New Framework for Analysis of Human Behaviour

[1]Shavet Bhagat,[2]Sukhwinder Singh
[1]Student,[2]Assistant Professor
[1]Electronics and Communication Department
[1]Punjab Engineering College(deemed to be university), Chandigarh,India

***Abstract*— As time passes computer applications are becoming vital part in every field. Due to which it has been used in several applications. Thus detection of object from a video and then tracking of that object is an important task in computer vision. Similarly, the human action recognition has also become as a prominent domain for research work to introduce novel advancements. This study develops a novel human action recognition system on the basis of the Deep Neural Network (DNN). For the purpose of the human recognition system a dataset of images is considered and then the feature extraction is performed from the selected colored and depth image from the dataset. Along with this the segmentation is also applied to the extracted image. After then, the DNN is applied to train the dataset. The reason behind implementing the DNN is to overcome the parametric independent issue of naïve bayes classifier. The simulation of proposed work is done in MATLAB simulation platform and obtained results shows that the proposed work performs the human action recognition with high accuracy.**

Keywords—Human Action Recognition, Feature Extraction, Segmentation, Classifiers, Deep Neural Network.

I. INTRODUCTION

Recently, due to the introduction of automatic human activity recognition techniques worldwide attention has started focusing on video analysis methods because of the increasing demand in various fields like surveillance, entertainment environments and [1] healthcare systems. In case of the surveillance system, to automatically determine the unusual activities and also to make the related authority aware about the unusual activity [2] like automatic getting alert about the person with a bag loitering at an airport. In the same way in case of the entertaining environment by implementing the activity recognition technique the human computer interaction (HCI) can be increased like automatically detecting the action [3] did by the different players while playing the table tennis in order to form an avatar in the computer who is capable to play tennis for the player. Moreover, in case of health care system after implementation of activity recognition technique [4], the patient's rehabilitation can be improved like automatic determination of action did by a patient to facilitate the rehabilitation processes [5].

In the human activity recognition mechanisms the low-level core technique can be efficiently integrated having unique individual activity recognition, many people interaction and crowd behavior [6], and abnormal activity recognition. Following are some of the actions that are possible to recognize by using the human action recognition system [7].

1.1 Trajectory

An individual mobiles as a function of time is the path that is called as trajectory. In order to examine the activity or behavior of the tracked individual the trajectory of a tracked individual in a scene is often applied. To track and identify sports videos, like hockey and soccer a PCA-HOG descriptor is utilized [8]. The trajectory is utilized to examine the player's run-left, run-right, runs in, or run-out when the player plays with a ball. Furthermore, through examining trajectories the loitering manner can be simply contingent [9]. In order to examine the pedestrian location and velocity by utilizing the Kalman filters that are mainly applied to categorize a walking pedestrian, a running pedestrian, a loitering pedestrian or a falling-down pedestrian. Through assuming a linear discriminative mechanism [10] on the basis of the clothing color to investigate loitering pedestrians an appearance-based mechanism is applied. In order to judge the current interval for a pedestrian the time stamps for every pedestrian are utilized as well as to investigate the loitering cases a strip-shifting paradigm is deliberated [11].

1.2 Falling Detection

The falling detection is the other major subject of the individual user activity recognition. For the protection and the secure atmospheres the falling detection is effectively important especially for the mature people who live alone. To extort the one-pixel-wide edge characteristics the human shape based paradigm is offered and for the falling detection the string matching paradigm can be utilized. A back-propagation neural network (BPNN) and a tri-axial accelerometer can also be utilized for the falling detection. The BPNN is utilized to identify the four daily activities and the four falling actions that are the walking, jumping, getting into bed, rising from the bed and the front fall, back fall, left fall, right fall [12].

1.3 Human Pose Estimation

In the computer vision community the human pose estimation is a significant topic. On the basis of the outcomes of the human poses the the activity of the human can be effectively identified. For this a 2D model can be utilized for torso detection and tracking, and a skin color model can be utilized for hands tracking. Also 3D reconstruction can be utilized by multi-view from synchronized many cameras [13]. Afterward the 2D and 3D coordinates are altered into a standardized feature space, and categorized through closest mean classifiers (NMC) for identifying the major poses.

1.4 Multiple People Interaction and Crowd Behavior

Because of the requirements of atmosphere protection currently several people interaction and the crowd manner have drawn more concentration. On the basis of the state-of-art cascade of boosted integral characteristics through investigating the head areas of the people and through examining they can be approximately counted. In order to estimate the presentation of the head counting mechanism the PETS 2012 and Turin metro station dataset can be applied. Also by applying the skeleton graph the people can be counted quickly. In the head, torso and limbs the skeleton silhouette can be decayed [14]. Moreover, an adaptive background subtraction mechanism merging color and gradient knowledge to eliminate shadows and unreliable color cues can be applied to track crowd of people. Through the color histogram knowledge as the group is splitting up into various sets and the various sets can be joined into a single set to be dealt with. Among objects it also investigated the various interactions. Therefore rather than the single persons only crowd of persons are tracked.

1.5 Abnormal Activity Recognition

Generally to describe the abnormal activity it is too complicated to explain that what an abnormal or unusual activity is, as its description is on the basis of the contexts and surrounding atmospheres. Furthermore, it is not normally experimented from the description or else it is normal or usual [15]. However, instead of constructing a single abnormal activity model, a deviation method has been commonly accepted. A normal model can be constructed as in the background subtraction having specified instances or earlier seen data and taken into an account novel examination as abnormal or unusual when it diverge very much from the trained model [16].

II. PROBLEM FORMULATION

In current era, with an exponential growth in big data and computing technology, the automatic human action recognition also gets advanced to the level where more challenging application domains are growing. Traditionally, the classifiers like naïve bayes and neural network were used for the purpose of recognizing the human actions. The objective of utilizing the neural network as classifier was to reduce the complexity and such was obtained. Other than this, the naïve bayes classifiers were also implemented. The naïve bayes classifier is a simple classifier that utilizes the bayes theorem and allocates the class labels to the problem instance. The naive bayes classifier is accepted as the most prominent classifier for different tasks. But it is not suitable for each and every domain. The major issue in the naïve bayes classifier is that it assumes that all the value of a feature is totally independent to the value of other feature. This strong feature independent assumption leads to the very bad results generated by the naive bayes classifier. Similarly the naïve bayes classifier has some other issues also like data scarcity and common features. Thus there is a need to upgrade the classifiers that could be capable to handle all these issue by resolving them efficiently.

III. PROPOSED WORK

On coming to the fourth generation it is evaluated or observed that the technology has changed a lot by introducing so many advancements in each and every domain. Similarly, in Human action recognition system the training of the data sets is shifted toward DBN (Deep Belief Networks or Deep Learning Neural Network) from Machine learning. Initially, neural Networks were used to reduce calculation complexity of machine based and statistical based learning system. But then DBN comes to the existence that is capable to reduce the evaluation and implementation complexity.Thus, in this study, we have decided to implement the Deep Learning Algorithm as a classifier for human action detection. The training of the data set is done by using the DNN (Deep Neural Network). The advantage of the DNN is that the DNN follows the same topology as multi layered neural network posses. The only difference is that in DNN the number of hidden layers is higher so that it can be considered as deep network. In DNN the training is done by using a deep belief neural network (DBN) instead of direct learning. In this the parameterized layers are optimized together thus it can be defined that in this the all the parameters are dependent to each other which was a lacking point of naïve bayes classifiers. The flow of the proposed human action recognition system is shown in figure 1.

The step wise flow of the proposed work is as follows:

1. Start.
2. For starting the process of human action recognition the top most work is to select the input image for the purpose of the simulation. In order to select the image, the proposed work empowers the users to select the color and depth image from the available dataset of the motion images.
3. After selecting the images, next step is to extract the information from the selected images. The information that is relevant or meaningful is extracted from the depth image.
4. After information extraction the segmentation is applied to the image. The image segmentation is done to detect the objects from the selected image.
5. On the segmented image, the objects are detected from the image.
6. After detecting the objects, the next step is to evaluate the histogram corresponding to the detected object.
7. After calculating the histogram, the Training of the Deep Neural Network takes place. In this phase the network is trained on the basis of the selected portion of the dataset.
8. After training the network, the testing is performed for the rest of the dataset on the basis of the trained dataset.
9. End.

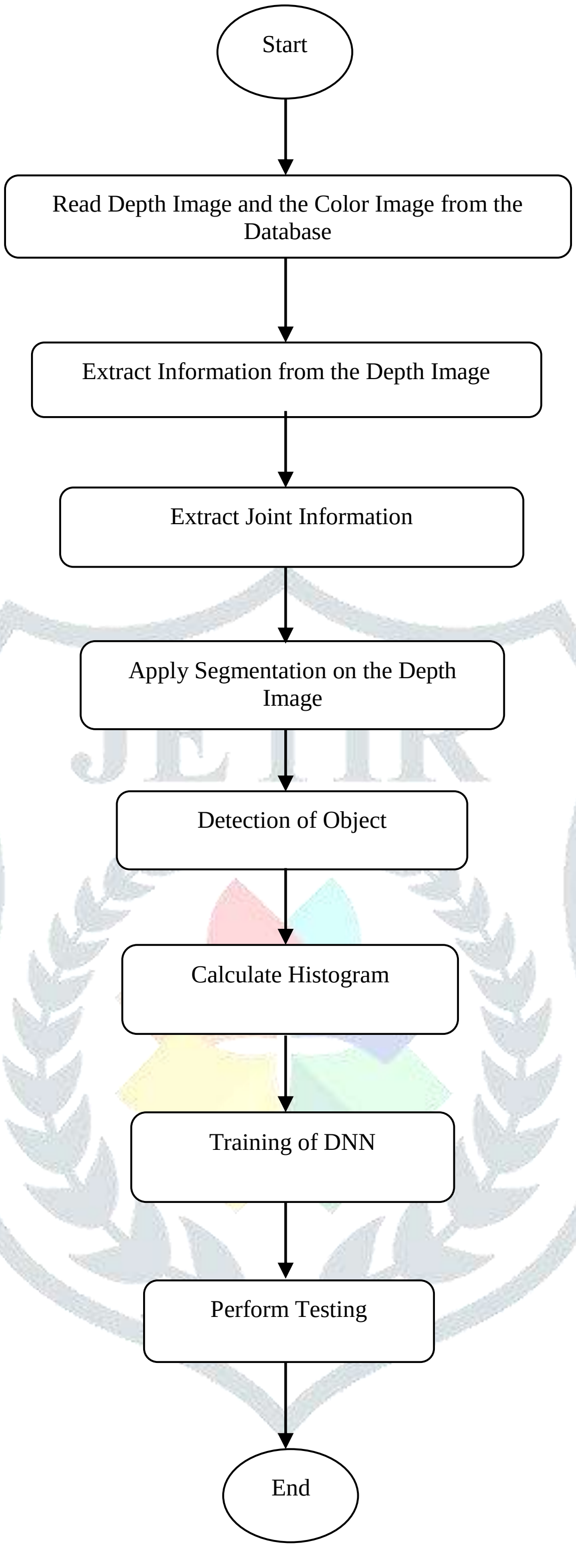

Figure 1 Framework for proposed Human Action Recognition System.

IV. RESULTS

This study develops a DNN based human action recognition system that performs various tasks like information extraction, image segmentation, training and testing of the dataset etc. In this work the human action recognition is done by recognizing the various actions performed by the humans such as pushing, to carry something etc.

In this work the projected mechanism is calculated by comparing it among the state- of-the-art methods through utilizing four public benchmark datasets. This section delineates some of the simulated results to ensure the accuracy of the proposed work.

The figure 2 depicts the input image that is used for the simulation purpose. The image depicts a person who is in the motion to push something. After selecting the image, the next step is to read the depth image from the available dataset.

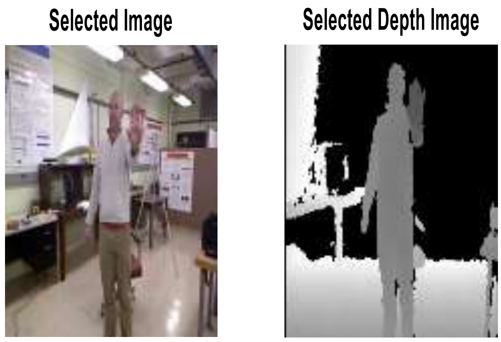

Figure 2 Selected Image of a person in the motion to push something.

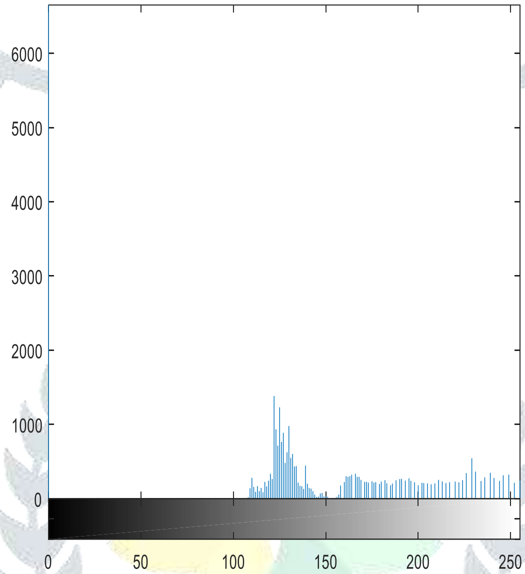

Figure 3 Histogram of the selected image in the motion to push something.

In the Figure 2 in the selected image of the person it is shown that the person performs an action of Push. So the histogram of that action of pushing something or movement of pushing is demonstrated in the figure 3.

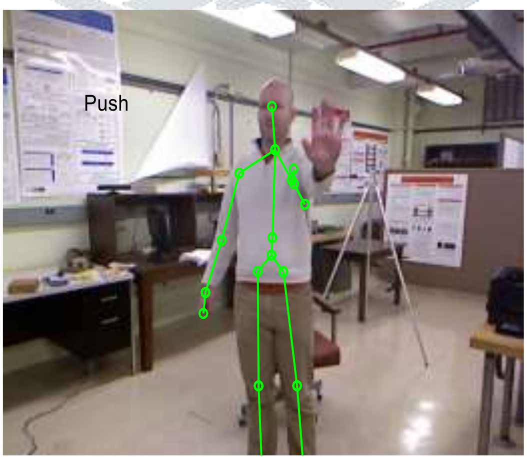

Figure 4 Action is detected by Skeleton.

In the figure 4 the action of push is detected by recognizing the skeleton of the person by using DNN. In the Figure 5 in the selected image of the person it is shown that the person performs an action of Carry something. So the histogram of that action of carrying something or movement of pushing is demonstrated in the figure 6.

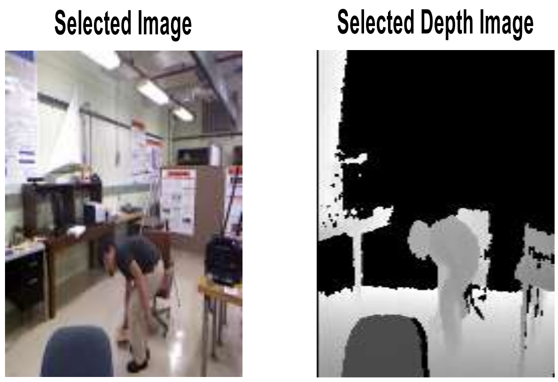

Figure 5 Selected Image of a person in the motion to carry something.

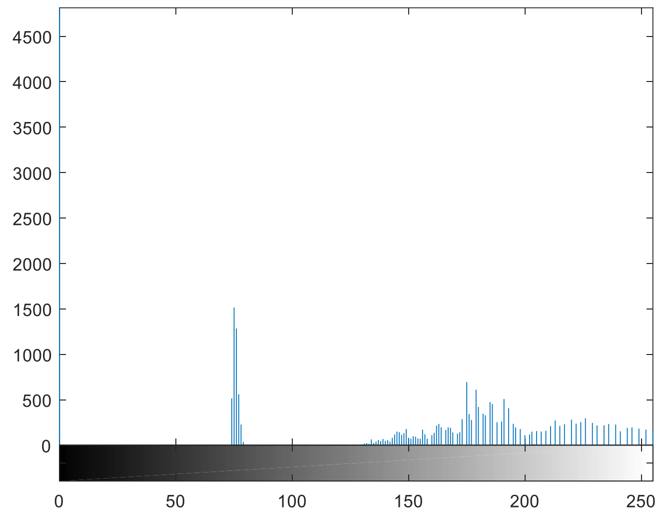

Figure 6 Histogram of the selected Image of a person in the motion to carry something.

As shown in the figure 2 and the figure 5 due to different actions performed by the person in the selected images so the histogram of both the images is also different as shown in the figure 3 and figure 6.

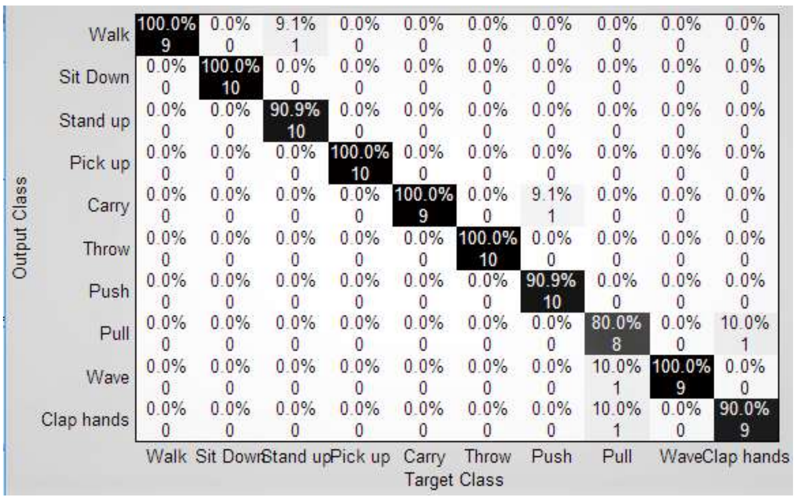

Figure 7 Confusion Matrix.

In the confusion matrix the output class of the different actions (fig 7) is observed by applying the projected paradigm that is Deep Neural Network. By utilizing the projected mechanism we obtain the accuracy in identifying the actions of the person in the image. Here the actions that are identified are Walk, Sit down, Stand up, Pick up, Carry, Throw, Push, Pull and wave. Now the accuracy shown in the confusion matrix of the actions like Walk, Sit down, Pick up, Throw and Wave is 100 % obtained by the projected mechanism that is DNN. For the action of standup it shows 90.9% accuracy whereas it shows the 9.1% as walk. Also for the action of the Push it shows 90.9% accuracy and here it shows 9.11% the action of carry and same as this the accuracy of the action of the Pull is 88.9% where 11.1 % is the action of wave is illustrated here. So the from the results it is verified that our projected mechanism offers the most accurate output class comparative to the other methods.

Table 1 Accuracy analysis

S. No.	Accuracy (%)	
	Methods	Depth+Skeleton
1.	DCSF[18]	-
2.	Moving Pose[19]	-
3.	ActionLet[20]	66.0
4.	DOM [21]	71.0
5.	Traditional Method [1]	80.9
6.	Proposed Method	92.9293

V. CONCLUSION

The human action recognition is the widely used application for the surveillance purpose. A large number of cameras are deployed at the public places as well as private places to record the human activities so that the unwanted activities could be recognized in case of emergency. For the purpose of human activity recognition a large number of researches have been conducted by the various scholars till now. For this purpose different classifiers are used such as Neural Network, Naïve Bayes classifiers etc. This study observed that the naïve bayes classifier fails to generate the results with high accuracy due to its assumption that all the parameters are independent to each other. Thus the DNN is applied to overcome this shortcoming of Naïve Bayes classifier. The simulated results show that the DNN leads to the more accuracy in action recognition generated results. In future the proposed work can be enhanced more to attain the high level of accuracy by working on the feature extraction process. The feature extraction could be done by applying any of advanced and prominent feature extraction technique.

REFERENCES

[1] Maxime Devanne, StefanoBerretti , PietroPala, HazemWannous ,Mohamed Daoudi , Alberto Del Bimbo. 2017. Motion segment decomposition of RGB-D sequences for human behavior understanding, ELSEVIER, Pp 222-223.

[2] Lu Xia, Chia-Chih Chen, and J. K. Aggarwal. 2012. View Invariant Human Action Recognition Using Histograms of 3D Joints, IEEE, Pp 1-8.

[3] Hossein Rahmani. 2017. 3D Human Action Recognition, IEEE, Issue 99.

[4] Xuelu Wang. 2017. Human Action Recognition from Gradient Boundary Histograms, research gate.

[5] Bo Lin ; Bin Fang. 2017. Spatial-temporal histograms of gradients and HOD-VLAD encoding for human action recognition, IEEE.

[6] Yang Qin, Lingfei Mo, Benyi Xie. 2018. Feature fusion for human action recognition based on classical descriptors and 3D convolutional networks, IEEE.

[7] Qiuping Chun ; Erhu Zhang.2017. Human action recognition based on improved motion history image and deep convolutional neural networks, IEEE.

[8] Jiali Jin ; Zhenhua Wang ; Sheng Liu ; Jianhua Zhang ; Shengyong Chen ; Qiu Guan. 2017. Joint label-interaction learning for human action recognition, IEEE, 2017.

[9] Kiran Talele, Dr. Kushal Tuckley. 2016. Human Action Unit Detection of Patient using Geometric Feature Analysis, 2016 IEEE Region 10 Conference (TENCON) - Proceedings of the International Conference, pp. 2156-2161.

[10] Okan Candir, M. Elif Karsligil. 2016. Detection of the Periodicity of Human Actions for Efficient Video Summarization, 2016 IEEE International Symposium on Multimedia, pp. 391-394, 2016.

[11] Riccardo Monica, Jacopo Aleotti, Stefano Caselli. 2015. GMM-Based Detection of Human Hand Actions for Robot Spatial Attention, 2015 International Conference on Advanced Robotics (ICAR), pp. 1-6, 2015.

[12] Zhong Zhou, Feng Shi, Wei Wu. 2015. Learning Spatial and Temporal Extents of Human Actions for Action Detection, IEEE TRANSACTIONS ON MULTIMEDIA, Pp. 512-525.

[13] Alessandro Pieropan, Giampiero Salvi, Karl Pauwels, Hedvig Kjellström. 2014. Audio-Visual Classification and Detection of Human Manipulation Actions, 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2014), Pp. 3045-3052.

[14] Nikhil Naikal, Pedram Lajevardi, Shankar. S. Sastry. 2014. Joint Detection and Recognition of Human Actions in Wireless Surveillance Camera Networks, 2014 IEEE International Conference on Robotics & Automation (ICRA), Pp. 4747-4754.

[15] Trevor Beugeling, Alexandra Branzan Albu. 2013. Sway Detection in Human Daily Actions Using Hidden Markov MODELS, 6th Annual International IEEE EMBS Conference on Neural Engineering, Pp. 1582-1585.

[16] Mohamad Hanif, Md Saad, Aini Hussain, Win Kong 2013. SESRG-InViSS : Image and Video Data Set For Human Pose, Action, Activity and Behaviour Detection, 2013 IEEE 3rd International Conference on System Engineering and Technology, Pp. 37-42.

[17] Wang Yaonan.2009. Effective method for tracking multiple objects in real-time visual surveillance systems, IEEE, 20(6), Pp.1167-1178.

[18] L.Xia,J.K.Aggarwal.2013. Spatio-temporal depth cuboids similarity feature for activity recognition using depth camera, in: CVPR Workshop on Human Activity Understanding from3DData,Portland,Oregon,Pp.2834–2841.

[19] M. Zanfir, M. Leordeanu, C. Sminchisescu. 2013. The moving pose: an efficient 3D kinematics descriptor for low-latency action recognition and detection, in: IEEE, Pp.2752–2759.

[20] J.Wang,Z.Liu,Y.Wu,J.Yuan. 2012. Mining action let ensemble for action recognition with depth cameras, IEEE, Pp.1–8.

[21] G.Yu,Z.Liu,J.Yuan. 2014. Discriminative order let mining for real-time recognition of human–object interaction, Asian Conference on Computer Vision (ACCV), Pp.50–65.