

OHAR: Optimized Human Action Recognition Paradigm using Optimized Type 2 Neuro-Fuzzy Classifier

DR. J.A. SMITHA[1,*], RAGHU RAMAMOORTHY[2] AND DR. R CHINA APPALA NAIDU[3]

[1]Professor, Department of Computer Science and Engineering, Sri Sairam College of Engineering, Bengaluru, Karnataka, India, (smithajeyakumar37@gmail.com)

[2]Assistant Professor, Department of Computer Science and Engineering, The Oxford College of Engineering, Bengaluru, Karnataka, India, 560 068. (raghuace85@gmail.com)

[3]Professor and Head, Department of Computer Science and Engineering, The Oxford College of Engineering, Bengaluru, Karnataka, India, 560 068 (chanr789@gmail.com).

[*] Corresponding author: DR. J.A. SMITHA

Abstract

Human activity recognition (HAR) is made to identify actions and goals of persons one or more from the images which contain sequence of actions related on environments and actions. However, different issues and challenges are increased in the applications of human activity recognition for improving detection accuracy with different activities. Hence, Optimized Human Action Recognition Paradigm (OHAR) is developed. In the paper, optimized type 2 fuzzy classifier is designed to classify human actions from the image database. The input video is transformed into multiple region in the initial stage. The collected frames are sent to the pre-processing stage for removing noise from frames. After that, the key frame is selected from the image frames by using the Structural Similarity Index (SSIM). Once key frames are selected, the three feature extraction methods are utilized like Space-Time Interest (STI) points, grid shape feature, and coverage factor. Finally, the proposed classifier is proceeding to human activity recognition with selected features set. Here, an optimized type 2 neuro-fuzzy classifier is used for detecting human action. The proposed classifier is enhanced Rider optimization algorithm (ROA). The presentation of proposed method is evaluated based on statistical computations such as accuracy, sensitivity, specificity, recall, and F_Score.

Key Words: *Human actions, Fuzzy classifier, Actions, Fuzzy rules, Riders and Membership functions.*

Introduction

Human activity recognition is fascinated research from together industry and researchers in human-computer interaction, human behaviour analysis, and ubiquitous computing. Human activity recognition is used in different real-time applications like indoor navigation, military, tactical, gaming, and healthcare to personal fitness [1]. Two major types are available in human activity detection such as wearable sensors and systems. In sensor-based detection, wearable sensors can be connected to human body, and human activity is interpreted into exact sensor signal designs that is identified as well as segmented [2]. In major cases, environmental devices needed an installation device in addition human activity translated like cameras can be utilized as invasive devices particularly through ageing people. Due to the reasons, human activity recognition has been considered important research with the utilization of wearable sensors [3].

 https://doi.org/10.37896/pd91.4/91445

There are many ways to achieve recognition of human activities from video frames. But that exiting methods are not attained the better results. The HAR database may contain events known as triplets that can easily identify human activity [4, 5]. Similarly, some databases contain video frames without very hard labels. Also, large data sets make it difficult to classify objects and humans in video frames. The shortcomings mentioned above are motivated by the need for machine learning and in-depth learning structure. Human activities are identified using machine learning techniques like Neural Network (NN) [6], Bayesian Network AND Support Vector Machine (SVM) [7]. However, machine learning techniques fails in providing better results. So today, in-depth learning is concentrated on recognizing human activities through video. Deep learning can manage a large number of databases to identify human activities.

In depth machine learning techniques such as Deep Neural Network (DNN) [8], Deep Belief Neural Network (DPNN) and Conventional Neural Network (CNN) [9] are used to identify human activities using videos. In-depth learning methods are widely used in variety of applications. Among all, CNN image-based processing methods are particularly suitable for detection and authentication. But CNN can suffer because of the undefined structure for identifying objects from humans and images. Algorithms like particle swarm optimization (PSO), cuckoo search algorithm (CA) and genetic algorithm (GA) [10] are used to improve the performance of CNN architecture, respectively. Each optimization algorithm captures the concentration. Therefore, an upgraded version of in-depth learning is needed to employ optimal human function detection from video frames.

Literature Review

HAR identification is processed using various methods. Among them few works are analysed here;

Mohammad Mehdi Hassan *et al.*, [11] have introduced a system based on cell phone passive sensors for human movement approval. Efficient highlights are first derived from initial information. Features extraction includes autoregressive coefficients, average and medium. For further processing the extracted features are given to DNN classifier. The approach introduced was different and conventional pronunciation approval approaches, for example, the conventional multiclass SVM and ANN overcame them.

Ignatov Andrey *et al.*, [12] Here, it introduces the use of CNN for HAR, with direct factual features that store data about the global type arrangement. Besides, researching the effect of time arrangement length on the accuracy of approval and pointing it up to a second creates a continuous static operating sequence. The accuracy of the introduced approach is estimated in two commonly used WISDM and UCI databases, which contain individually marked accelerometer information from 36 and 30 clients, and analyze in cross database. The results showed that proposed model is better in class operation, while low computational cost and manual component design are not required.

Abdulmajit Murad *et al.*, [13] have introduced Deep Recurrent Neural Networks (DRNNs) to generate detection models adapted to capture long-distance conditions in factor length input heirs. Here, long short-term memory (LSDM) can provide unilateral, bilateral, and degraded formats based on DRNs and evaluate their reliability across various benchmark datasets. Study

 https://doi.org/10.37896/pd91.4/91445

results show that our presented models have overcome techniques that use conventional methods.

Seokong Zhao *et al.,* [14] have introduced human movement endorsement that enhances deep learning in IoT conditions. A semi-directed in-deep learning framework was planned to provide most accurate HAR, which productively uses and breaks the labeled sensor information to produce categorized learning model. To make the problem of less marked example easier totake care of a brilliant auto-naming scheme based on Deep Q-Network (DQN) was recently developed with a planned distance-based award rule that can improve learning performance in IoT conditions.

Nafiul Rashid *et al.,* [15] have introduced Adaptive convolutional neural network (CNN) is an energy-productive Human Activity Recognition (AHAR) appropriate for low-power edge gadgets. AHAR utilizes a versatile engineering that chooses a part of pattern design to use during the induction stage. Here, approves the introduced strategy in characterizing velocity exercises from two different datasets like Opportunity and w-HAR.

Design of the proposed methodology

Recently, human activity recognition is as considered one of the most essential study areas in computer vision technology. Human activity recognition is used in different applications like multimedia industry, consumer agencies, robotics, sports, content-based image retrieval, human-computer interaction, smart technologies, video indexing, video surveillance, and security. Hence, in this paper, the Optimized Type 2 Neuro-Fuzzy Classifier to identify human actions from the images. The proposed method is proceeding related to five phases as video to frame conversion, pre-processing, keyframe extraction, feature extraction, and classification. The complete block diagram of proposed method is illustrated in Figure 1.

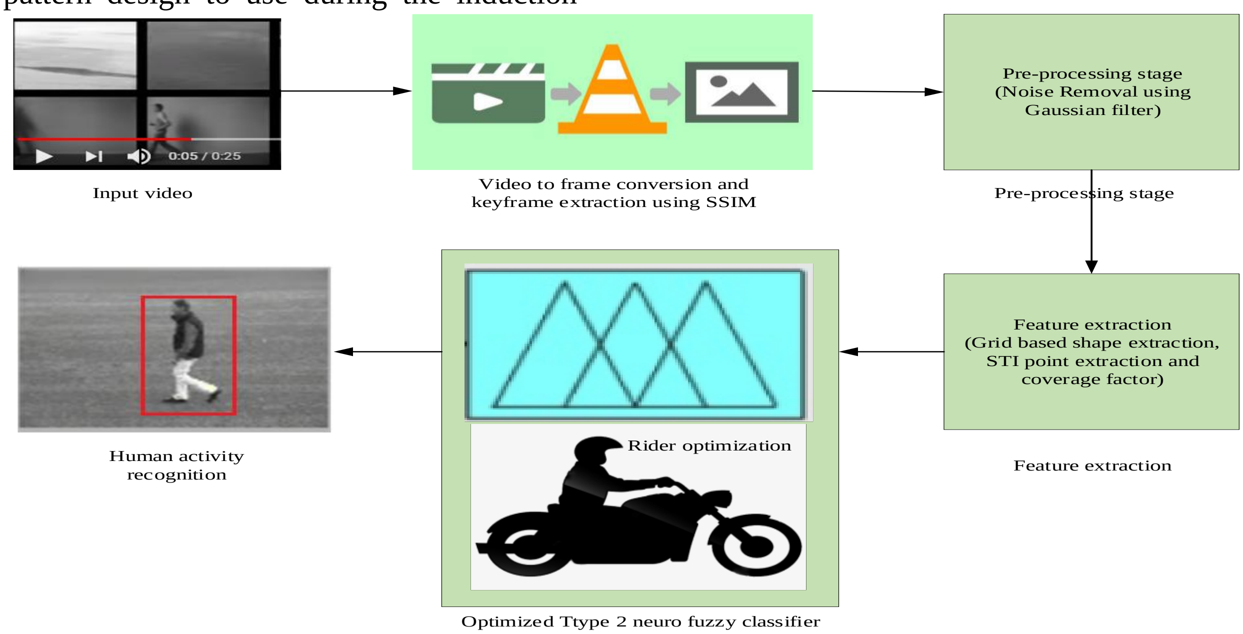

Figure 1. Block diagram of proposed methodology

 https://doi.org/10.37896/pd91.4/91445

ISSN: 0369-8963

At the beginning, video dataset is collected from internet. After collection, video is converted into frame. The frame is used for further processing. Before, processing, pre-processing is applied for each frame. Because, the frames may contains noise and unwanted information. In this paper, pre-processing step includes background subtraction, Gaussian model, and noise removal. The pre-processed image is sent to keyframe extraction. The keyframe extraction is proceeding with the SSIM index which is utilized to compute similarity among the frames. The similarity value is compared with the threshold value which has a larger value it is designated as the keyframes. After that, the features are selected by different methods such as coverage factor, STI points and grid-based shape. Finally, extracted features are sent to proposed classifier for human activity recognition.

(i) Video to frame conversion

For human action recognition, the input video is converted into different frames. Using single frame, we cannot able to find the human action. Each and every frame are analysed for recognition process. The extracted frames are used for further processing.

(ii) Pre-processing stage

Pre-processing is performed in order to reduce the noise frames in video and to enhance image contrast. Human mobility changed in every frame of the video. So we process all the frames. Due to noise, the detection may fail to achieve best results. For noise removal and enhancement Gaussian filter is utilized. The Gaussian filters works, based on the pixel strength [16]. The intensity of a Gaussian action's probability and a frame's size are combined to form pixel strength. The Gaussian model can be created as follows:

$$M(f) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{f^2}{2\sigma^2}} \quad (1)$$

Where σ represent the standard deviation and f represent the inter-frame distance. The benefits of Gaussian filter includes reducing noise, blurring the edges from images and reduces hassle.

(iii) Key frame extraction

Frames are containing the major information of the images, but some of the frames are provided repeat information of actions. These frames have occupied the spaces and also increase the burden of the classifier. Hence, the keyframes are selected based on SSIM index values [17]. The keyframe extraction is essential requirement for effective human activity recognition. The sequence of video frames is mathematically formulated as follows,

$$F = \{f_1, f_2, \dots, f_j, \dots, f_m\} \quad (2)$$

Where, f_j can be represented as a jt^{h} frame of the video, m is described as several frames in the video. From the number of frames, the keyframes are selected based on SSIM which formulated as follows,

$$SSIM(f_j, f_{j+1}) = \frac{(2\mu_j\mu_{j+1} + S_1)(2\sigma_{jj+1} + S_1)}{(\mu_j^2 + \mu_{j+1}^2 + S_1)(\sigma_j^2 + \sigma_{j+1}^2 + S_2)'} \quad (3)$$

Where, S_1, S_2 can be described as two variables, σ_{jj+1} can be described as the covariance of f_j and f_{j+1}, σ_{j+1}^2 can be described as a variance of f_{j+1}, σ_j^2 can be described as a variance of f_j, μ_{j+1} can be described as an average of f_{j+1} and μ_j can be described as an average of f_j respectively. From the frames, the keyframes are selected

https://doi.org/10.37896/pd91.4/91445

which are utilized to compute the features for human activity recognition.

(iv) Feature extraction

Once, the keyframes are selected, feature extraction procedure is initialized. The projected method is extracting the features like STI points, coverage factor, and shape features. The grid-based method is used to extract shape features from selected keyframes. The feature extraction process is mathematically formulated as follows,

Grid-based shape extraction

The shape of human activity is an essential feature extraction, utilized by grid method for human activity detection. In a grid-based method, the human body shape is extracted. The grid-based method is a pixel-oriented, in which the shape is planned to a grid of ideal measurement. In proposed methodology, 4*4 grid size is selected and the human shape in every keyframe is split into a cell with specified measurements. After that, the grid is skimmed from the top to bottom and left to the right frame. This process is initiated from the top left corner. In the keyframes, every cell is fixed as a binary parameter is 1, if the cell is changed into a human shape and 0 for the remaining cases [18]. Hence, the binary sequence of 0 and 1 with a matrix, it constructed as a matrix element. With the consideration of matrix parameter, the shape and thereby, the human activity is detected.

STI point extraction

To analyse spatiotemporal measures in video, local structure in space and time-based features are essential. This feature of image values is changing in combined domains [19]. The image noted that have a local change of strengths incline to be information which named as 'interest points'. To compute local features of the frame, the scale-space symbol is formulated as follows,

$$k(.,\eta^2,\tau^2) = k * U(.,\eta^2,\tau^2) \quad (4)$$

Where U can be described as Gaussian kernel, η,τ can be represented as spatial and temporal parameters. The Gaussian kernel is mathematically formulated as,

$$U = \frac{\exp\left(-\frac{b^2+c^2}{2\eta_x^2} - n^2/2\tau_x^2\right)}{\sqrt{(2\pi)^3\eta_x^4\tau_x^2}} \quad (5)$$

Where n can be described as a total number of frames. After that, a second-moment matrix can be computed by utilizing the spatiotemporal image gradients which formulated as follows,

$$\mu(.,\eta^2,\tau^2) = k * U(.,\eta^2,\tau^2) * (\nabla k(\nabla k)^T) \quad (6)$$

The local maxima feature is computed by defining the feature position which formulated as follows,

$$h^s = det(\mu) - trace^3(\mu) \quad (7)$$

In the feature extraction process, the spatiotemporal extent of the frame is utilized to calculate the feature size of the frame by choosing parameters. The spatiotemporal neighbor is formulated as follows,

$$X = (k^a, k^b, k^n, k^{aa}, \dots, k^{nnn}) \quad (8)$$

This feature extraction process is finding out the center of features with the consideration of normalized derivatives which can be computed related on the specified parameters. The velocity is computed, the neighborhood features also attuned to calculate the invariance related to the camera gestures. Hence, the STI point is extracted from the frames, a vector of measurement 1*40 is achieved.

(v) Coverage factor

This feature is utilized to compute the coverage factor-related to the coverage or

https://doi.org/10.37896/pd91.4/91445

shape of the human activity from the images. To achieve human activity, the mean of the interesting point is extracted. Related to the mean value, the center opinion of human activity can be achieved. Additionally, the location of the human activity is identified with the basis of vertical length and horizontal length. Hence, the coverage of the human activity, the vector size is fixed as 1*2. So, in the feature extraction procedure, the feature vector is achieved. The collected vector is sent to the classification phase [20].

Proposed Optimized Type 2 Neuro-Fuzzy Classifier

In the proposed classifier, a combination of type-2 neuro-fuzzy classifier and ROA is utilized to recognize human activity from the images. The ROA is utilized to empower the type 2 neuro-fuzzy classifier efficiency in human activity recognition. The membership function of fuzzy classifier is improved with the help of the ROA algorithm. The process of type 2 neuro-fuzzy classifier and ROA is presented in this section.

(i) Type 2 neuro-fuzzy classifier

This method is used to make human activity recognition systems intelligent. Additionally, it reduces the burden of humans and it does not consider rigid distance metrics to decide the image similarity for activity recognition. To achieve efficient HAR, fuzzy logic is an efficient tool. It is a most efficient and flexible technique to manage the combination of measurements with their degree of uncertainty in HAR. In this study, an attempt is made to model the uncertainty of the similarity of image features with the help of Type 2 ambiguous model [21]. The Type 2 fuzzy model features a pacifier, rule-based operation, fuzzy inference engine and output processor, and includes a type reducer and diffuser. FL is again characterized by IF-THEN rules but not its effect or precursor packages. Obscure packages are processed by an uncertainty trail (FOU), which is characterized by their boundaries such as bottom and top member functions. In general, type 2 is more severe and computationally more severe due to type-reduction. In the Type-2 ambiguous classifier, the secondary member is either one or zero. The type 2 fuzzy classifier is mathematically formulated as follows,

$$\tilde{S} = \int_{x \in X} \int_{U \in J^X \sqsubseteq [0,1]^{1/(x,u)}} = \int_{x \in X} \int_{U \in J^X \sqsubseteq [0,1]^{(1/u)/6}} \tag{9}$$

Where, $\tilde{S}$ can be described as a union of all primary memberships, J^X can be described as the primary membership of x, U is the second variable and X is the second variable. The uncertainty of $\tilde{S}$ is considered as equal to one value. The union of the primary membership function is named FOU which presented as follows,

$$FOU(\tilde{S}) = U_{x \in X} J_x \tag{10}$$

$$\tilde{\mu}^{\tilde{A}}(x) = \overline{FOU(\tilde{A}) \forall_x} \,\epsilon X \tag{11}$$

Where, $\tilde{\mu}^{\tilde{A}}(x)$ can be represented as upper membership function. For a continuous universe of distance U and X, the embedded interval of the primary membership function is denoted as follows,

$$\tilde{S}_e = \int_{x \in X} \left[\frac{1}{\theta}\right] x \quad \theta \epsilon J_x \sqsubseteq U = [0,1] \tag{12}$$

$$S_e = \int_{x \in X} \left[\frac{\theta}{x}\right] \quad \theta \epsilon J_x \sqsubseteq U = [0,1] \tag{13}$$

Where, $\tilde{S}_e$ can be described as a union of all primary memberships of the set, which is considered as embedded set and S_e can be described as the domain of fuzzy sets. In type 2 fuzzy set theory, the symmetric set is formulated as follows,

 https://doi.org/10.37896/pd91.4/91445

$$\tilde{\mu}^{\tilde{A}}(m+x) = \tilde{\mu}^{\tilde{A}}(x)(m-x) \quad (14)$$

After that, input and output of the system are fuzzified. The input image is A and the image from database is B. The similarity distance $L1, L2, L3$ among A and I are three inputs of the human activity recognition. Three types of fuzzy variables are considered to enable the optimal human activity recognition such as "Similar", "Very Similar" and "not similar" respectively. These fuzzy variables are utilized to describe the feature differences and the output of the system. From the description, the similar action of the images and characterize. In this proposed classifier, the trapezoidal MF is utilized which formulated as follows,

$$\tilde{\mu}^{\tilde{A}}(x) = \begin{cases} \dfrac{(x+b)}{b-d}, & if -b \leq x \leq -d \\ 1, & if -d \leq x \leq d \\ \dfrac{b-x}{b-d}, & if\ d \leq x \leq -b \\ 0, & otherwise \end{cases} \quad (15)$$

Based on the membership functions, human activity is recognized with related feature differences of images. The fuzzy descriptions are utilized to analysis the feature distance enabled by the rules of the fuzzy inference system. The fuzzy interference system is built by fuzzy IF-THEN rules which presents a degree of absence or presence of interaction or association among the elements of two or more sets. The fuzzy rules can be designed explicitly through the expert itself as is the case in the system. Additionally, Mandeni fuzzy interface method is utilized. In the proposed method, human activity recognition is achieved with the help of rules. the different objects have similar feature variance scopes, the rule is detected the human activity recognition. After that, the fuzzy values are defuzzifier to create a crisp value for the output variable [22]. The proposed type 2 fuzzy system, is proceeding with two steps such as type reduction and defuzzification. Different types of methods are available to type reduction in fuzzy sets, in this method, a center of sets are utilized which formulated as follows,

$$Y_{cos}(X) = \left[C_l(\tilde{A}) C_r C_l(\tilde{A})\right] \quad (16)$$

$$C_l(\tilde{A}) = \frac{\sum_{i=1}^{M} \mu_l^i C_l^i}{\sum_{i=1}^{M} \mu_l^i} \quad (17)$$

$$C_r(\tilde{A}) = \frac{\sum_{i=1}^{M} \mu_r^i C_r^i}{\sum_{i=1}^{M} \mu_r^i} \quad (18)$$

Where M can be represented as several rules. From the consideration of type reduction, the interval set is computed. After that, this interval set is defuzzified in type 2 fuzzy set which formulated as follows,

$$y = \frac{C_l(\tilde{A}) + C_r(\tilde{A})}{2} \quad (19)$$

At last, the human activity is recognized with type 2 fuzzy classifier. The test image is sent to the proposed classifier which computes the similarity index through the rules with the consideration of the feature set. The possible similarity is compared with the image's features. With the help of features and proposed classifier, human activity recognition is achieved. In the fuzzy classifier, ROA algorithm is used in selecting membership function variables. A detailed ROA description is presented below in this section.

(ii) Rider Optimization Algorithm

ROA algorithm is utilized to optimize the fuzzy membership function which empowers the recognition the human activity from the images. The general description of the ROA is presented in this The ROA algorithm is used to improve the ambiguous membership

 https://doi.org/10.37896/pd91.4/91445

function, which helps to recognize human activity from images. A general description of the ROA is provided in this section. Rider optimization is developed based on rider behaviours. The ROA consists of two riding punches that travel to the common objective space for the winner of the race. In calculation, all four meetings are considered, usually the size of riders at each meeting is selected similarly from full number of riders [23]. Side Step Rider, Devotee, Overtaker and Occupier are the 4 groups of riders. These four meetings followed a different strategy to achieve the goal introduced as follows,

- ❖ Side ride thinking process comes to the last part except the policy course.
- ❖ The follower follows the main riding devotion of the centre of devotion.
- ❖ The boss can follow his situation to get to the last area identified with the main ride stop area.
- ❖ The Extreme Occupier will come to the last part thinking of the riding situation too fast.

The pseudo-code of ROA is illustrated in Figure 2. The step by step process is explained below;

Algorithm 1: Pseudocode of ROA for membership function variable selection

Input: Random membership function variables.

Output: Optimal membership function variables.

Begin

Initialize the population with random membership function variables

Initialize the rider parameters, brake, accelerator, gear, steering angle

Find success rate

Evaluate the fitness function

While t< off time

For i=1 to R

Update the position of bypass rider

Update the position of bypass rider

Update the position of bypass rider

Rank the riders

Select the maximum leading rider

Update the parameters

Save the results of optimal membership function parameters.

Step 1: Initialization: The ROA is created by four-rider clusters which are denoted as G whose locations are generalized randomly. The generalization process can be mathematically formulated as follows,

$$X^t = \{X^t(i,j)\}; \quad 1 \le i \le R; 1 \le j \le Q \qquad (20)$$

Where, $X^t(i,j)$ can be represented as a location of i^{th} rider at a time instant t, the dimension of number of coordinates is denoted by G,number of riders and Q is denoted by R which is equivalent to G. The number of riders computed, is related based on the number of riders in every group are formulated as follows,

$$R = B + F + O + A \qquad (21)$$

Where, A can be represented as several attackers, O can be represented as several over-takers, F can be represented as several followers and B can be represented as several bypass riders. The relation among the aforesaid parameters is computed below,

$$B = F = O = A = \frac{R}{4} \qquad (22)$$

Related on the above conditions, the rider locations are computed. The attacker position, overtaker position, follower position, and bypass rider positions are presented in the ranges $\left[X^{\frac{3R}{4}+1}, X^R\right]$ $\left[X^{\frac{R}{2}+1}, X^{3R/4}\right]$, $\left[X^{\frac{3R}{4}+1}, X^{R/2}\right]$ and $\left[X^1, X^{R/4}\right]$ respectively. Once completed the group initialization, the rider value is initialized like brake, accelerator, gear, in addition steering. The steering angle (T) of the rider vehicle at the time is presented as follows,

$$T^t = \{T^t_{i,j}\}: \; 1 \le i \le R; \; 1 \le j \le Q \qquad (23)$$

Where, $T^t_{i,j}$ is represented as the steering angle of i^{th} rider vehicle. The starting point of steering angle at the starting time which formulated as follows,

$$T^t_{i,j} = \begin{cases} \theta^i & if\ j = 1 \\ T^{i,j-1} + \varphi & if\ j \ne 1 \& T^{i,j-1} + \varphi \le 360 \\ T^{i,j-1} + \varphi - 360; & otherwise \end{cases} \qquad (24)$$

Where, φ can be described as coordinate angle, θ^i can be represented as the position angle of the i^{th} rider vehicle. The position angle of i^{th} rider can be computed based on maximum angle of 360^0 and the number of riders, which is formulated as follows,

$$\theta^i = i * \frac{360^0}{R} \qquad (25)$$

Where R can be represented as several riders. The coordinate angle utilized to compute the steering angle is formulated as,

$$\varphi = \frac{360}{Q} \qquad (26)$$

Step 2: Finding the success rate and fitness function evaluation: The parameters and rider groups are generalized, the success rate of every ride can be calculated. The success rate can be mathematically formulated related to the distance that can be computed based on the below equation,

$$r^i = \frac{1}{\|X^i - L^T\|} \qquad (27)$$

 https://doi.org/10.37896/pd91.4/91445

Where, L^T can be represented as final location position, X^i can be described as the position of i^{th} rider. To empower success rate, the distance of final location is reduced. Hence, the reciprocal of distance calculation provides the success rate of rider. The ROA is utilized to select optimal membership functions of the type 2 neuro-fuzzy classifier. The fitness evaluation of the ROA is formulated as follows,

$$FF = Min\ (M^E) \tag{28}$$

$$M^E = \frac{M^r}{T^r} \tag{29}$$

Where, M^E is described as classification error, M^r as classification records and T^r is described as total number of records. This fitness evaluation is proceeding by selecting the optimal membership function values.

Step 3: Position update: The location of each set of rider is updated periodically to compute the leading rider in addition, which is considered as the winner [24]. The rider positions are updated based on the characteristics. The bypass rider position updating process is presented as follows,

$$X_{t+1}^{i,j}(i,j) = \delta\left[X^t(\eta,j) * \beta(j) + X^t(\xi,j) * [1-\beta(j)]\right] \tag{30}$$

Where, β can be represented as a random value, this value takes among 0 and 1 but size $1 \times Q$, ξ can be described as random value 1 and R, η can be described as a random value which ranges from 1 to R and δ can be represented as a random number and parameter ranges from 0 and 1 respectively. The follower position is also updated based on the leading rider because it reaches target quickly and effectively. In the ROA algorithm, the follower position is updated based on the below equation,

$$X_{t+1}^{F}(i,k) = X^L(L,K) + \left[cos\left(T_{i,k}^t\right) * X^L(L,k) * D_i^t\right] \tag{31}$$

Where, D_i^t can be described as distance traveled by the i^{th} rider, $T_{i,k}^t$ can be described as the steering angle of i^{th} rider in the k^{th} coordinate, L can be described as leading rider index, X^L can be described as leading rider position and K can be described as coordinate selectors respectively. The travel distance is computed by multiplying the rider velocity with the rate of off time. The position update of overtaker is computed based on the below equation,

$$X_{t+1}^{F}(i,k) = X^t(i,K) + [D_t^I(i) * X^L(L,K)] \tag{32}$$

Where, $D_t^I(i)$ can be described as the direction indicator of i^{th} rider at the time t, position of i^{th} a rider can be denoted by $X^t(i,K)$ respectively. The direction indicator calculates, direction of relative success rate which formulated follows,

$$D_i^I(i) = \left[\frac{2}{1-log(S_t^R(i))}\right] - 1 \tag{33}$$

Where, $D_i^I(i)$ can be described as indicator of direction which presented in the range [-1,1], S_t^R can be described as interval time [0,1] and $S_t^R(i)$ can be described as the success rate of i^{th} rider respectively. The updating process of the attacker is mathematically formulated as follows,

$$X_{t+1}^{A}(i,k) = X^L(L,j) + \left[cos\left(T_{i,j}^t\right) * X^L(L,j) * D_i^t\right] \tag{34}$$

Where, D_i^t can be described as distance calculation of i^{th} rider, $T_{i,k}^t$ can be described as steering angle of i^{th} rider in k^{th} coordinate and $X^L(L,j)$ as the position of leading rider respectively.

https://doi.org/10.37896/pd91.4/91445

ISSN: 0369-8963

Step 4: Calculate success rate: **T**he achievement pace of every rider can be Figured, on fruition of position update measure. The situation of rider can be refreshed with the main race up until now. The pioneer position is supplanted with another rider position to such an extent that the achievement pace of the new rider is in the greatest reach [25]. Consequently, the rider, who gave the greatest achievement rate which considered as a main rider.

Step 5: Rider parameter update: The rider parameters are required to compute an efficient optimal solution. The rider value updating proceeds with consideration of additional parameters such as the activity counter. The steering angle and gear are examined related to the activity count which can be updated following the success rate. The activity count is mathematically formulated as follows,

$$A_c^{t+1}(i) = \begin{cases} 1; & if r^{t+1}(i) > r^t(i) \\ 0; & otherwise \end{cases} \quad (35)$$

Similarly, the steering angle can be updated related on the activity counter which formulated as follows,

$$T_{i,j}^{t+1} = \begin{cases} T_{i+1,j}^t; & if A_c^{t+1}(i) = 1 \\ T_{i-1,j}^t; & if A_c^{t+1}(i) = 0 \end{cases} \quad (36)$$

The vehicle gear is upgraded based on maximum value of the gear a rider and activity counter can take as follows,

$$e_I^{T+1} = \begin{cases} e_I^T + 1; & if A_c^{t+1}(i) = 1 \;\&\; e_i^T \neq |e| \\ e_I^T - 1; & if A_c^{t+1}(i) = 0 \;\&\; e_i^T \neq 0 \\ e_I^T; & otherwise \end{cases} \quad (37)$$

The vehicle accelerator can be updated based on the vehicle gear which formulated as follows,

$$e_I^{T+1} = \frac{e_I^{T+1}}{|e|} \quad (38)$$

Where, $|e|$ can be described as a number of gears.

The brake of the vehicle can be updated based on the below equation,

$$k_i^{T+1} = \left[1 - \frac{e_i^{T+1}}{|e|}\right] \quad (39)$$

Step 6: Riding off time: The above process is recurrent continuously, till the time moves the off time, withing that leading rider can be computed. In the final process, the leading rider can be considered a winner. Thus, with the help of proposed classifier, human activity is recognized efficiently.

Results and discussion

The experimental results of proposed approach are analysed in this section. The proposed approach is analysed based on the efficiency of specificity, precision, accuracy, recall, sensitivity and F_Measure respectively. Comparison is made between proposed approach performance and the existing methods like ANN, DNN, and FLC-PSO respectively. The validation is performed for projected method using the databases. In the proposed methodology, the KTH dataset is considered to authenticate the performance of the projected methodology. The KTH dataset consists of six classes like boxing, hand waving, clapping, running, walking and jogging. The 25 person activity are collected with six classes. This database contains 2391 sequences that are taken by utilizing the fixed camera with a 25fps frame rate and down samples to pixel resolution of $160 * 120$ which is an average length of four seconds. The KTH database is collected from [26]. In table 1, the simulation parameter is presented. The KTH dataset and human

 https://doi.org/10.37896/pd91.4/91445

activity recognition outputs are illustrated in Figure 2.

Table 1: Paramter used in this work

S. No	Method	Description	Values
1	Proposed method	FIS type	Mamdani
2		Defuzzification method	Centroid
3		Population size	200
4		Number of iterations	100
5		Random vectors	0,1
6		Coefficient vector	0,2
7	ANN	Architecture	Normal feedforward MLP
8		Hidden layer	1
9		Training algorithm	Backpropagation
10		Number of an input layer	3
11		Number of hidden layers	7
12		Number of the output layer	1
13		Transfer function	Sigmoid

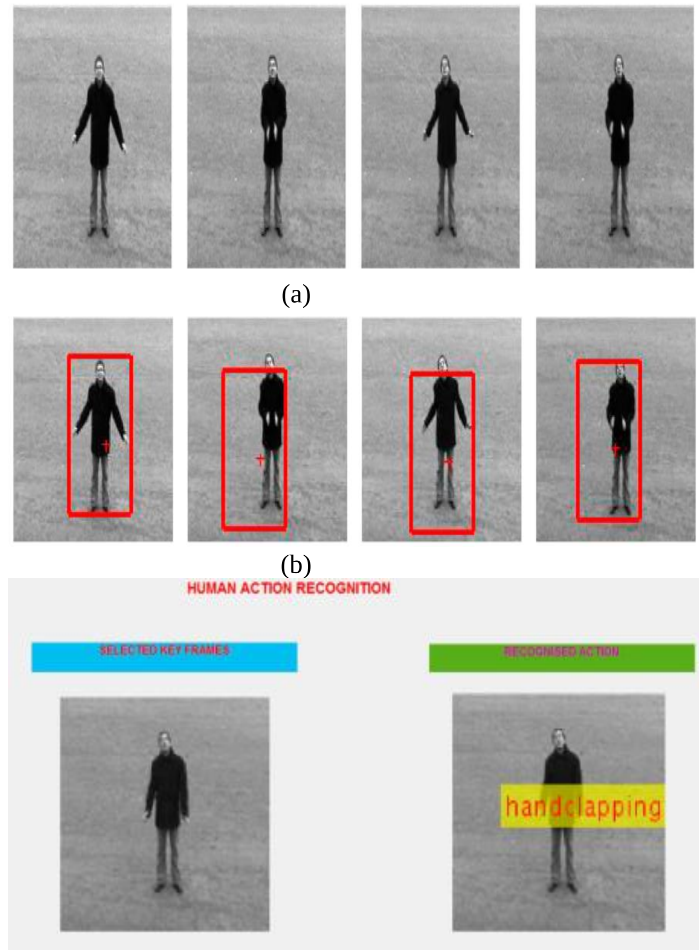

Figure 2. Analysis of (a) databases, (b) feature extraction, and (c) classification phase

 https://doi.org/10.37896/pd91.4/91445

The dataset can be used to diagnosis human activity from the images. The proposed method is validated with the help of statistical measurements. The measures are given below;

Accuracy: The ratio of number of true patterns to the sum of all patterns is known as accuracy. The given formula is described for accuracy in equation 40,

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \quad (40)$$

Sensitivity: The two essential proportions of evaluating symptomatic precision of a test are affectability and explicitness. The affectability of a symptomatic test measures its capacity to accurately recognize subjects with human movement. It is the extent of genuine positives that are accurately recognized by the test, given by the condition 41,

$$sensitivity = \frac{TP}{TP + FN}$$

(41)

Specificity: The explicitness is the capacity of a test to effectively recognize subjects without condition. It is the extent of genuine negatives that are accurately recognized by the test. Tee formula for sensitivity is described in equation (42)

$$specificity = \frac{TN}{FP + TN} \quad (42)$$

Recall: Recall is the function of correctly classified error (TPs) and incorrectly classified error (FNs). The recall calculation is given in equation 1.

$$Recall = \frac{TP}{TP + FN} \quad (43)$$

F_Score: F_Score measure calculated based on precision and recall measure. It is defined by Equation (44) as

$$F_{Score} = 2.\frac{Precision \times recall}{Precision + recall} \quad (44)$$

The projected method is classified the human activity with six activities which are presented in Table 2.

Table 2: Confusion matrix with proposed classifier

	Clapping	Running	Hand waving	Walking	Jogging	Boxing
Clapping	97	1	0	0	0	1
Running	0	95	0	0	1	0
Hand waving	0	0	97	1	0	1
Walking	0	1	1	98	0	0
Jogging	3	3	0	1	99	1
Boxing	0	0	2	0	0	97

The comparison analysis is essential to validate the proposed methodology. The suggested approach is compared using conventional methods, such as FLC-PSO, DNN, and ANN respectively. The comparison of specificity measure is demonstrated in Figure 3. The specificity of suggested approach is 0.95 at the first class. Similarly, the existing methods attains the specificity values of 0.7, 0.68, and 0.71 respectively. From the contrast analysis of specificity measures, the suggested approach has been attained the best results in HAR. The comparison analysis of accuracy measure is presented in Figure 4. The accuracy of projected method is 0.97 at first class. Similarly, the existing methods attained accuracy values are 0.84, 0.82, and 0.85 respectively. The comparison analysis of sensitivity measure is presented in Figure 5.

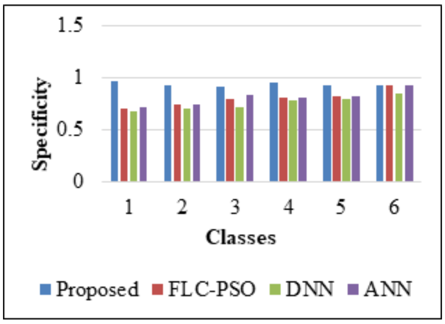

Figure 3: Specificity Analysis

 https://doi.org/10.37896/pd91.4/91445

Figure 4: Accuracy Analysis

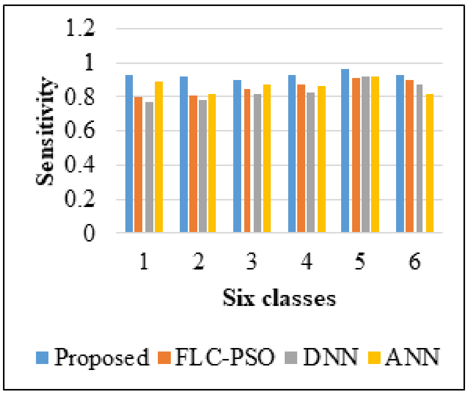

Figure 5: Precision Analysis

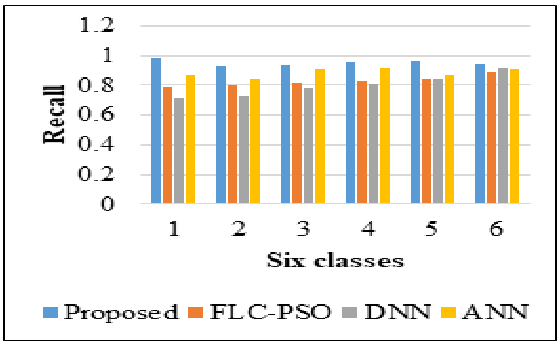

Figure 6: Recall Analysis

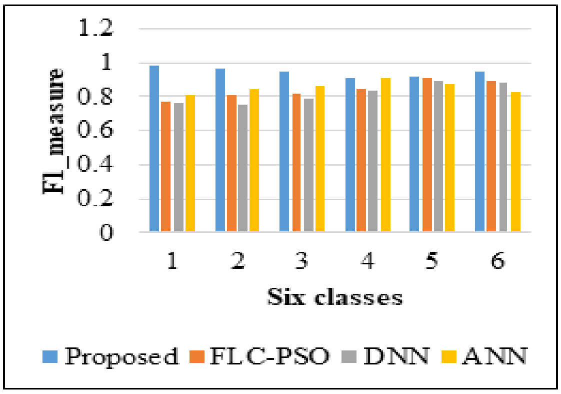

Figure 6: F1_Measure Analysis

The sensitivity of the suggested approach is 0.93 at the first class. Similarly, the existing methods are attained sensitivity values of 0.8, 0.77, and 0.89 respectively. From contrast analysis of sensitivity measures, the projected method has been achieved optimal results in human activity recognition. A comparative analysis of the recall measurement is given in Figure 6. The recall of the planned system is 0.98 in the first class. Similarly, the current methods achieve memory values of 0.79, 0.72 and 0.87, respectively. From diverse analysis of recall measures, the proposed method has reached best results in human activity recognition. A comparative analysis of F_Measure is demonstrated in Figure 7. The F_Measure of proposed work is 0.98 in the first class. Similarly, the current methods achieve F_Measure values of 0.77, 0.76 and 0.81, respectively. From the different analysis of F_Measure, the planned method has attained optimal results in human activity recognition.

Conclusion

In the proposed work, OHAR is established to identify human activity from images. The proposed method is a combination of type 2 fuzzy classifier and ROA algorithm. The type 2 fuzzy classifier process is enhanced with utilization of ROA and it is utilized to classify the human actions from the databases. In this approach, initially, the KTH databases are collected from the internet. The collected video databases are converted into frames. After that, in pre-processing stage all unnecessary features and noise are completely removed from images. From the pre-processing frames, essential frames are selected by using SSIM index values. The selected keyframes are considered for the feature extraction phase. Three different feature extraction methods are considered in the proposed method such as grid shape features, STI points, and coverage factor. Finally, extracted features are sent to the classifier for human activity recognition. The suggested approach has been validated using different metrics. The suggested approach has been contrasted with the

https://doi.org/10.37896/pd91.4/91445

conventional methods such as ANN, DNN, FLC-PSO respectively. From the contrast analysis, the projected technique has achieved the optimal solutions in terms of accuracy is 97%.

References

[1] Henry Friday Nweke, Ying Wah Teh, Mohammed Ali Al-Garadi, and Uzoma Rita Alo. (2018) - 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, 233-261.

[2] Elnaz Soleimani, and Ehsan Nazerfard. (2021) - Cross-subject transfer learning in human activity recognition systems using generative adversarial networks. *Neurocomputing* 426 26-34.

[3] Raffaele Gravina, Congcong Ma, Pasquale Pace, Gianluca Aloi, Wilma Russo, Wenfeng Li, and Giancarlo Fortino. (2017) - Cloud-based Activity-aaService cyber–physical framework for human activity monitoring in mobility. *Future Generation Computer Systems*, 75, 158-171.

[4] Lukas Koping, Kimiaki Shirahama, and Marcin Grzegorzek. (2018) - A general framework for sensor-based human activity recognition. *Computers in biology and medicine*, 95, 248-260.

[5] K.S.Gayathri, K. S. Easwarakumar, and Susan Elias. (2017) - Probabilistic ontology-based activity recognition in smart homes using Markov Logic Network. *Knowledge-Based Systems*, 121, 173-184.

[6] Haihua Liu, Na Shu, Qiling Tang, and Wensheng Zhang. (2017) - Computational model based on neural network of visual cortex for human action recognition. *IEEE transactions on neural networks and learning systems*, 29 (5), 1427-1440.

[7] Majd Latah, (2017) - Human action recognition using support vector machines and 3D convolutional neural networks. *International Journal of Advances in Intelligent Informatics*, 3 (1) 47-55.

[8] Lei Wang, Yangyang Xu, Jun Cheng, Haiying Xia, Jianqin Yin, and Jiaji Wu. (2018) - Human action recognition by learning spatio-temporal features with deep neural networks. *IEEE access*, 6, 17913-17922.

[9] Aouaidjia Kamel, Bin Sheng, Po Yang, Ping Li, Ruimin Shen and David Dagan Feng. (2018) - Deep convolutional neural networks for human action recognition using depth maps and postures. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 49 (9), 1806-1819.

[10] Alaa Tharwat, Hani Mahdi, Mohamed Elhoseny, and Aboul Ella Hassanien. (2018) - Recognizing human activity in mobile crowdsensing environment using optimized k-NN algorithm. *Expert Systems with Applications*, 107, 32-44.

[11] Mohammed Mehedi Hassan, Md Zia Uddin, Amr Mohamed, and Ahmad Almogren. (2018) - A robust human activity recognition system using smartphone sensors and deep learning. *Future Generation Computer Systems*, 81, 307-313.

[12] Andrey Ignatov, (2018) - Real-time human activity recognition from accelerometer data using Convolutional Neural Networks. *Applied Soft Computing*, 62, 915-922.

[13] Abdulmajid Murad, and Jae-Young Pyun. (2017) - Deep recurrent neural networks for human activity recognition. *Sensors*, 17 (11) 2556.

[14] Xiaokang Zhou, Wei Liang, I. Kevin, Kai Wang, Hao Wang, Laurence T. Yang, and Qun Jin. (2020) - Deep-learning-enhanced human activity recognition for Internet of healthcare things. *IEEE Internet of Things Journal*, 7 (7), 6429-6438.

[15] Nafiul Rashid, Berken Utku Demirel, and Mohammad Abdullah Al Faruque. (2021) - AHAR: Adaptive CNN for Energy-efficient Human Activity Recognition in Low-power Edge Devices. arXiv preprint arXiv: 2102.01875

[16] Ken Cai, Rongqian Yang, Huazhou Chen, Lihua Li, Jing Zhou, Shanxing Ou, and Feng Liu. (2017) - A framework combining window width-level adjustment and Gaussian filter-based multi-resolution for automatic whole heart segmentation. *Neurocomputing*, 220 138-150.

[17] Yiming Tang, Fuji Ren, and Witold Pedrycz. (2020) - Fuzzy C-means clustering through SSIM and patch for image segmentation. *Applied Soft Computing*, 87, 105928.

[18] Shouji Du, Yunsheng Zhang, Zhengrong Zou, Shenghua Xu, Xue He, and Siyang Chen. (2017) - Automatic building extraction from LiDAR data fusion of point and grid-based features. *ISPRS journal of photogrammetry and remote sensing*, 130, 294-307.

[19] Paul T Sheeba and S. Murugan. (2018) - Hybrid features-enabled dragon deep belief neural network for activity recognition. *The Imaging Science Journal* 66 (6), 355-371.

[20] Gellért Máttyus, Wenjie Luo, and Raquel Urtasun. (2017) - Deeproadmapper: Extracting road topology from aerial images. *In Proceedings of the IEEE International Conference on Computer Vision*, 3438-3446.

[21] Saad M Darwish, and Raad A. Ali. (2015) - Observations on Using Type-2 Fuzzy Logic for Reducing Semantic Gap in Content–Based Image Retrieval System. *International Journal of Computer Theory and Engineering*, 7 (1), 12015.

[22] Xiaomei Li, and Yang Chen. (2021) - Discrete Non-iterative Centroid Type-Reduction Algorithms on General Type-2 Fuzzy Logic Systems. *International Journal of Fuzzy Systems*, 1-12.

[23] D.Binu, and B. S. Kariyappa. RideNN: (2018) - A new rider optimization algorithm-based neural network for fault diagnosis in analog circuits. *IEEE Transactions on Instrumentation and Measurement*, 68 (1), 2-26.

[24] Guohu Wang, Yongliang Yuan, and Wenwen Guo. (2019) - An improved rider optimization algorithm for solving engineering optimization problems. *IEEE Access*, 7, 80570-80576.

[25] Ambaji S Jadhav, Pushpa B. Patil, and Sunil Biradar. (2020) - Optimal feature selection-based diabetic retinopathy detection using improved rider optimization algorithm enabled with deep learning. *Evolutionary Intelligence*, 1-18.

[26] https://www.csc.kth.se/cvap/actions/