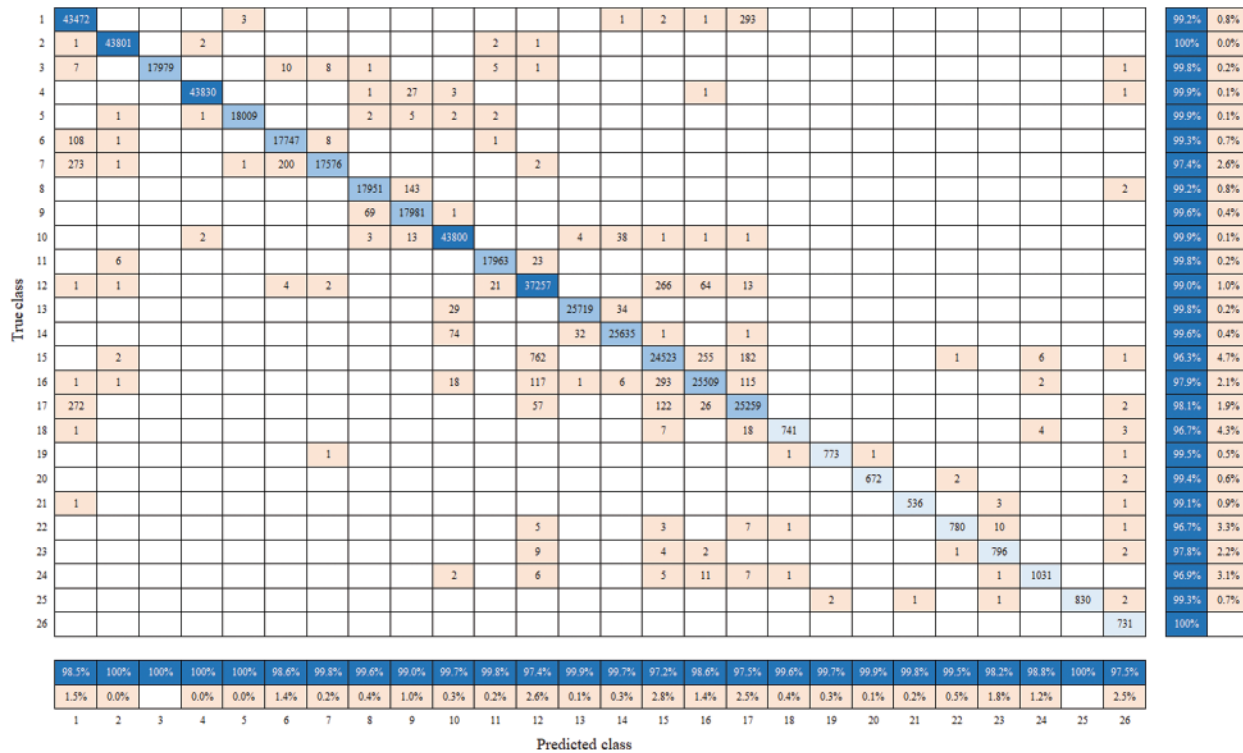


Figure 4: Confusion matrix of K-nearest neighbor classifier

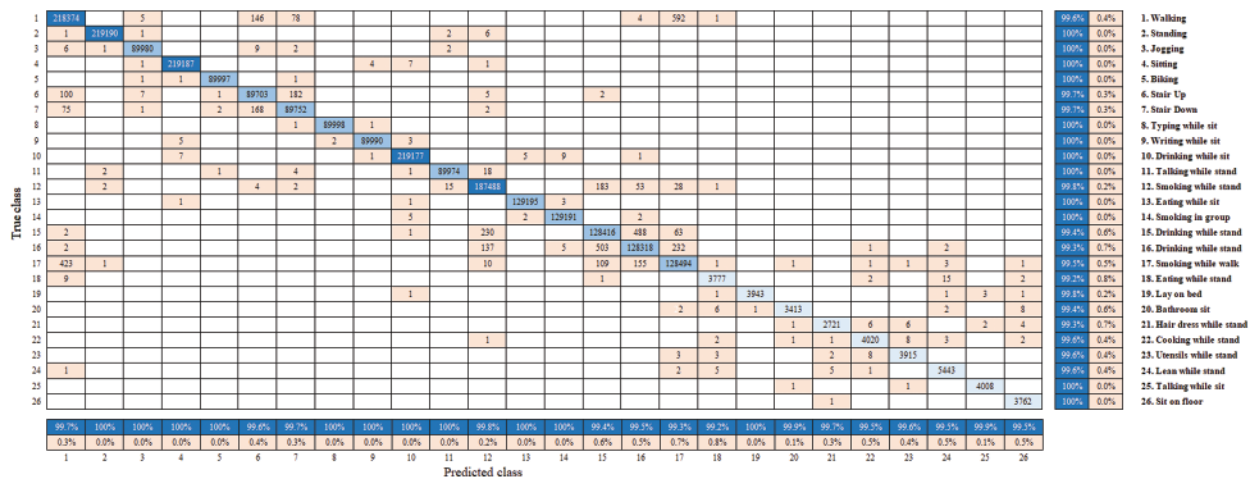


Figure 5: Confusion matrix of neural network classifier

5 Conclusion

The reliable recognition of several human physical activities of daily life can be very helpful for many applications like eHealth, remote monitoring and tracking of human for awareness feedback, coaching, human machine interaction, bad habits motoring. etc. In this paper, a multisensory learning approach of complex human activity recognition is proposed that provide

better recognition performance for a large number of complex human physical activities of daily life. Neural network (NN), Naive Bayes (NB), K-Nearest Neighbor (KNN) and Ensemble method AdaBoost classifiers are used along a proposed mathematical model of preprocessing and feature extraction. Neural network and KNN classifiers outperformed other classifiers. It is further concluded that the smart device based multisensory based human activity recognition is a cost effective and more practical solution rather than vision based or dedicated sensor-based approaches. Furthermore, in this work a new data set against 26 human physical activities of daily life is formulated using built-in sensors of smart phone and smart watch, that will be helpful for future research in this field. The data of smartphone and smartwatch for a large number of complex human physical activity will serve as a benchmark.

Funding Statement: This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2018R1D1A1B07042967) and the Soonchunhyang University Research Fund.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] F. Afza, M. Sharif, S. Kadry, G. Manogaran, T. Saba *et al.*, “A framework of human action recognition using length control features fusion and weighted entropy-variances based feature selection,” *Image and Vision Computing*, vol. 106, pp. 104090, 2021.
- [2] F. Zahid, J. H. Shah and T. Akram, “Human action recognition: A framework of statistical weighted segmentation and rank correlation-based selection,” *Pattern Analysis and Applications*, vol. 23, no. 1, pp. 281–294, 2020.
- [3] M. Zahid, F. Azam, M. Sharif, S. Kadry and J. R. Mohanty, “Pedestrian identification using motion-controlled deep neural network in real-time visual surveillance,” *Soft Computing*, vol. 1, pp. 1–17, 2021.
- [4] Z. Qin, Y. Zhang, S. Meng, Z. Qin and K. K. R. Choo, “Imaging and fusing time series for wearable sensor-based human activity recognition,” *Information Fusion*, vol. 53, no. 1, pp. 80–87, 2020.
- [5] A. Henriksen, J. Johansson, G. Hartvigsen, S. Grimsgaard and L. Hopstock, “Measuring physical activity using triaxial wrist worn polar activity trackers: A systematic review,” *International Journal of Exercise Science*, vol. 13, pp. 438, 2020.
- [6] J. Lu, X. Zheng, M. Sheng, J. Jin and S. Yu, “Efficient human activity recognition using a single wearable sensor,” *IEEE Internet of Things Journal*, vol. 7, no. 11, pp. 11137–11146, 2020.
- [7] O. Barut, L. Zhou and Y. Luo, “Multitask LSTM model for human activity recognition and intensity estimation using wearable sensor data,” *IEEE Internet of Things Journal*, vol. 7, no. 9, pp. 8760–8768, 2020.
- [8] M. Sharif, T. Akram, M. Raza, T. Saba and A. Rehman, “Hand-crafted and deep convolutional neural network features fusion and selection strategy: An application to intelligent human action recognition,” *Applied Soft Computing*, vol. 87, pp. 105986, 2020.
- [9] N. Hussain, M. A. Khan, S. A. Khan, A. A. Albeshir, T. Saba *et al.*, “A deep neural network and classical features based scheme for objects recognition: An application for machine inspection,” *Multimedia Tools and Applications*, vol. 1, pp. 1–23, 2020.
- [10] I. Haider, M. Nazir, A. Armghan, H. M. J. Lodhi and J. A. Khan, “Traditional features based automated system for human activities recognition,” in *2020 2nd Int. Conf. on Computer and Information Sciences*, Al-Jouf, SA, pp. 1–6, 2020.
- [11] Y. D. Zhang, S. A. Khan, A. Rehman and S. Seo, “A resource conscious human action recognition framework using 26-layered deep convolutional neural network,” *Multimedia Tools and Applications*, vol. 1, pp. 1–23, 2020.

- [12] A. Mehmood, M. Sharif, S. A. Khan, M. Shaheen, T. Saba *et al.*, “Prosperous human gait recognition: An end-to-end system based on pre-trained CNN features selection,” *Multimedia Tools and Applications*, vol. 3, pp. 1–21, 2020.
- [13] T. Akram, M. Sharif, N. Muhammad, M. Y. Javed and S. R. Naqvi, “Improved strategy for human action recognition; Experiencing a cascaded design,” *IET Image Processing*, vol. 14, no. 5, pp. 818–829, 2019.
- [14] K. Aurangzeb, I. Haider, T. Saba, K. Javed, T. Iqbal *et al.*, “Human behavior analysis based on multi-types features fusion and Von Nauman entropy based features reduction,” *Journal of Medical Imaging and Health Informatics*, vol. 9, no. 4, pp. 662–669, 2019.
- [15] K. Javed, S. A. Khan, T. Saba, U. Habib, J. A. Khan *et al.*, “Human action recognition using fusion of multiview and deep features: An application to video surveillance,” *Multimedia Tools and Applications*, vol. 4, pp. 1–27, 2020.
- [16] A. Sharif, K. Javed, H. Gulfam, T. Iqbal, T. Saba *et al.*, “Intelligent human action recognition: A framework of optimal features selection based on Euclidean distance and strong correlation,” *Journal of Control Engineering and Applied Informatics*, vol. 21, pp. 3–11, 2019.
- [17] S. Kadry, P. Parwekar, R. Damaševičius, A. Mehmood, J. A. Khan *et al.*, “Human gait analysis for osteoarthritis prediction: A framework of deep learning and kernel extreme learning machine,” *Complex and Intelligent Systems*, vol. 11, pp. 1–19, 2021.
- [18] H. Arshad, M. I. Sharif, M. Yasmin, J. M. R. Tavares, Y. D. Zhang *et al.*, “A multilevel paradigm for deep convolutional neural network features selection with an application to human gait recognition,” *Expert Systems*, vol. 4, no. 3, pp. e12541, 2020.
- [19] A. Bayat, M. Pomplun and D. A. Tran, “A study on human activity recognition using accelerometer data from smartphones,” *Procedia Computer Science*, vol. 34, pp. 450–457, 2014.
- [20] G. Chetty and M. Yamin, “Intelligent human activity recognition scheme for eHealth applications,” *Malaysian Journal of Computer Science*, vol. 28, pp. 59–69, 2015.
- [21] J. J. Guiry, P. Van de Ven and J. Nelson, “Multi-sensor fusion for enhanced contextual awareness of everyday activities with ubiquitous devices,” *Sensors*, vol. 14, no. 3, pp. 5687–5701, 2014.
- [22] H. M. Ali and A. M. Muslim, “Human activity recognition using smartphone and smartwatch,” *International Journal*, vol. 3, no. 10, pp. 568–576, 2016.
- [23] S. K. Polu and S. Polu, “Human activity recognition on smartphones using machine learning algorithms,” *International Journal for Innovative Research in Science & Technology*, vol. 5, pp. 31–37, 2018.
- [24] M. Shoaib, S. Bosch, O. D. Incel, H. Scholten and P. J. Havinga, “Complex human activity recognition using smartphone and wrist-worn motion sensors,” *Sensors*, vol. 16, no. 4, pp. 426, 2016.
- [25] Y. Dong, J. Scisco, M. Wilson and A. Hoover, “Detecting periods of eating during free-living by tracking wrist motion,” *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 4, pp. 1253–1260, 2013.
- [26] S. Sen, V. Subbaraju, A. Misra and Y. Lee, “The case for smartwatch-based diet monitoring,” in *2015 IEEE Int. Conf. on Pervasive Computing and Communication Workshops*, St. Louis, MO, USA, pp. 585–590, 2015.
- [27] P. M. Scholl and K. Van Laerhoven, “A feasibility study of wrist-worn accelerometer based detection of smoking habits,” in *2012 Sixth Int. Conf. on Innovative Mobile and Internet Services in Ubiquitous Computing*, Palermo, Italy, pp. 886–891, 2012.
- [28] L. Köping, K. Shirahama and M. Grzegorzec, “A general framework for sensor-based human activity recognition,” *Computers in Biology and Medicine*, vol. 95, no. 8, pp. 248–260, 2018.
- [29] M. B. Tahir, K. Javed, S. Kadry, Y. D. Zhang, T. Akram *et al.*, “Recognition of apple leaf diseases using deep learning and variances-controlled features reduction,” *Microprocessors and Microsystems*, vol. 1, pp. 104027, 2021.
- [30] M. S. Sarfraz, M. Alhaisoni, A. A. Albeshir, S. Wang and I. Ashraf, “StomachNet: Optimal deep learning features fusion for stomach abnormalities classification,” *IEEE Access*, vol. 8, pp. 197969–197981, 2020.