TEST Engineering & Management

January - February 2020
ISSN: 0193 - 4120 Page No. 3141 - 3147

# Robust Feature Based Automated Multi-View Human Action Recognition System Using Machine Learning

Dr. M. Mahalakshmi, Associate Professor, Department of ECE, CMRCET, Hyderabad, ,Telangana, India. (Email: strimaha@gmail.com)
Mr. M. Dayanidhy, Assistant Professor, Department of ECE, Velammal Institute of Technology, Chennai, Tamil Nadu, India. (Email: dayanidhy.phd@gmail.com)

***Article Info***
***Volume 82***
***Page Number: 3141 - 3147***
***Publication Issue:***
***January-February 2020***

***Article History***
***Article Received:*** *14 March 2019*
***Revised:*** *27 May 2019*
***Accepted:*** *16 October 2019*
***Publication:*** *19 January 2020*

***Abstract***
An automatic human action recognition technique is proposed in this paper. The main intention of this paperwork is to provide a new approach for image recognition using an artificial neural network. In multimedia processing and traditional techniques, human-machine interaction is a research topic for communication that allows disabled people to communicate easily with machines. Human body shapes which cause the change in appearance is also in the study for human action recognition. Our work focuses around human action recognition from a video scene. The given input image is pre-processed by Gaussian filter and a mean Eigen space is produced by taking a mean of the similar postures. In addition to these observances, we used Artificial Neural Network which produces the robustness.

***Keywords:*** *Image extraction, Human action recognition, K-nearest neighbour, Artificial neural network.*

## I. INTRODUCTION

The analysis of human behavior is important for detection, monitoring, and understanding of the physical behavior of the people. Actions can be described in various and different ways. The widely adopted interpretation describes actions as single periods of human motion patterns. Action execution styles and human body size variations appearing between different persons are addressing the action recognition methods. The video is subdivided intoframes. The processes involved in human motion analyses are given in the Fig.1.

The mean and standard deviation of the frames are calculated to find the threshold image. The frames are pre-processed to remove the unwanted noise present in it and transform the image as necessary for further processing. Convert the RGB image into a grey scale image. The frames are given a value for feature extraction using Laplacian Smoothening Transform. Most of the videos are unlabeled or weakly labeled. Collect well-labeled videos for consuming time and labor intensive. So as to utilize unlabeled videos, a semi-supervised classifier is trained dependent on the heterogeneous features in video space.

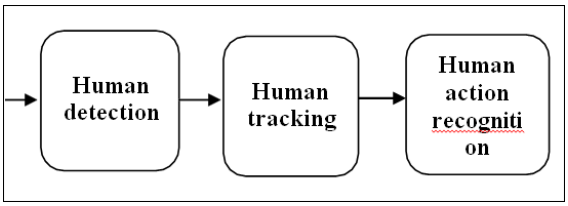


Fig.1. Human Motion Analyses steps

K-Nearest Neighbor is a classifier used here to find the nearest image in the dataset. The value of K should always be odd for any two classes otherwise there may be a tie. Hence it is favorable to choose an odd value. The value of K should not be equal to the multiple of the number of classes. When the dataset is large, it is complex to find the nearest neighbor. Hence we employ Artificial Neural Network which is useful for fast computation and less stimulation time. Artificial neural networks have been chosen for image

TEST Engineering & Management
January - February 2020
ISSN: 0193 - 4120 Page No. 3141 - 3147

compression due to their massively parallel and distributed architecture. The idea following this training command is the back propagation algorithm. Multilayer perceptions are used for action classification. The different viewing angle with multiple recognition [1], [2], [3] outcomes leads to action recognition [4] with high recognition accuracy.

## II. RELATED WORKS

The various action recognition system of human is a functioning exploration field because of its importance in a wide scope of applications, such as intelligent surveillance, human-computer interaction, content-based video compression, and retrieval, augmented reality, etc. An action or motion is alluded to as a solitary time of a human movement design, such as a mobile advancement. By perceiving the human body from different viewing angles, a view-invariant action representation is obtained. This portrayal is subsequently used to describe and recognize actions. The background subtraction is broadly utilized in video analysis. It rearranges further processing by locating regions of interest in the image. Along these lines, so as to perform action identification at high frame rates, the utilization of a less difficult activity depiction is required. Neurobiological examinations [5] have reasoned that the human mind can see activities by watching just the human body poses [6].

A few analysts have been as of now given an account of the estimation of human behaviors. For instance, there were a few methodologies in [7], [8], [9], [10] as strategies with utilizing 2D camera, which went for developing a framework that can comprehend the human behaviors utilizing the time series images for the motion of the human, and gauge the human activities by examining the action and the object target. These examinations assign that human behaviour recognition must be built up in human activity monitoring and the environment related with human action.

## III. METHODOLOGY

The objective of human action recognition methodology is to determine the states of the desired object parts in a video sequence and its generic framework is depicted in Fig.2.

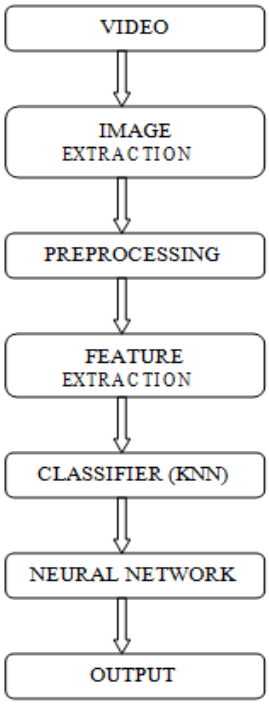


Fig.2. A generic framework for human action recognition algorithms

### A. ImageExtraction

In our technique, we segregate the image (static) feature from the two pictures and key edges of input video. Thinking about computational proficiency, we concentrate key frames by a shot boundary recognition algorithm. The color histogram of each five frames is determined. The histogram is subtracted with that of the past frame. The frame amidst the shot is considered as a key frame [11].

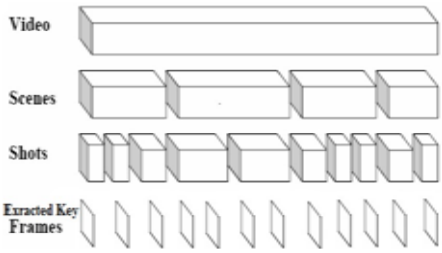


Fig.3. Overview of key frame Extraction

Key frame extraction as given in the Fig.3 is a necessary part in video analysis and management,

January - February 2020
ISSN: 0193 - 4120 Page No. 3141 - 3147

providing a suitable video summarization for video indexing, browsing and retrieval [12]. The use of key frames reduces the amount of data required in video indexing and provides the frame work for dealing with the video content. Thus the redundant frames are removed which reduces the computational complexity and improve recognition efficiency.

Shot Boundary Detection Algorithm:

Let N (x) be the xth frame in video sequence, x = 1, 2... Nn（Nn signifies the total number of frames in the video). The shot edge identification algorithm is portrayed as pursues:

Step 1:

Dividing each frame into blocks with n columns and m rows, and F(i, j, x) stands for the block at (i, j) in the xth Frame.

Step 2:

Computing x2 histogram coordinating contrast between the comparable blocks between back to back frames in video input. The frames (i, j) in the xth and (x+1) th alluded as h(i, j, x) and h(i, j, x+1) individually, which represents the histogram of blocks and Block's difference is estimated measured by the accompanying condition:

$$D_b = (x, x+1, i, j) = \sum_{i=0}^{L-1} \frac{[h(i,j,x) - h(i,j,x+1)]^2}{h(i,j,x)}$$

Here, the count of gray image is accounted as L

$$STD = \sqrt{\sum_{x=1}^{N_{v-1}} (D(X, X+1) - MD)^2 / N_{v-1}}$$

Step 3:

Histogram Distinction x2 calculation between two sequential frames:

$$D(x, x+1) = \sum_{i=1}^{m} \sum_{j=1}^{n} w_{ij} D_B(x, x+1, i, j)$$

Here, the block (i, j) weight is referred as $w_{ij}$.

Step 4:

Automatic Thresholding Measurement:

The entire video input subjected to k2 histogram distinction for the Mean and standard variance computation.

Mean and standard variance are characterized as follows :

$$\text{Mean Deviation MD} = \sum_{x=1}^{N_{v-1}} D(x, x+1)/N_{v-1}$$

Step 5:

Detection of Shot Edge:

Threshold Value= (Standard Deviation × $\alpha$ + Mean Deviation

Where $\alpha$ is a constant varies from 0 to 1.

**B.** Feature Extraction

Feature extraction is a notable form of dimensionality reduction. It is done after pre-processing techniques in character recognition. It is the process by which certain features of curiosity within an image are detected and represented for further processing. The main aim of feature extraction is to procure the most pertinent information from the original data and represent the information in a lower dimensionality space. It involves clarifying the number of resources required to describe a large set of data accurately by using Laplacian Smoothening Transform algorithm. Feature extraction is a habitual term for methods of constructing combinations of the variables. It reduces the number of random variables. If the feature is analysed with large number of variables, then it requires large amount of power computation and memory. A good feature setcontains discerning information which can differentiate one object from the other. It has been used in many applications suchas character recognition, document verification, script recognition, checksortingetc.

**C.** Classifier

K-Nearest Neighbour [13], is one of the machine learning algorithms which is shown in Fig.4. KNN is the simplest machine learning algorithm.

The preparation procedure of this calculation comprises of putting away component vectors and marks of the preparation pictures.

TEST Engineering & Management
January - February 2020
ISSN: 0193 - 4120 Page No. 3141 - 3147

The training process of this calculation comprises of storing labels of the training images and feature vectors. The KNN binary is given more accurate data classification. It is an object which is classified through a mainstream selection of its neighbors. Euclidean distance is used as the distance metric. It works built on a minimum distance from the interrogation instance to the training samples to regulate the K-nearest neighbors. The information of KNN can be any dimension scale from insignificant, to a measurable scale. When there are two classes, k must be an odd integer. However, there will be a tie when k is an odd integer when performing multiclassclassification.

When K=1,

$R_i = \{x: d(x, x_i) < d(x, x_j)\}; i \neq j$

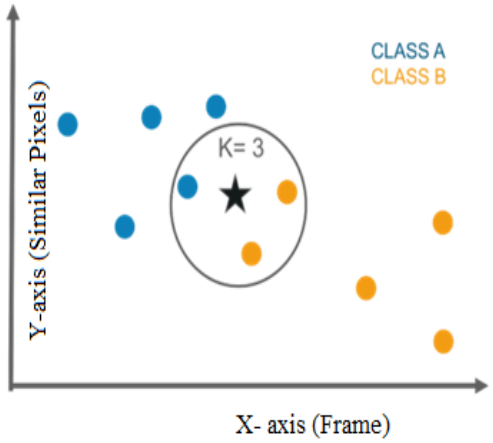


Fig.4. K-Nearest Neighbor Classification

A main advantage of the KNN algorithm is that it performs well with multi-modal2 classes because the basis of its decision is based on a little neighborhood of comparable items. In this manner, regardless of whether the target class is multi-modular, the algorithm can even now prompt great accuracy. The significant weakness of the KNN calculation is that it utilizes every one of the features similarly in processing for similarities. This prompts classification errors, particularly when there is just a little subset of features that are valuable for classification.

The K-NN technique is an apathetic algorithm dissimilar to numerous different machines learning strategies, for example, artificial neural networks, kernel methods [14], and wavelet systems [15], [16] and so on. When the dataset is large, it is complex to find the nearest neighbor. Hence we employ Artificial Neural Network because it is useful for fast computation and less stimulation time.

### D. Artificial Neural Network

The computers can solve a problem by a set of instructions which should be programmed by a human. But neural network acts like a brain. It need not to be programmed, it learns through examples or experiences. Hence, it needsdata set to train. Artificial neural network [shown in Fig.5 is a computing system i.e., designed to stimulate the way the human brain analysis and processes the information. It has self- learningcapability that enables as to produce better results. A single artificial neuron called perceptron is used. The modes of perceptron are training mode and usingmode.

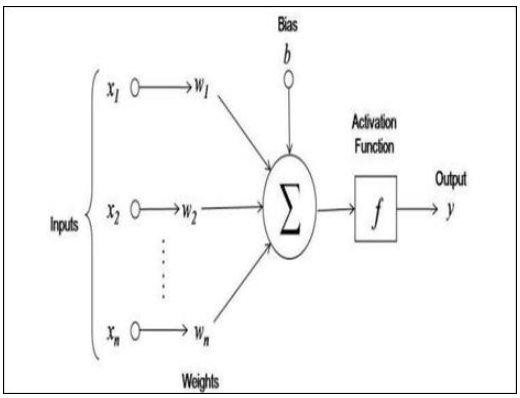


Fig.5. Artificial Neural Network input/output function

The multilayer perceptron is used for image recognition. The input image is dispatched to the input layer where the patterns of local contrast are recognized. The multilayer perceptron is composed of two hidden layer. The hidden layer one recognizes the features of the face such as eyes, nose and mouth where the hidden layer two is used for reconization of face. The output of the hidden two is forwarded to the output layer which is an ANN output.

## IV. RESULTS AND DISCUSSION

The video from the KTH database is converted into frames which is used for further processing.

January - February 2020
ISSN: 0193 - 4120 Page No. 3141 - 3147

The key frame extraction is based on frame difference. The Fig.6 shows the frames which are extracted from the video. The image contains noise. The noise in the image is suppressed using Gaussian filter. The figure 7 shows the preprocessed image

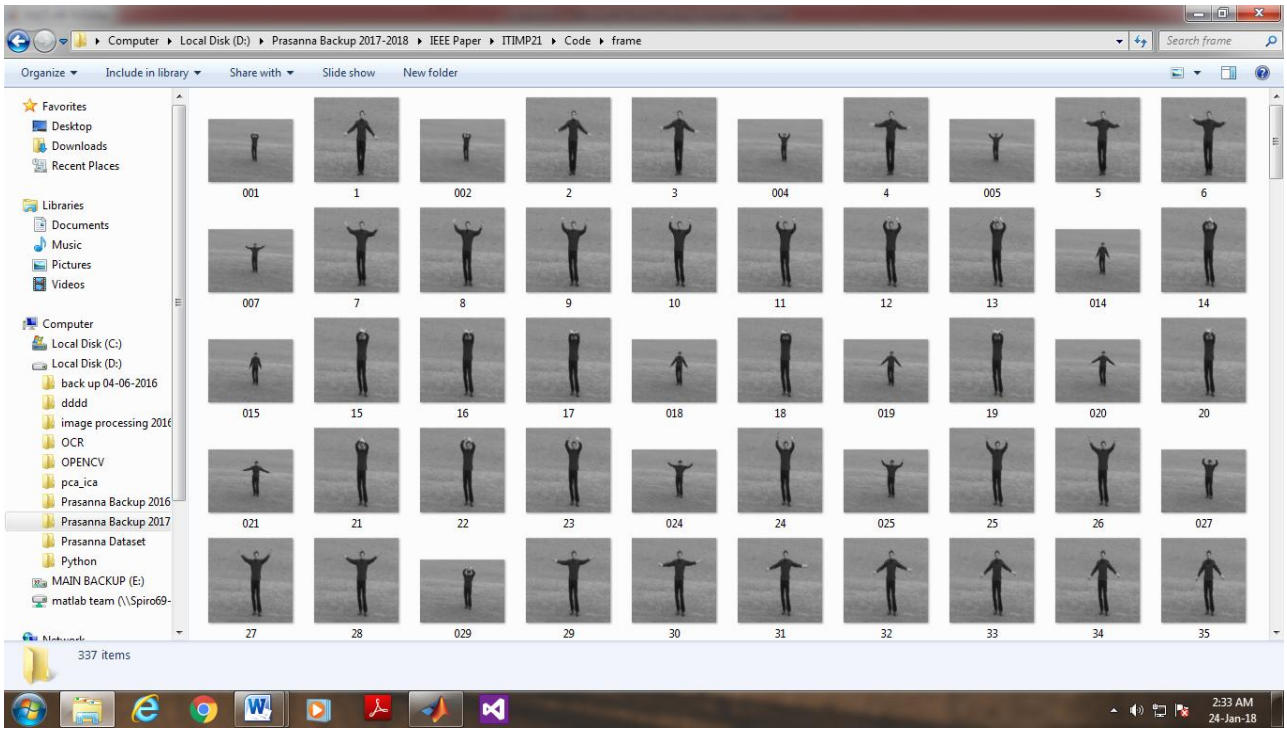


Fig.6. Frame Conversion Output

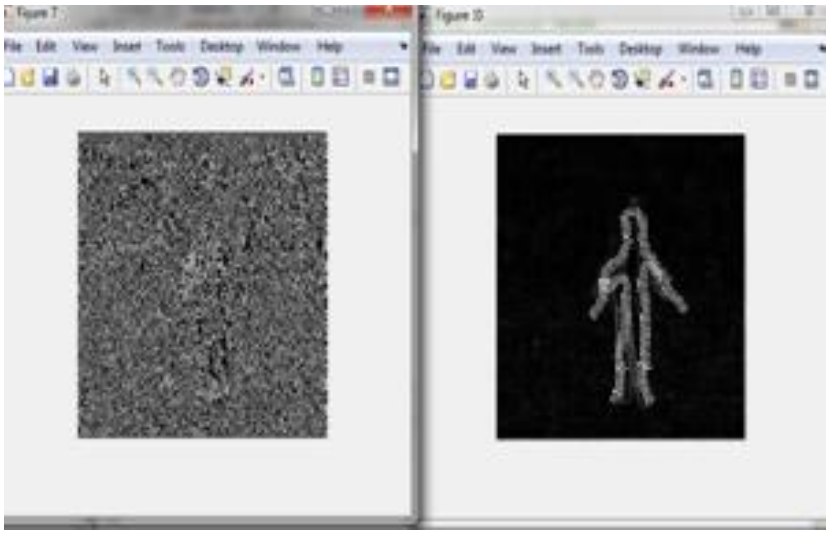

Fig.7. Preprocessing Output

The magnitude of the gradient value in x direction is used to find the action of the image. The angle is used to find the direction of the image. The K-Nearest Neighbor algorithm is applied to the preprocessed image. KNN is based on similarity measure. The Artificial Neural Network algorithm is then applied to the K-nearest neighbor output and then further processed. The Artificial Neural Network output is shown in figure 8.

January - February 2020
ISSN: 0193 - 4120 Page No. 3141 - 3147

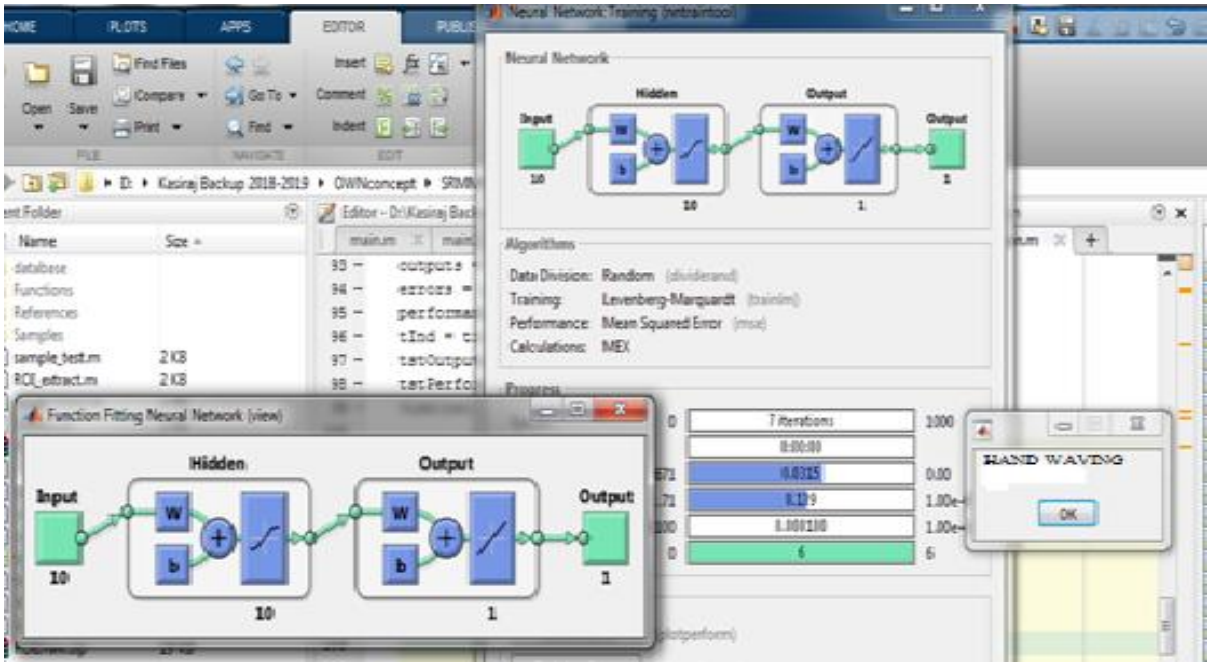


Fig.8.Artificial Neural Network Output

## V. CONCLUSION

In this paper, we have elaborated the various steps to examine the behaviour of an individual human from the sequence of videos with their action recognition. Particularly, we have concentrated on the classification through videos. In this part we utilized the classification of KNN technique on the KTH basic of videos. Artificial neural network can be applied to train and test the image for the purpose of recognition. The test image is recognized and profitably matched with original image. ANN reduces the time for training and testing. It is more suitable for practical application. Compared with other works, our approach is easier to be implemented and has better performance.

## VI. REFERENCES

[1] AndrzejDziech, Ali Amuri, "contour recognition using neural network application", Communications Dept.,AGH Cracow University, Cracow.

[2] NilamNur Amir Sjarif, NilamNur Amir Sjarif, "Human Action Invarianceness for Human Action Recognition" 2015 IEEE.

[3] Abhinav Gupta, AniruddhaKembhavi and Larry S. Davis "Observing Human-Object Interactions: Using Spatial and Functional Compatibility for Recognition" pattern analysis and machine intelligence, vol. 31, no. 10, October 2009.

[4] Chen Chen, RoozbehJafari and Nasser Kehtarnavaz "A Real-Time Human Action Recognition System Using Depth and Inertial Sensor Fusion" IEEE sensors journal, vol. 16, no. 3, February 1, 2016.

[5] Chandni J. Dhamsania and Tushar V. Ratanpara, "A Survey on Human Action Recognition from Videos"2016 IEEE.

[6] Qingfeng Liu and Chengjun Liu "A Novel Locally Linear KNN Method with Applications to Visual Recognition" neural networks and learning systems, pp 2162-237X © 2016 IEEE.

[7] AlexandrosIosifidis, AnastasiosTefas, and Ioannis Pitas, "View- Invariant Action Recognition Based on Artificial Neural Networks" neural networks and learning systems, vol. 23, no. 3, march 2012.

[8] MaimaitiminMaierdanh, KeigoWatanabel and ShoichiMaeyamal "Estimation of Human Behaviors Based on Human Actions Using an ANN", Control, Automation and Systems pp 22-25, Oct 2014.

[9] Md. IqbalQuraishi, J Pal Choudhury and Mallika De "Image Recognition and Processing Using Artificial Neural Network" 978-1-4577-0697-4 ©2012 IEEE.

[10] M.Egmont-Petersena, D. de Ridderb, H. Handelsc, L. Beaurepaire, K.Chehdi; B.Vozel "Image processing with neural networks—a

January - February 2020
ISSN: 0193 - 4120 Page No. 3141 - 3147

review", Pattern Recognition 35 (2002), 2002, PP-2280–2288.

[11] A.V.Kumthekar, Prof.Mrs.J.K.Patil, "Key frame extraction using color histogram method". Volume 2 Issue 4 pp 207-214 July 2013.

[12] Nick C. Tang, Yen-Yu Lin, Ju-HsuanHua, Ming-Fang Weng, and Hong-Yuan Mark Liao "Robust Action Recognition via Borrowing Information across Video Modalities" image processing ,vol. 24, no. 2, February 2015.

[13] Shichao Zhang, Xuelong Li, Ming Zong, Xiaofeng Zhu, and Ruili Wang, "Efficient KNN Classification with different Numbers of Nearest Neighbors", neural networks and learning systems 2162-237X © 2017 IEEE.

[14] Kenji Suzuki, "Pixel-Based Artificial Neural Networks in Computer-Aided Diagnosis", Department of Radiology, Division of Biological Sciences, the University of Chicago USA,2015.

[15] Berend Jan van der Zwaag , Kees Slump , and Lambert Spaanenburg" Extracting Knowledge from Neural Networks in Image Processing" PP-143-145.

[16] Adilson Gonzaga, Armando Marin, Evandro A. Silva, Fabiana C. BertoniKelton A.P. Costa, Luciana A.L. Albuquerque "Neutral Facial Image Recognition Using Parallel Hopfield Neural Networks", Universidade de São, Paulo,2013.