JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 2, 2022, p. 2392 - 2403
https://publishoa.com
ISSN: 1309-3452

Hybrid Adaptive Deep Learning Based Human Action Recognition

Siva Nagi Reddy Kalli[1], V. Sidda Reddy[2], Rangaiah Leberu[3]

[1]Professor, Department of ECE, Sridevi Women's Engineering College, Hyderabad, T.S, India

[2]Professor, Department of IT, CVR College of Engineering, Hyderabad, T.S, India

[3]Professor, Department of ECE, Rajarajeswari College of Engineering, Bengaluru, KA, India

[1]sivanagireddykalli@gmail.com, [2]siddareddy.v@gmail.com, gmail.com,[3]rleburu@gmail.com

ABSTRACT

Many Human Action Recognition (HAR) systems are used to recognize everyday human behaviors such as standing, walking, sitting, running, cooking, and exercising. The existing human action recognition technology is not perfect. It has poor adaptability and low accuracy. Deep learning is used to improve action recognition accuracy and motion detection. Deep neural network models can be difficult to train, such as gradient disappearance, gradient explosion, and over-fitting. This paper will help to reduce the difficulty in training the above deep neural model parameter initialization method based upon the activation function for the multi-layer networks. It also addresses the problem of model training deep neural networks. The presented hybrid algorithm Improved Spotted Hyena algorithm and Seagull Optimization Algorithm, (ISHO-SOA), is used to solve optimization algorithms and obtain minimum optimum values. The segmented image is further classified using different inputs by Adaptive Deep Convolution Networks (ADCNN), with feature maps. The classification method provides performance metrics such as Accuracy and Precision, F1 measure, along with Recall that can be used to recognize the most accurate results and the best classifiers. This algorithm is highly efficient in improving classification accuracy and solving global optimization problems. Experimental results show that the proposed algorithm has a 98.21% higher accuracy than other meta-heuristic algorithms.

Keywords: Action recognition; optimization; Neural Network; Classification and Accuracy.

1. INTRODUCTION

Human Action Recognition (HAR) is an important topic in computer vision, machine learning, and artificial intelligence. This study aims to recognize human behavior in video and images. It has wide applications in security monitoring and disability surveillance, multimedia content understanding, human computer interaction, virtual reality, and human-computer interaction. HAR is able to recognize basic and complex behaviors in real-world situations. It is also difficult when there are more appliances, chart inspection, human CPU Communication, autonomous driving, and activity. There are many real-world uses for HAR, including healthcare, gaming, interior navigation, tactical military applications and personal fitness. HAR is especially useful in industries that still require manual labor [1-3].

There are three types of action feature extraction algorithms. They can be classified as: traditional action recognition techniques [4]-[6], deep learning action recognition methods [7]-[9], hybrid feature extraction method for action recognition [10]. This hybrid feature-based method for action recognition is a fusion of artificial and deep learning feature extraction. It is essentially a hybrid feature extraction that uses a combination artificial design features and deeplearning strategies.

This study proposed adaptive deep learning (ADL), which is a more advanced form of deep learning. This study uses the three ADL processes to detect human activity in real-time. The first phase involves pre-processing to extract human activity using the Top-Hot filter. The second phase uses Otsu technique to segment the human activity. In the third phase, features are used to extract the human activity using hybrid Improved Spotted Hyena Optimization - based on Seagull Optimization Approach. ADCNN (Classification Method Adaptive Deep Convolution Neural Network) is used to categorize human behavior and calculate better outcomes.

Below topics are described through: Section 2 gives a quick review of related works employing deep learning for HAR. The proposed models, performance findings, and comparisons are discussed in Sections 3 and 4 correspondingly. Lastly, the outcomes are explained in Section 5.

2. LITERATURE REVIEW

Recent years have seen a lot of research in the area of human action detection using Deep learning technology. This is because it has proven to be a great success. Deep learning techniques have been used by many scholars to recognize action. This section will examine a variety of techniques that can be used to recognize human activity.

Hassan and colleagues [11] presented the smartphone sensor data to establish a reliable human activity detection system. For recognition and training purposes, the robust-based Deep Belief Network has been integrated. The twelve different physical activities can be identified with greater accuracy and higher recognition rates than other methods. Author failed to highlight more precise characteristics in real-time situations to recognize more complex activities. Murad et. al. [12] used Deep Recurrent Neural Networks to detect extended-choice dependence for differential input data. Three unique LSTM-based DRNN Frameworks were presented for HAR challenges. DRNN technology was used to detect temporal dynamics among input data in activity sequences. A DRNN that can be used on low-power devices was developed to ensure successful implementation. Refer to [13] for an example. This technique uses stacking and folding to extract the Invariant spatiotemporal characteristics directly from video Data. The function can then be used to produce a good action recognition effect. Refer to [14] for a method of detection based on convolutional neural networks architecture. This method can learn latent features from images. Reference [15] uses 3D convolutional neural networks, which detect human behavior in video without relying on design features and produces overwhelming motion recognition results. Deep learning has become very complex in 3D video. Many operations are now extremely complicated. The computational complexity of convolution operations increases exponentially. The convolution operation, its development and application based upon deep learning have made tremendous progress. However, there are still problems with falling off into local optima and slow train. The training of deep neural network models isn't convex. Non-convex optimization and solution Optimization algorithms (e.g. Stochastic gradient descent depends on initialization of model parameters [16, 17]. It is an important step in the model training process. It directly impacts the distribution of each hidden level and the initial stage in the neural network model. It is possible to ensure that each hidden layer node follows the same gradient propagation distribution. This can increase the speed of non-convex optimization and prevent it from falling into a local optimal process. Incorrect parameter model initialization could lead to gradient disappearance and gradient explosion if the gradient is propagated accordingly. Large differences in the distribution of hidden layers can lead to slow convergence, making it difficult to find the local optimal solution [18]. Julieta et al. [19] Using recurrent neural network to focus on human actions, they targeted time-varying representations to accomplish tasks such as short-term prediction or long-term human movement synthesis. They also compared these methods to other state-of the-art methods Recurrent Neural Net Method. Chao li et al. [20] They propose a framework that is an end to-end CNN feature-learning framework. They use a layered approach where Co-occurrence features are learned with different contextual information. They first encode Point-level information and render semantic representations in time and space. They propose a global space approach to learning great common information. Maosen Li et al. Two scalable graphs were proposed by Maosen Li et al. [21]. They capture the relationships between body parts and joints. They employ the spatial features of multi-scale CNN. Common structure and scale diagrams Capture motion or physical limitations Comparison of performance Deep learning models outperform machine-learning based algorithms. HAR data is time series data which contains the required spatial or temporal information. This data can be used to build powerful models capable of learning from human activities. This work develops a HAR model that combines spatiotemporal and

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 2, 2022, p. 2392 - 2403
https://publishoa.com
ISSN: 1309-3452

feature extraction to provide effective HAR. Xu et al. [22] Combine a traditional RNN and a wearable sensor-based, HAR-aligned initial neuron (INN) model in order to create InnoHAR. The INN architecture is made up of several layers of convolutional layers, which run parallel to the pooling layer and form the initial layers. Multiple datasets have been used to test the INN architecture. It performed better than the Deep Convolutional LSM model. This framework has a disadvantage: the INN is slow to initialize and can be computationally complex. Small changes may need expensive retraining. Damodaran N. et al. Damodaran N. et al. Two classification algorithms were introduced: support vector machines, (SVM), and long-term memory (LSTM). Wavelet analysis is used to preprocess and extract features. They can recognise walking, sitting, standing and running activities.

3. HUMAN ACTIVITY RECOGNITION USING ISHO-SOA

The current behavior of the human body is to infer HAR through a series of analyses and observations of its surroundings. HAR is currently valuable in different fields, wherever it is expensive. This is to ensure an individual's efficiency and lifestyle. This work uses deep knowledge classification to create HAR. It includes four stages: segmentation, feature selection and classification. The first step is to collect the images from the database. Next, the images go through the pre-processing phase. Top-hat filtering, which is used to enhance the image's contrast and noise removal, is part of the preprocessing phase. The image is then segmented using an optimized Otsu technique. A novel contribution to feature selection is made by a hybrid Improved Spotted Hyena Optimization algorithm and Seagull Optimization Algorithm. This was done in order to resolve global optimization problems. We also achieved the optimal weight for the ACDNN process. The image then goes to the classification process. Here the Adaptive Deep Consvolution Neural Network is used to extract features from the segmented images using N-feature maps. Below is the pseudo code for ACDNN. Our proposed method improves classification accuracy by combining it with the hybrid algorithm process. Figure 1 shows the proposed methodology diagram..

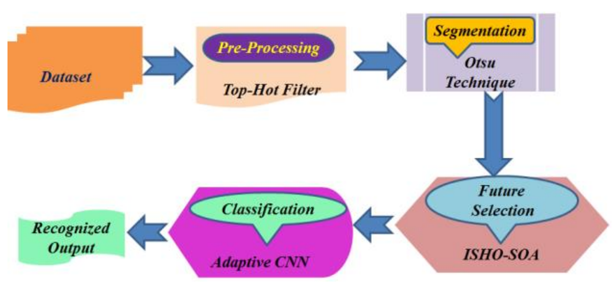

Fig1: Proposed architecture of HAR

3.1 Pre-Processing

Pre-processing step: First improve the original image by Bicubic interpolation. It transforms a superior-resolution image into a lower-resolution one. Gaussian, top hat and bottom hat filters, as well as mean and guided filters, are all used to remove noise from the image. We will then be able to process the image. The following explains the pre-processing equation:

Step 1: Perform Bicubic Interpolation to form I_{r*t}

$$I(i,j) \leftarrow a_{ij}x^iy^j \qquad (1)$$

Where

$I(i,j) -$ *image at location*

$a_{ij} -$ *low resolution image*

$x^iy^j -$ *high resolution image*

Step 2*:* Improving image utilizing top hat – bottom hat

$$I_{Tho} \leftarrow I_{guassian} - [I_{guassian}\theta\ strel) \oplus strel] \qquad (2)$$

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 2, 2022, p. 2392 - 2403
https://publishoa.com
ISSN: 1309-3452

$$I_{Bhc} \leftarrow [(I_{guassian} \oplus strel)\theta strel] - I_{guassian} \quad (3)$$
$$I_{Th-Bh} \leftarrow I_{guassian} + I_{Tho} - I_{Bhc} \quad (4)$$
$$I_{enhance} \leftarrow I - I_{guassian} + I_{Th} - Bh \quad (5)$$
$$I \leftarrow I_{enhance} \quad (6)$$

Where

$Strel - strcturing\ element$

$I_{guassian} \gg Gaussian\ filter\ (I) perform\ Gaussian\ filter\ on\ I$

$I_{Tho} -\ image\ after\ performing\ top\ hat\ opening$

$I_{Bhc} - image\ after\ performing\ bottom\ hot\ closing$

$I_{Th-Bh} - image\ after\ performing\ subtract\ (Th - Bh)$

Step 3: Calculate sharpening through deduct a weighted Laplacian from the blurred image

$$I_{laplacian} \leftarrow I_{grad\ X} + I_{grad\ Y} \quad (7)$$
$$I_{sharpen} \leftarrow I_{enhance} - C * I_{laplacian} \quad (8)$$

$C - is\ multiplicative\ coefficient$

$I_{grad\ X} - image\ gradient\ in\ X - direction$

$I_{grad\ Y} - image\ gradient\ in\ mage\ gradient\ in\ Y - direction$

3.2 Image Segmentation utilizing Optimized Otsu's Model

Consider the input picture evaluated like $I(X,Y)$ by grey point $0,1 \ldots \rho - 1$. where I is evaluated as$F = \{f_0, f_1 \ldots \ldots f_\rho - 1\}$, where $f_0, f_1 \ldots \ldots f_\rho - 1$ characterize the frequency of every greyscale stage in the input image I.

Let $N = \sum_{i=1}^{\rho} f_i$ is the entire quantity of pixels in the image.

The probability of the ith greyscale intensity is calculated as

$$Probi = \frac{f_i}{N}; \Pr obi > 0, \qquad \sum \Pr obi = 1 \quad (9)$$

The segmentation model partitions the image into $T+1$ segments as $S_k = \{S_{0,}S_1 \ldots\ldots S_k\}$. Here T is chose as of the collection thresholds $T = \{t[r], t[r+1], \ldots\ldots t[r+1] - 1\}$ every subdivision part is the position of greyscale points as of T. The biased probability along with biased gray-scale levels η for every subdivision S be able to designed through eqn(10) and (11).

$$ws_k = \sum_{i\epsilon S_k} \Pr obi \quad (10)$$
$$\eta = \sum_{i\epsilon S_k} \frac{\Pr obi.\max\{i, g*\log(i)\}}{ws_k}; \qquad g\epsilon GC \quad (11)$$

Weighted mean greyscale intensity $\mu W1$ and mean interclass variance σ_{VS}^2 between the segments of the whole image is given by

$$\mu W1 = \sum_{i=0}^{\rho} i * \Pr obi \quad (12)$$
$$\sigma_{VS}^2 = \sum_{i=0}^{\rho} ws_k.\eta^2 - \mu_{W1}^2 \quad (13)$$
$$(\sigma_{VS}^2) * = \sum \max\{\sigma_{VS}^2(S_i), \sigma_{VS}^2(S_j)\} \quad (14)$$

By maximizing the inter-class variance, optimal thresholds are definite toward every greyscale level of the segmented region and expressed in underneath eqn (15).

$$\{t*[1], t*[2] \ldots t*[n]\} = argmax\{\sigma_{VS}^2 * (t[1], t[2] \ldots . t[T])\} \quad (15)$$

In the segmentation stage, the aim of Otsu's method is to perform the automatic image thresholding for all thresholds with possible values. The pixels separate the image into the inter-class variation among center as well as surroundings image to approximate the noise intensity in the over-segmentation procedure.

3.3 Hybrid Improved Spotted Hyena Optimization with Seagull Algorithm (ISHO-SOA)

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 2, 2022, p. 2392 - 2403
https://publishoa.com
ISSN: 1309-3452

SHO is a metaheuristic algorithm and the main concept is the public connection among Spotted Hyenas as well as their joint performance. In spotted hyenas, they were lives in clan and their female members are clan. Though, male component go away their clan while they turn into adult along with join a new clan. They are skillful seeker along with major three extra hyena types that is lined, tanned also aardwolf). The main reason of SOA is able to change the direction of assault constantly with velocity through migration with it has tough comprehensive seeking capability, which mainly combines the spotted hyenas attacking prey method of the ISHO with the curved assault behaviors, extensively humanizing general seeking along with universal seeking capability of algorithm. The steps for SHO algorithms is explained as follows

3.3.1 Encircling the prey: Spotted hyenas are well-known among prey position also surround. Toward scientifically represent the group chain of command spotted hyenas, present greatest competitor explanation be there the aim of the prey otherwise intention which near toward the most favorable seeking gap doesn't identified a priori. Afterward better seeking candidate solution is defined the next seeking representative resolve towards revamp their place regarding better applicant explanation. The numerical model is representing through underneath eqn (16) and (17).

$$\overrightarrow{D_h} = \left|\overrightarrow{B}.\overrightarrow{P_p}(x) - \vec{P}(x)\right| \qquad (16)$$

$$\vec{P}(x+1) = \overrightarrow{P_p}(x) - \overrightarrow{E}.\overrightarrow{D_h} \qquad (17)$$

where $\overrightarrow{D_h}$ describe the space among the prey with spotted hyena, x signify near iteration, $\overrightarrow{B}$ as well as $\overrightarrow{E}$ co-efficient vectors, $\overrightarrow{P_p}$designate the prey location vector, $\vec{P}$is the dotted hyena location vector. Though$\left|\overrightarrow{B}.\overrightarrow{P_p}(x) - \vec{P}(x)\right|$ explicate complete value in addition to multiplication through vectors correspondingly. The vectors $\overrightarrow{B}$ with $\overrightarrow{E}$ were considered underneath eqn:

$$\vec{B} = 2..\overrightarrow{rd1} \qquad (18)$$

$$\vec{E} = 2..\vec{h}..\overrightarrow{rd_2} - \vec{h} \qquad (19)$$

$$\overrightarrow{h} = 5 - (Iteration * (5/\max_{iteration}) \qquad (20)$$

Iteration = 1,2, 3…, $Max_{iteration}$.For correctly matching the exploration as well as exploitation, here h is linearly downward as of 5 to 0 more the path of most numeral iterations $Max_{iteration}$ .The vectors $\overrightarrow{B}$ and $\overrightarrow{E}$ are calculated as follows: Additionally, this mechanism enhanced further development the iteration rate goes up. However, $\overrightarrow{rd_1}$, $\overrightarrow{rd_2}$are random vectors in [0, 1]. In this figure, the dotted hyena (A, B) could renovate its place towards the prey position $(A *, B *)$. change through importance of vector $\overrightarrow{B}$ with $\overrightarrow{E}$, here various numeral of positions was reaching concerning the present position. By using equation (11) and (12), a spotted hyena could renovate its location at random approximately the prey. Consequently, the similar theory is able to expand by n-dimensional seeking space.

3.3.2 Hunting: Spotted hyenas survive as well as hunt in grouping, relying under a system of true friends along with the ability to spot prey. Toward characterize the behavior of spotted hyenas randomly; we assume that the better seek representative, either is optimum, known the place of prey. The more seek representative group a cluster or group of true friends about the most excellent seek representative and save better explanation they've found far-off toward renovate their locations.

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 2, 2022, p. 2392 - 2403
https://publishoa.com
ISSN: 1309-3452

$$\overrightarrow{D_k} = \left|\overrightarrow{B.p_h}\right| - \overrightarrow{p_k} \tag{21}$$

$$\overrightarrow{p_k} = \overrightarrow{p_h} - \overrightarrow{E}.\overrightarrow{D} \tag{22}$$

$$\overrightarrow{C_h} = \left|\overrightarrow{p_k} + \overrightarrow{p_{h+1}} + ... + \overrightarrow{p_{K+N}}\right| \tag{23}$$

Where $\overrightarrow{ph}$ defines the first location of best dotted hyena, $\overrightarrow{pk}$ indicates location of other spotted hyenas. Here, N indicates the numerical spotted were computed as follows:

$$N = count_{nos}(\overrightarrow{p_h} + \overrightarrow{p_{h+1}} + \overrightarrow{p_{h+2}} .... + (\overrightarrow{p_h} + \overrightarrow{M} \tag{24}$$

here $\overrightarrow{M}$ is a random vector in [0.5, 1], defines the amount of explanation with count up all applicant explanation, subsequent to count by $\overrightarrow{M}$, relate toward better best explanation within particular seek space, and $\overrightarrow{C_h}$ is a group or N numerical of optimal explanation cluster.

3.3.3 Attacking prey: In the procedure of attacking prey, spiral motion behavior in the air of seagull with $X, Y and Z$ in planes is calculated by underneath eqn (25) to (28):

$$x' = r \times \cos(k) \tag{25}$$

$$y' = r \times \sin(k) \tag{26}$$

$$z' = r \times k \tag{27}$$

$$r = u \times e^{kv} \tag{28}$$

where r is radius of every rotate of coiled and k is formal number in the range [0, 2π], $u and v$ are constants to specify the coiled form, plus e is the support of the natural logarithm. The renovated location of every agent is computed utilizing eqn (20) – (23).

$$\overrightarrow{P_s}(x) = (\overrightarrow{D_s} \times x' \times y' \times z') + \overrightarrow{P_{bs}}(x) \tag{29}$$

here $\overrightarrow{P_s}(x)$ assign the better explanation along with renovate the location of new seek representatives.

3.3.4 Diversity mutation operation: The SHO algorithm goes into local optima, if the present optimal member is the local optima. This is also a natural attribute of new set intellectual optimization algorithms. To lessen the probability of premature convergence for the SHO algorithm, a diversity mutation operation is defined on the present optimal spotted hyena individuals. The steps are: suppose that an individual $X_i = X_{i1,} X_{i2,} X_{i3,,} ....X_{id,}$ of the spotted hyena species choose one element $X_K (k = 1,2,...,d)$ at random from the individual X_i with a probability of 1/d along with randomly construct a original number in the range l_i, u_i instead of the element X_{ik} from the individual X_i, therefore turn out a new individual $X_i^{'} = X_{i1,}^{'} X_{i2,}^{'} X_{i3,,}^{'} ....X_{id,}^{'}$ The variation mutation operation is

$$X_i^{'} = \begin{cases} l_i + \lambda(u_i - l_i) & i = k, Otherwise \\ X_i \end{cases} \qquad (30)$$

Where l_i, u_i are the higher as well as lower bounds of the variable X_i, respectively, and λε [0, 1] is a random.

3.3.5 Search for prey: Spotted hyenas generally seek the prey in relation to cluster location of spotted hyenas would inherent in vector $\overrightarrow{C_h}$ .They shift since one another in order to search as well as attack prey. Consequently, we utilized by way of formal rate as +1or -1 to oblige the seek representative to go far-off from the prey. These systems consent to the SHO algorithm toward seek worldwide.

3.4 Adaptive Deep Convolution Neural Network

In ACDNN, Convolution Nets were trained in an entirely supervised mode through eminent back propagation technique like deep learning models. The two important appearance of DCNN is weight distributions along with thin connections, which can diminish the parameters trained, tested and decrease the computational intricacy. The model scheduled on LeNet5, though utilizing max-pooling as an alternative of standard-collection. The form contains four layers: the initial two layers are convolution layers, third layer is hidden layers they are total link layer along with final layer is logistic regression layer. Structure of ADCNN is showed in below figure 3.

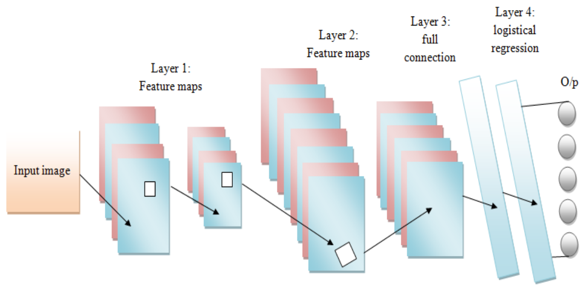

Fig 3: Structure of ADCNN

Input layer: Let Z be group of input functions, and every input together of many proportions. The function have written in the underneath eqn (31).

$$Z = Z_1 , Z_2 , Z_3 ....Z_n \qquad (31)$$

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 2, 2022, p. 2392 - 2403
https://publishoa.com
ISSN: 1309-3452

Activation layer: On convolution layer, the output characteristic map is created via past layers element chart were convolved with learnable kernels and were place via activation purpose. In basic, input and output junction is

$$P_j^{\,1} = g\left(\sum_{i \in N_j} P_i^{l-1} * W_{ij}^1 + b_j^1\right) \tag{32}$$

Where N_j express a collection of input maps, p_i^{l-1} correspond to the output of the l-1 layer along with input of next layer via a feature map i .therefore l is the index of the layer. W_{ij}^1 weight as of neuron in map i to neuron in map j in layer 1. $*$ be the convolution calculation. An additive bias b was known toward every output map. Moreover, in support of exacting output map, the input maps would be convolved among many kernels. Accordingly, if output map m and map k both add beyond input map i , then the kernels useful toward map i were various meant for outcome maps m with k.

Weight updation : With the gradient of the weight, the updated weight of DCNN is expressed in the below mathematical equation (33). The new weights were calculated in equation (34).

$$W_{ij}^{\;l} = W_{ij}^{l-1} - \frac{\partial E}{\partial w_{ij}^l} r_{ij}^l \tag{34}$$

The knowledge rate is a worldwide constant in conventional models. However, an enormous knowledge time will promote defeat inaccuracy among massive weights. Above troubles be able to restrict less quantity although it resolves more instruction instant. So as to resolution troubles this effort launches an algorithm by adaptive amount describe ADCNN. For every $W_{ij}^{\;l}$ renovating contain inconsistent $\alpha \, or \, \beta$ to adjust the knowledge rate updated in underneath eqn (35)

$$r_{ij}^l = \left\{ \begin{matrix} \alpha_{ij}^l \\ \beta_{ij}^l \end{matrix} \left\{ \begin{matrix} if\,(E^m - E^{m-1})\;\; (E^m - E^{m-2}) \geq 0 \\ if\,(E^m - E^{m-1})\;\; (E^m - E^{m-2}) < 0 \end{matrix} \right\} \right\} \tag{35}$$

Where $1 < \alpha < 2, 0 < \beta < 1$ the parameter hence $\alpha \;\; and \;\; \beta$ have chosen base on the test character as well as construction of the model. The knowledge time is higher toward speed up the divergence, if the loss role decreases continuously. Knowledge rate is decrease even as if the loss function altered late.

5. RESULT AND DISCUSSION

The part described the human behavior such as walking down-stairs, walking, eating, running, jogging, crying and praying. The graph is plotted with each of the activity compared with different algorithms. Some of the performance parameters like accuracy, precision; F1 and recall were calculated to achieve better results.

5.1 Dataset description

In this section we were discussed elaborately about human activity recognition dataset and it was taken from ourselves. Further, that dataset contains total of 7767 and 3162 events were used to get our desired result. Each event consists of 561 basic highlights. In this dataset, there are various types of images were taken. Each image represents different activities according to their positions.

5.2 Performance analysis

Here, Some of the below performance measures have taken for our proposed model and compared with our existing techniques. The measures are mathematically explained in the below formula.

***Accuracy**:* The term of accuracy is described as the amount of detection made correctly. The mathematical representation is below,

$$Accuracy = \frac{tp+tn}{tp+fp+fn+tn} \qquad (36)$$

***Precision**:* The precision is defined as the detection of positive class values by the proposed model accurately. The equation for the precision is given below,

$$Precision = \frac{tp}{tp+fp} \qquad (37)$$

Recall: The sensitivity or measure recall is defined as the correctly detected the given activity by the proposed approach. The formula for the recall is follows,

$$Recall = \frac{tp}{tp+fn} \qquad (38)$$

F-measure: F-measure means computation of mean value using precision and recall which provides the F-measure. It varies from 0 (worst performance) to 1 (highest performance). The formula for F-measure as follows,

$$F1 - measure = \frac{2\times(recall\times precision)}{recall+precision} \qquad (39)$$

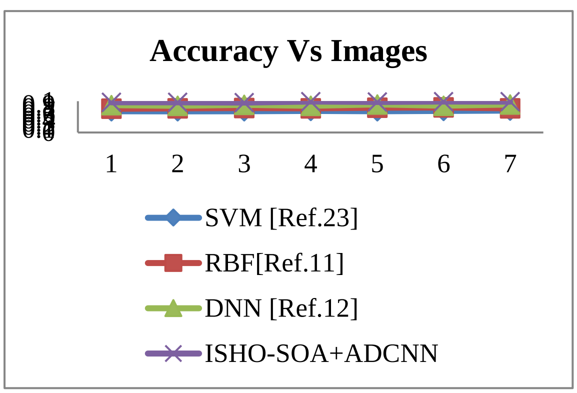

Fig 4: Accuracy

Figure 4 demonstrates the accuracy calculation for various input images and it was compared with classification techniques. Here, the activities we took as 1,2,3,4 etc. The given classification techniques such as SVM, RBF, DNN and hybrid (ISHO-SOA)-ADCNN. In SVM, we got above 40 for all the activities and next to this we got minimum accuracy for RBF. In Hybrid-ADCNN, accuracy for first activity is 98, second activity is 97.5, third activity is 99, and fourth activity is 97 and so on. Compared to all techniques our proposed method gave the better results.

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 2, 2022, p. 2392 - 2403
https://publishoa.com
ISSN: 1309-3452

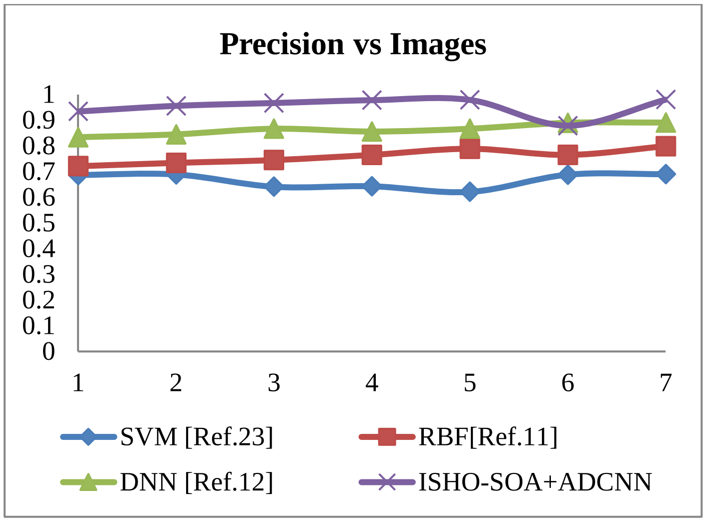

Figure 5: Precision

The above Figure 5 explains the precision calculation for various human activities. In SVM, we got minimum precision value for all inputs. Moreover, in RBF, we got 62 for first image, 64 for second image, 68 for third image, 79 for fourth image, and 67 for fifth image and so on. Compared to all techniques, we got better precision outcome by using Hybrid-ADCNN.

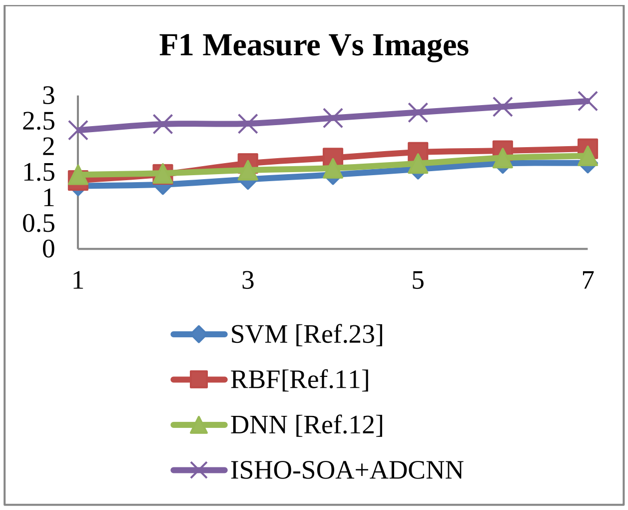

Figure 6: F1 measure

The above figure 6 explains the F1 measure plotted by using various classification techniques compared with proposed and existing techniques. X-axis represents human activities from 1 to 7 and y-axis represents F1 measure range from 0 to 3. Each activity gives different F1 measure corresponding to proposed and existing techniques. Compared to all techniques, our proposed model Hybrid-ADCNN gave better results.

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 2, 2022, p. 2392 - 2403
https://publishoa.com
ISSN: 1309-3452

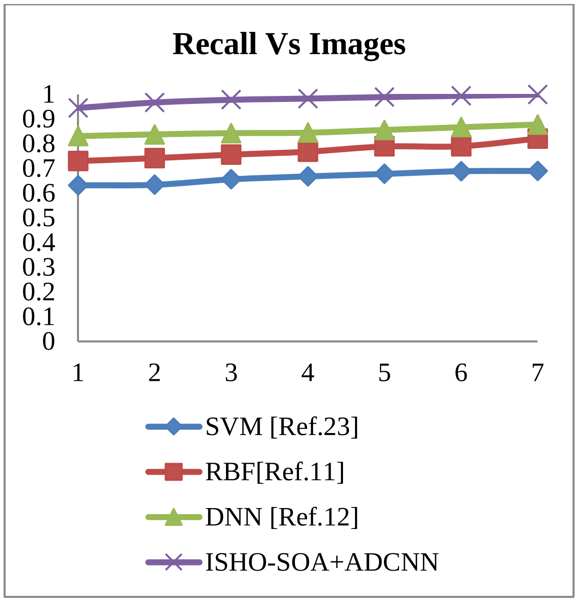

Fig 7: Recall

Figure 7 demonstrates the recall performance calculated for our input data. X-axis represents various human activities and y-axis represents recall measures. The compared techniques are SVM, RBF, DNN and Hybrid-ADCNN. Further, the hidden layer in ADCNN gave successfully predict our results with maximum values.

6. CONCLUSION

The proposed article is a hybrid multi-model framework for the efficient recognition of basic as well as transitional activities. The framework employed the hybrid algorithm of ISHO-SOA module followed with Adaptive deep learning models for the final classification of activities. The proposed approach has been tested on the available datasets for both basic and transition activities and compared with other state-of-the art approaches employing the same datasets. By improving the performance, the feature maps in ADCNN gives better results. Finally, our proposed method gave better results 98.21% accuracy. The performance is validated and evaluated with the existing technique that shows the proposed model is highly effective than existing models. In future, the identification of human activity will find with various creative algorithms and enhanced techniques.

REFERENCES:

1) Asim, Yusra, Muhammad Awais Azam, Muhammad Ehatisham-ul-Haq, Usman Naeem, and Asra Khalid. "Context-aware human activity recognition (CAHAR) in-the-Wild using smartphone accelerometer." IEEE Sensors Journal 20, no. 8 (2020): 4361-4371.

2) Aslan, Damla, Bengü Bayyurt Çetin, and İzzet Gökhan Özbilgin. "An innovative technology: Augmented reality based information systems." Procedia Computer Science 158 (2019): 407-414.

3) Wan, Shaohua, Lianyong Qi, Xiaolong Xu, Chao Tong, and Zonghua Gu. "Deep learning models for real-time human activity recognition with smartphones." Mobile Networks and Applications 25, no. 2 (2020): 743-755.

4) X. Peng, L. Wang, X. Wang, and Y. Qiao, ''Bag of visual words and fusion methods for action recognition: Comprehensive study and good practice,'' Comput. Vis. Image Understand., vol. 150, pp. 109–125, Sep. 2016.

5) J. C. Niebles, H. Wang, and L. Fei-Fei, ''Unsupervised learning of human action categories using spatial-temporal words,'' Int. J. Comput. Vis.,vol. 79, no. 3, pp. 299–318, 2008.

6) C. Thurau and V. Hlavac, ''Pose primitive based human action recognition in videos or still images,'' in Proc. IEEE CVPR, Jun. 2008, pp. 1–8.

7) T. Qi, Y. Xu, Y. Quan, Y. Wang, and H. Ling, ''Image-based action recognition using hint-enhanced deep neural networks,'' Neuro-computing, vol. 267, pp. 475–488, Dec. 2017.

8) B. B. Amor, J. Su, and A. Srivastava, ''Action recognition using rateinvariant analysis of skeletal shape trajectories,'' IEEE Trans. Pattern Anal. Mach. Intell., vol. 38, no. 1, pp. 1–13, Jan. 2016.

9) M. Oquab, L. Bottou, I. Laptev, and J. Sivic, ''Learning and transferring mid-level image representations using convolutional neural networks,'' in Proc. IEEE CVPR, Jun. 2014, pp. 28–37.

10) D. Wu and L. Shao, ''Leveraging hierarchical parametric networks for skeletal joints based action segmentation and recognition,'' in Proc. IEEE CVPR, Jun. 2014, pp. 724–731.

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

12) Murad, Abdulmajid, and Jae-Young Pyun. "Deep recurrent neural networks for human activity recognition." Sensors 17, no. 11 (2017): 2556.

13) Q. V. Le, W. Y. Zou, S. Y. Yeung, and A. Y. Ng, ''Learning hierarchical invariant spatio-temporal features for action recognition with independent subspace analysis,'' in *Proc. IEEE CVPR*, Jun. 2011, pp. 3361–3368.

14) G. W. Taylor, R. Fergus, Y. LeCun, and C. Bregler, ''Convolutional learning of spatio-temporal features,'' in *Proc. Eur. Conf. Comput. Vis.*, 2010, pp. 140–153.

15) S. Ji, W. Xu, M. Yang, and K. Yu, ''3D convolutional neural networks for human action recognition,'' in *Proc. Int. Conf. Mach. Learn.*, 2010, pp. 495–502.

16) H. Robbins and S. Monro, ''A stochastic approximation method,'' *Ann. Math. Statist.*, vol. 22, no. 3, pp. 400–407, 1951.

17) L. Bottou and O. Bousquet, ''The tradeoffs of large scale learning,'' in *Proc. Adv. Neural Inf. Process. Syst.*, 2008, pp. 161–168.

18) Y. Jia *et al.*, ''Caffe: Convolutional architecture for fast feature embedding,'' in *Proc. 22nd ACM Int. Conf. Multimedia*, 2014, pp. 675–678.

19) Martinez, J.; Black, M.J.; Romero, J. On human motion prediction using recurrent neural networks. *Comput. Aided Geom. Des.***2021**, *86*, 101964.

20) Li, C.; Zhong, Q.; Xie, D.; Pu, S. Co-Occurrence Feature Learning from Skeleton Data for Action Recognition and Detection with Hierarchical Aggregation. In Proceedings of the Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018.

21) Li, M.; Chen, S.; Chen, X.; Zhang, Y.; Wang, Y.; Tian, Q. Symbiotic graph neural networks for 3d skeleton-based human action recognition and motion prediction. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *1*, 10010.

22) Xu C., Chai D., He J., Zhang X., Duan S. InnoHAR: A deep neural network for complex human activity recognition. *IEEE Access.* 2019;**7**: 9893–9902.

23) Yogesh Hole et al 2019 J. Phys.: Conf. Ser. 1362 012121

24) Damodaran N, Haruni E, Kokhkharova M, Schäfer J (2020) Device free human activity and fall recognition using WiFi channel state information (CSI). CCF Trans Pervasive Comput Interact 2:1–7.