Human Motion Detection and Action Recognition

Bhargav Patel[1], Jaykumar Shah[2], Raj Dattani[3], Shalini Shah[4]

[1-4]*G H Patel College of Engineering & Technology, Gujarat*

---***---

Abstract— *In the discipline of image processing, detecting human motion and distinguishing their behaviours from video feeds is critical. However, identifying and recognising the many sorts of human activities with more precision and accuracy is a difficult process. An enormous amount of work has gone into developing systems that can be trained to alert about individuals whose actions are suspicious. Applications in the realm of security and surveillance that involve human action recognition, such as airport monitoring, retail surveillance, and other key places where security is a top priority.*

Intelligent video surveillance and environmental home monitoring, video storage and retrieval, intelligent human–machine interactions, and identity recognition are just a few of the applications for human action recognition. Human action recognition encompasses a wide range of computer vision research areas, including human identification in video, human position estimation, human tracking, and time series data processing and understanding. In the fields of computer vision and machine learning, it's also a difficult problem. Many fundamental difficulties in human action recognition remain unsolved now. Despite its widespread use in a variety of applications, accurate and efficient human action identification remains a difficult topic of research in computer vision. Human action recognition approaches using depth data, 3D-skeleton data, still picture data, spatiotemporal interest point-based methods, and human walking motion recognition have been the focus.

Keywords— Human Motion Analysis, Human Motion Representation, Human Motion Recognition, Recognition Methods

I. INTRODUCTION

Healthcare professionals can utilise Human Activity Recognition to get clinically important information about a person's mobility. When using a smartphone to characterise immobile states, Human Activity Recognition generally uses phone orientation to distinguish between sit, stand, and lie. While phone orientation is useful for determining when a person is sleeping, sitting, bending, and standing can be misclassified due to comparable pelvis alignment. As a result, it's tough to train a classifier using this data. To improve sit-stand categorization, a hierarchical classifier that includes the transition stages into and out of a seated state is proposed in this study. Young and senior participants wore a Blackberry Z10 smartphone on their right front waist for evaluation and completed a continuous series of 16 daily tasks. The data from the Z10 accelerometer and gyroscope was processed with a proprietary Human Activity Recognition classifier that classified immobile states based on past state awareness and transition recognition.

The findings of immobile state classification were compared with and without transition identification and previous state awareness (WT) and (WOT). WT's sit sensitivity and F- score were much higher than WOT's. For seniors, the stand specificity and F-score for WT were much higher than for WOT. In the young group, WT sit sensitivity was slightly higher than WOT, but not significantly so. For the young population, all outcomes improved. These findings suggest that studying the period leading up to an immobile state can help with immobile state recognition. Without the use of computationally demanding feature spaces or classifiers, sit-stand categorization on a continuous daily activity data set was like the current literature. The ability to comprehend this data and make improvements in our daily lives grows as more sensors are put into mobile phones to measure our movements, positions, and orientation.

The goal of our project is to analyse sensor data from mobile phones in the context of activity recognition. Our goal is to develop a model that uses sensor data to reliably classify whether a person is sleeping, sitting, bending, and standing. Studying activity recognition has several advantages and opens a slew of new possibilities. Mobile health apps that track a user's activities over time can be useful for geriatric care or personal health tracking. This research includes linkages to several fields of study, including medical, human-computer interaction, and sociology, in addition to offering psychological assistance. Human activity recognition research is in high demand because of its applications in health care, computer vision, household safety, and robot learning. If sensors collect and monitor patient data, a significant amount of money can be saved. In the event of any odd behaviour, the system can transmit reports to the doctor automatically.

We identified human behaviours using sensors from low-cost Smartphones. Smartphones have grown in popularity, accessibility, and computing capacity, making them a perfect choice for non-invasive body-worn sensors. Smartphones have become an indispensable element of everyday life. Smartphones are carried by many people throughout the day. As a result, smartphone sensors may collect data and the system can recognise human behaviour. Activity of Humans The goal of recognition is to recognise a person's activities based on a set of observations of him or her and the surrounding environment. Information retrieved from numerous sources, such as ambient or body-worn sensors, can be used to perform recognition. Our goal is to categorise the actions in a dataset into four categories: sleeping, sitting, bending, and standing. We offer a study of a method for identifying activities using data from a gyroscope and accelerometer, such as sleeping or standing.

A depiction of the data informs the analysis. We analyse the differences in error rates between different methods. Human motion recognition based on vision is a method for understanding and analysing the movement of individuals in video collected footage. Biomechanics, Machine Vision, Image Processing, Artificial Intelligence, and Pattern Recognition are among the fields covered. It is a multidisciplinary field with broad applications in social, commercial, and educational settings. Human motion recognition is required in a wide range of applications. Sports, medical, surveillance, content-based video storage and retrieval, man-machine interfaces, video conferencing,
art and entertainment, and robotics are among the applications. Different human motion depiction and recognition algorithms are used in a wide range of

applications. The number of body parts involved, and the duration of movement are determined through the application of human motion analysis. In most cases, human-computer interaction is limited to hand movements, although more complex activities or applications, such as sports or dancing, may need the use of all body parts. Human motion is conceptually divided into gestures, actions, activity, interactions, and group activities, depending on its complexity.

The tracking and initialization of the human body in video determines representation and recognition approaches. 2-D kinematic or stick figure, 3-D kinematic or form model, and picture model are three broad approaches to representation. Humans are represented by attributes such as the number of joints, degree of freedom, and limb length when they are initialised using the Kinematic technique. Humans are represented as images in image models, and properties such as shape and region are retrieved and preserved. Furthermore, the recognition can be determined by the motion's representation as well as its complexity. Simple actions employ single-layered sequential or space-time techniques. Multi-layered techniques are required for complicated actions.

The study of human behaviour recognition using depth pictures has gotten a lot of press in recent years. Because depth picture pixels capture distance information and are colour-independent, they avoid the above-mentioned difficulties to some extent as compared to standard optical images. Numerous academics have merged the nature of depth images and used many traditional algorithms to such images as optoelectronic technology has advanced.

This research offers a generative adversarial network to handle the problem of massive deformations of body parts while also considering the complexity of various layers of body parts. The internal structure of the proposed generator and discriminator in the article has been optimized, and it can imitate the hierarchical link between body parts. To standardize the interaction between parents and children, hierarchical perception terms are also added in the objective function. Hierarchical adversarial networks aid in the correct estimation of the positions of various body components, particularly those that are distorted or heavily obstructed.

II. LITERATURE SURVEY

Human activity recognition aims to detect a person's actions using a set of data about him or her and his surroundings. Human activity recognition is a field of study in which human behaviour is deduced from traits derived from movement, location, physiological signals, and information from the environment, among other things. The information generated by the environment and sensors worn by the person is used to understand the activity. Sensors worn on the waist, wrist, chest, and thighs provide excellent precision. However, these sensors are inconvenient and do not give long-term solutions.

Human Activity Recognition is a subset of the broader notion of context-aware computing, often known as ubiquitous computing. Ubiquitous computing works in a similar way to Human Activity Recognition in that it collects data from users and assists them. In the fields of
nursing, military, entertainment, and everyday life, Human Activity Recognition is widely used.

Human Activity Recognition, for example, can be used to aid soldiers with their action reports, to report a person's irregularity to hospital care, and so on. Athletes will benefit from the use of integrated motion sensors, which will supply them with vital data that will help them enhance their performance. Human Activity Recognition is clearly becoming an important aspect of our daily lives. Smartphone-based Human Activity Recognition has several advantages, including device portability, comfort, and unobstructed sensing. The disadvantage of this technique is that it uses and shares services with other apps on the phone, which could be an issue for low-resource handsets.

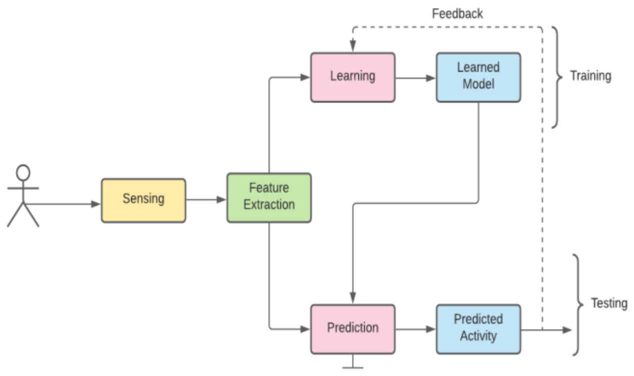


Fig1. Human Activity Recognition Development Process

In the previous two decades, there has been a lot of progress in human motion recognition and analysis, and there is a lot of material available in the form of journals, transaction papers, patents, reviews, and surveys. The researchers used criteria such as the type of models and the dimensionality of the tracking space to classify past work in the subject (Gavrila, 1999 & Poppe 2007). Some evaluations categorise the literature according to the complexity of the action to be identified (Aggarwal & Rhoo, 2011). For the classification of the available literature, some surveys employ sensor modality, sensor multiplicity, various applications, number of humans, number of monitored limbs, and assumptions (Aggarwal & Cai 1997; Aggarwal & Nandhakumar 1998; Moeslund, T. B., & Granum, 2006; Morris & Trivedi, 2008).

Aggarwal has been working in the human motion recognition domain since the 1970s, and he has published periodic evaluations of the sector with a large amount of work. Human motion recognition approaches are categorised into two types in his latest review study with M. S. Rhoo (2011): single layered approaches and hierarchical approaches. Cedras and Shah (1995) described motion-based recognition as consisting of two steps: the extraction of motion information in the first phase, and the matching of an unknown input with the created model in the second.
The literature has been divided into three categories by Gavrila (1999):

(a) 2-D approaches without explicit shape models,
(b) 2-D approaches with explicit form models.
(c) Three-dimensional techniques in the form of two survey articles, T. B. Moeslund (2001, 2006) has provided a compressive assessment of publications in computer vision-based human motion analysis.

III. METHODOLOGY

In this section, we'll go over the steps we took to create our system. The machine learning problem we're working on fits into the classification challenge category. To carry out the full experiment, we used R Studio version 3.2.5 on a Windows 10 platform.

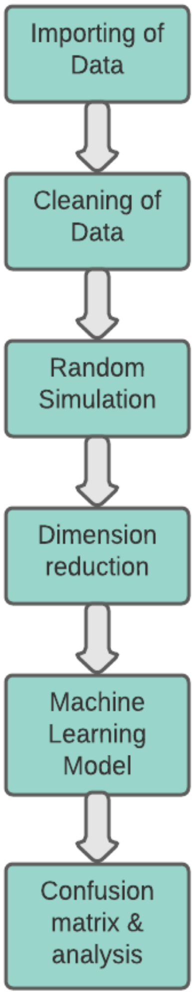


Fig.3 Flow of System Design of Human Activity Recognition

Cleaning and Importing Data

Data was gathered from a group of 30 individuals who participated in an experiment. Six activities were required of each participant (sleeping, sitting, bending, and standing). To acquire data, they employed embedded sensors from a Samsung Galaxy SII smartphone. As a comma separated file, the dataset was imported into R studio. There were 586 features in the imported dataset.

The following are the dataset's original characteristics:

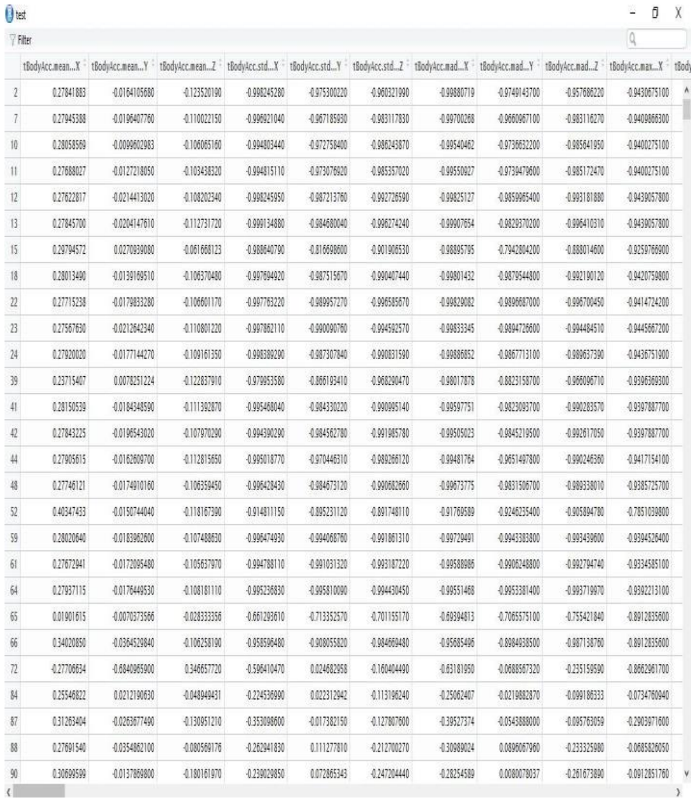
1. A 561-feature vector with temporal and frequency domain variables is included in the dataset.
2. Its designation as an activity.
3. A unique identity for the person who conducted the experiment.

The names of the features were changed and utilised as headers. For that feature, missing data and NA values were imputed using the mean.

The following formula was used to standardise the data:

$$zi = \frac{xi - \bar{x}}{sd}$$

Where xi is an instance of data, $\bar{x}$ is mean of that feature column and sd is standard deviation of that feature column.



	tBodyAcc.mean...X	tBodyAcc.mean...Y	tBodyAcc.mean...Z	tBodyAcc.std...X	tBodyAcc.std...Y	tBodyAcc.std...Z	tBodyAcc.mad...X	tBodyAcc.mad...Y	tBodyAcc.mad...Z	tBodyAcc.max...X
2	0.27841883	-0.0164105680	-0.123520190	-0.998245280	-0.975300220	-0.960321990	-0.99880719	-0.9749143700	-0.957686220	-0.9430675100
7	0.27945388	-0.0196407760	-0.110022150	-0.996921040	-0.967185930	-0.983117830	-0.99700268	-0.9660967100	-0.983116270	-0.9409866300
10	0.28058569	-0.0099602983	-0.106065160	-0.994803440	-0.972758400	-0.986243870	-0.99540462	-0.9736632200	-0.985641950	-0.9400275100
11	0.27688027	-0.0127218050	-0.103438320	-0.994815110	-0.973076920	-0.985357020	-0.99550927	-0.9739479600	-0.985172470	-0.9400275100
12	0.27622817	-0.0214413020	-0.108202340	-0.998245950	-0.987213760	-0.992726590	-0.99825127	-0.9859965400	-0.993181880	-0.9439057800
13	0.27845700	-0.0204147610	-0.112731720	-0.999134880	-0.984680040	-0.996274240	-0.99907654	-0.9829370200	-0.996410310	-0.9439057800
15	0.29794572	0.0270939080	-0.061668123	-0.988640790	-0.816698600	-0.901906530	-0.98895795	-0.7942804200	-0.888014600	-0.9259766900
18	0.28013490	-0.0139169510	-0.106370480	-0.997694920	-0.987515670	-0.990407440	-0.99801432	-0.9879544800	-0.992190120	-0.9420759800
22	0.27715238	-0.0179833280	-0.106601170	-0.997763220	-0.989957270	-0.996585670	-0.99829082	-0.9896687000	-0.996700450	-0.9414724200
23	0.27567630	-0.0212642340	-0.110801220	-0.997862110	-0.990090760	-0.994592570	-0.99833345	-0.9894726600	-0.994484510	-0.9445667200
24	0.27920020	-0.0177144270	-0.109161350	-0.998389290	-0.987307840	-0.990831590	-0.99886852	-0.9867713100	-0.989637390	-0.9436751900
39	0.23715407	0.0078251224	-0.122837910	-0.979953580	-0.866193410	-0.968290470	-0.98017878	-0.8823158700	-0.966096710	-0.9396369300
41	0.28150539	-0.0184348590	-0.111392870	-0.995468040	-0.984330220	-0.990995140	-0.99597751	-0.9823093700	-0.990283570	-0.9397887700
42	0.27843225	-0.0196543020	-0.107970290	-0.994390290	-0.984562780	-0.991985780	-0.99505023	-0.9845219500	-0.992617050	-0.9397887700
44	0.27905615	-0.0162609700	-0.112815650	-0.995018770	-0.970446310	-0.989266120	-0.99481764	-0.9651497800	-0.990246360	-0.9417154100
48	0.27746121	-0.0174910160	-0.106359450	-0.996428430	-0.984673120	-0.990682660	-0.99673775	-0.9831506700	-0.989338010	-0.9385725700
52	0.40347433	-0.0150744040	-0.118167390	-0.914811150	-0.895231120	-0.891748110	-0.91769589	-0.9246235400	-0.905894780	-0.7851039800
59	0.28020640	-0.0183962600	-0.107488630	-0.996474930	-0.994068760	-0.991861310	-0.99729491	-0.9943383800	-0.993439600	-0.9394526400
61	0.27672941	-0.0172095480	-0.105637970	-0.994788110	-0.991031320	-0.993187220	-0.99588986	-0.9906248800	-0.992794740	-0.9334585100
64	0.27937115	-0.0176449530	-0.108181110	-0.995236830	-0.995810090	-0.994430450	-0.99551468	-0.9953381400	-0.993719970	-0.9392213100
65	0.01901615	-0.0070373566	-0.028333356	-0.661293610	-0.713352570	-0.701155170	-0.69394813	-0.7065575100	-0.755421840	-0.8912835600
66	0.34020850	-0.0364529840	-0.106258190	-0.958596480	-0.908055820	-0.984669480	-0.95685496	-0.8984938500	-0.987138760	-0.8912835600
72	-0.27706634	-0.6840965900	0.346657720	-0.596410470	0.024682958	-0.160404490	-0.63181950	-0.0688567320	-0.235159590	-0.8662961700
84	0.25546822	0.0212190630	-0.048949431	-0.224536990	0.022312942	-0.113196240	-0.25062407	-0.0219882870	-0.099186333	-0.0734760940
87	0.31263404	-0.0263677490	-0.130951210	-0.353098600	-0.017382150	-0.127807600	-0.39527374	-0.0543888000	-0.095763059	-0.2903971600
88	0.27691540	-0.0354862100	-0.080569176	-0.262941830	0.111277810	-0.212700270	-0.30989024	0.0896067960	-0.233325980	-0.0685826050
90	0.30699599	-0.0137869800	-0.180161970	-0.239029850	0.072865343	-0.247204440	-0.28254589	0.0080078037	-0.261673890	-0.0912851760

Fig. 3 Snapshot of cleaned Dataset

Simulation at Random

Random Simulation is a technique for determining the accuracy of predictive models and preventing data overfitting and underfitting. The method entails randomly partitioning the dataset into seven training and three testing sets in a 7:3 ratio. According to the statistics Central limit theorem, the entire simulation is repeated 50 times to increase model correctness. The testing dataset gives us with a close approximation of real-time data as well as a mechanism to assess the model's stability in a real-world setting.

Dimensionality Reduction

To minimise the number of features in our dataset, we performed Principal Component Analysis. The purpose of principle component analysis is to minimise the dimensionality of a large data set with many connected variables while retaining as much variance as possible. Principal component analysis transforms the variables in a dataset into a new collection of variables called the principal components. The Principal Components are uncorrelated and arranged according to the amount of variance in all the original variables. The Eigen value decomposition of data co-variance reveals this behaviour.

Random Forest

Random Forest is a method for classification, regression, and other tasks that uses an ensemble learning approach. It works by building many decision trees during training and then producing a class that is the mean forecast of those trees. Overfitting of data is avoided using a random forest. 500 decision trees are generated by default in the R programming environment. We discovered that there was no need to build 500 decision trees because of our experiments.

K-Nearest Neighbour (KNN)

KNN is a classifier that is based on instances. It works on the premise that classification of unknown instances can be accomplished by linking the unknown instance to a known instance using some function. This function is a distance or similarity function.

To approximate our learning function, we employed the Euclidean distance function. The value of K was established by producing a graph of mistake rate vs. K value. We discovered that the error rate reduces from K=40 to K=65 and reaches a minimum at K=65.

Support Vector Machine

Decision hyperplanes define decision boundaries, and Support Vector Machines are built on them. A decision plane divides two groups of objects that belong to different classes. The goal of a Support Vector Machine is to increase the decision boundary between hyperplanes as much as possible. To train our dataset, we utilised the “e1071” SVM library in R. Kernel type, cost, and gamma are all parameters in the SVM method. As a similarity function, we picked the Gaussian(radical) kernel.

The cost value is set to 1 to keep the regularisation term constant and avoid data overfitting. The shape of the vector hyperplane is determined by the gamma value. Because the value of gamma is equal to 1/, it is retained at 0.013758. (number of features).

Artificial Neural Network
Artificial Neural Networks use a model developed on our understanding of how a human brain reacts to stimuli from inputs to describe the relationship between a set of input signals and an output signal. To train our dataset, we utilised the nnet package, which is for feed-forward neural networks with a single hidden layer. The backpropagation method is used to train the artificial neural network in the nnet package. This approach determines the output error and then propagates it back into the network. The weights are modified to minimise the mistake caused by each neuron. Because we're dealing with a classification problem, the Linout argument was set to False by default.

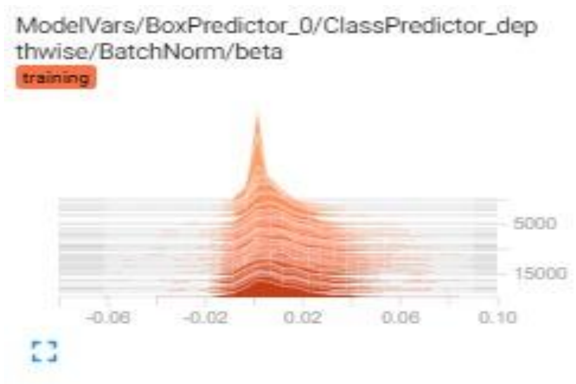
Prediction of Testing Dataset
After generating the Principal Components of the training set, we must use these Principal Components to predict testing data. This appears to be quite straightforward, but there are a few things to keep in mind:

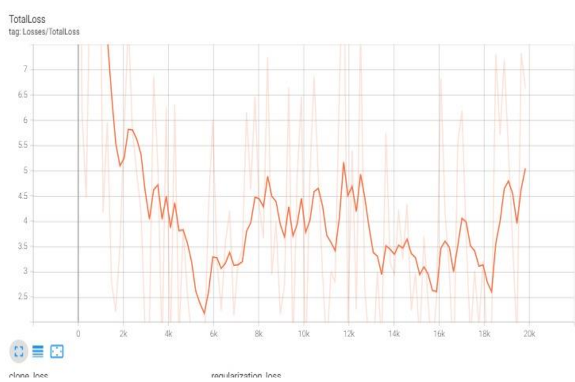
1. Applying PCA to the entire dataset at once will "leak" characteristics from the training dataset into the testing dataset, affecting our model's prediction abilities.
2. When PCA is applied to the testing and training datasets individually, the vectors will have different directions. As a result, the results will be incorrect.

The testing set will be transformed in the same way as the training set, with the same centre and scaling feature.

IV. Conclusion

We learned Machine Learning Algorithms such as Random Forest, K-Nearest Neighbour, Support Vector Machine, and Artificial Neural Networks after studying the Human Activity Recognition dataset. Using Principal Component Analysis, we were able to minimise the number of features in our dataset from 586 to 100. After conducting data analysis, we discovered that the Support Vector Machine was the most effective at predicting Human Activities. Because we utilised a Gaussian kernel with a smaller dataset, SVM is the most efficient. Because it is the simplest and uses the Euclidian distance function, K-Nearest Neighbour took the least amount of time to train the machine learning model. Support Vector Machine took the longest to train the model.





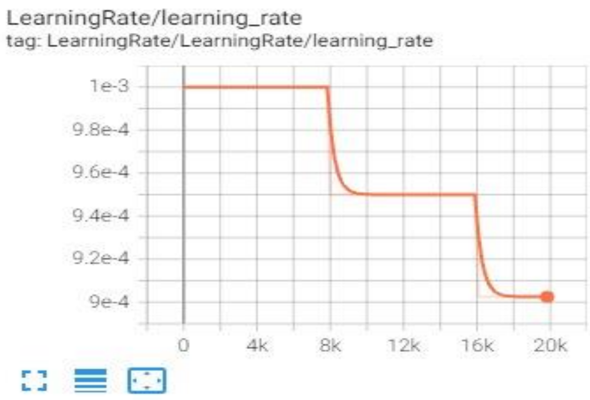


Fig 1. Graphical Representations of the Model

Aspects in the Future: The model developed can be used to anticipate human activity in real time. To send measurements and execute the model on those measurements, an Android application might be used. This application has features for tracking athlete health and performance, among other things. We discovered that using Principal Component Analysis to analyse data resulted in the production of optimised models, which improved performance.

The recognition of human action from video content is an important topic of computer vision research, and it has made great progress. The research addressed various representation and recognition systems that were categorised according to the level of action complexity. For the recognition of sophisticated actions and interactions, hierarchical techniques have had a lot of success. Techniques like bag-of-words and HMM, which have proven successful in voice and text recognition, are also being used to recognise actions. Human motion recognition requires advances in domains such as Artificial Intelligence and Machine Learning.

This research offers a created confrontation network based on stacked hourglasses that can deduce the structure and hierarchy of various bodily parts without explicitly doing so. The suggested network in this paper demonstrates the ability to learn to predict the spatial relationship between body parts. It's been trained from start to finish and put through its paces on three different benchmark data sets. The network has demonstrated its capacity to estimate malformed and obstructed body parts. This paper's solution can overcome these challenges and improve performance.

It also collects and compares the most recent experimental results from a range of data sets. Furthermore, the network suggested in this paper can be used to encode other related properties between nearby body parts. Furthermore, the network can process several hierarchical layers of information between body parts. Our algorithm is currently more sensitive to light, which is also one of its limitations. Our next task is to address the issue of the algorithm's sensitivity to light.

Fig 2. Output of standing human action

Fig. 3 Output of Sitting human action

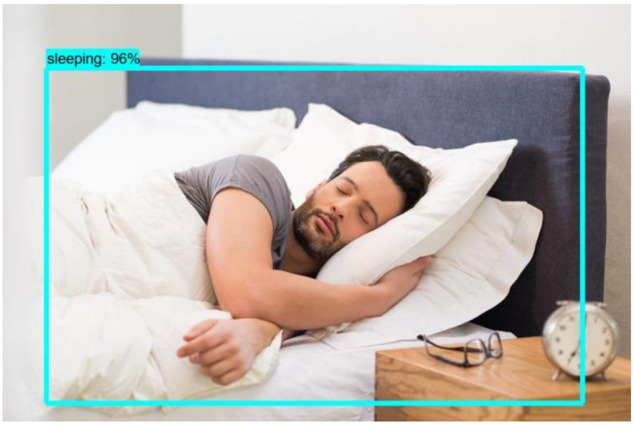


Fig. 3 Output of Sleeping human action

Fig. 3 Output of Bending human action

V. References

[1] Adil Mehmood Khan, Young-Koo Lee, Sungyoung Y. Lee, and Tae-Seong Kim,A triaxial accelerometer-based physical - activity recognition via augmented-signal features and a hierarchical recognizer, Information Technology in Biomedicine, IEEE transactions on information technology in biomedicine, Volume 14, NO. 5, September 2010

[2] Murat EKINCI, Eyup GEDIKLI, 2005. Silhouette Based Human Motion and Action Detection and Analysis for RealTime Automated Video Surveillance. Turk J Elec Engin. Volume 13, No.2.

[3] Nazh Ikizler and Põnar Duygulu, 1999. Human Action Recognition Using Distribution of Oriented Rectangular Patches. Computer vision and pattern recognition (CVPR.05).Volume1, pp 886-893.

[4] Sminchisescu, C., & Triggs, B. (2003). Estimating articulated human motion with covariance scaled sampling. The International Journal of Robotics Research, 22(6), 371–391. doi:10.1177/0278364903022006003

[5] Aggarwal, J. K., & Cai, Q. (1997, June). Human motion analysis: A review. Proceedings of theNonrigid and Articulated Motion Workshop '97 (pp. 90-102). IEEE. doi:10.1109/NAMW.1997.609859

[6] Aggarwal, J. K., & Nandhakumar, N. (1988). On the computation of motion from sequences of images-a review. Texas Univ at Austin.

[7] Aggarwal, J. K., & Ryoo, M. S. (2011). Human activity analysis: A review. ACM Computing Surveys, 43(3), 16. doi:10.1145/1922649.1922653

[8] Bertini, M., Del Bimbo, A., & Nunziati, W. (2003, November). Model checking for detection of sport highlights. Proceedings of the 5th ACM SIGMM international workshop on Multimedia information retrieval (pp. 215-222). ACM. doi:10.1145/973264.973298

[9] Hoffmann, M., Marques, H. G., Arieta, A. H., Sumioka, H., Lungarella, M., & Pfeifer, R. (2010). Body schema in robotics: A review. IEEE Transactions on Autonomous Mental Development, 2(4), 304–324.

[10] Moeslund, T. B., Hilton, A., & Krüger, V. (2006). A survey of advances in vision-based human motion capture and analysis. Computer Vision and Image Understanding, 104(2), 90–126. doi: 10.1016/j.cviu.2006.08.002

[11] Obdrzalek, S., Kurillo, G., Ofli, F., Bajcsy, R., Seto, E., Jimison, H., & Pavel, M. (2012, August). Accuracy and robustness of Kinect pose estimation in the context of coaching of elderly population. Proceedings of the 2012 annual international conference of the IEEE on Engineering in medicine and biology society (EMBC) (pp. 1188-1193). IEEE. doi:10.1109/EMBC.2012.6346149

[12] Rohr, K. (1994). Towards model-based recognition of human movements in image sequences. CVGIP. Image Understanding, 59(1), 94–115. doi:10.1006/ciun.1994.1006

[13] Saba, L., Dey, N., Ashour, A. S., Samanta, S., Nath, S. S., Chakraborty, S., & Suri, J. S. (2016). Automated Stratification of Liver Disease in Ultrasound: An Online Accurate Feature Classification Paradigm. Computer Methods and Programs in Biomedicine, 130, 118–134. doi: 10.1016/j.cmpb.2016.03.016 PMID:27208527

[14] Yamato, J., Ohya, J., & Ishii, K. (1992, June). Recognizing human action in time-sequential images using hidden markov model. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR '92 (pp. 379-385). IEEE. doi:10.1109/CVPR.1992.223161

[15] Yang, X., & Tian, Y. (2014). Effective 3d action recognition using eigenjoints. Journal of Visual Communication and Image Representation, 25(1), 2–11. doi: 10.1016/j.jvcir.2013.03.001

[16] Campbell, L. W., & Bobick, A. E. (1995, June). Recognition of human body motion using phase space constraints. Proceedings of the Fifth International Conference on Computer Vision (pp. 624-630). IEEE. doi:10.1109/ ICCV.1995.466880

[17 Gorelick, L., Blank, M., Shechtman, E., Irani, M., & Basri, R. (2007). Actions as space-time shapes. IEEE Transactions on Pattern Analysis and Machine Intelligence, 29(12), 2247–2253. PMID:17934233

[18] Hoffmann, M., Marques, H. G., Arieta, A. H., Sumioka, H., Lungarella, M., & Pfeifer, R. (2010). Body schema in robotics: A review. IEEE Transactions on Autonomous Mental Development, 2(4), 304–324.