IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Health Care Monitoring System using Action Recognition

Paka Akshaja, Pushpendra Kumar Pateriya

School of Computer Science and Engineering

Lovely Professional University

Phagwara, Punjab

Abstract:

Video Surveillance is the most prominent area in the present world for enhancing security and safety. It has been explaining the presence of technology for taking care of elderly people with the help of computer vision, sensors, and machine learning techniques. This work can enhance the safety of old people at home. In this work, a monitoring system has been proposed by the modest method by using computer vision and visual sensors such as cameras. There are various methods for action recognition, every technique has pros and cons. In this paper, we propose human activity recognition and classification using Pose estimation. We compile various algorithms' accuracy on the dataset. It helps us to keep track of the overall health condition of the elder in real-time and ease getting feedback and support from healthcare services. We will be performing a video analysis on the datasets for action recognition. Finally, we develop a computer-based application for monitoring elderly people in their homes.

Key Words:

Human Activity Recognition, Pose Estimation, Body Key points

1. Introduction

The primary aim of this research will be concerned with the video monitoring of the person and privacy living issues for the successful detection and generation of alert messages in case of fall detection and health status, which will allow the prompt detection and delivery of medical emergency to mitigate the long-term effects of fatal and non-fatal attacks.

This topic was inspired by many useful real-time applications like understanding human behavior, video surveillance, etc. Action Recognition using smartphones and wearable sensors is common; a major disadvantage of the wearable sensor-based system is that the elder may forget to wear or feel discomfort when attach to the body [1]. Another

commonly used technique is the vision-based detection system.

One of the popular vision-based systems is HAR which uses human pose information to extract and analyze the action of the people. We extract human pose using 18 body key points located in a 2Dimensional Plane from videos and images using Open Pose Library. Ultimate, activity is classified using pose information through a supervised machine learning algorithm.

The rest of the paper is structured as in Section 2 literature survey of some research papers. Section 3 has the method and architecture of the proposed approach. A brief description of the dataset and evaluation metrics.

2. Related Work

In this section, several approaches for fall detection described in the literature are discussed.

Four types of approaches are addressed for the classification of actions, including image structure-based methods, pose-based models, model-based approaches, and example-based methods.

The Pose-based method trains each pose using an annotated 3D image [2]. The model-based method uses a known parametric body model to match posture variables [3]. The example-based model uses classical machine learning algorithms to find actions. In the method based on image structure, the posture represents features in the classification of the action [4].

[5] Applied open pose and Kalman filter to track the target body, and then a one-dimensional full CNN is used for the classification of activity. [6] proposed a novel approach to activity recognition by simultaneously extracting features from objects used to perform the activity and human posture.

Moreover, a single person activity can also be recognized by using smartphone sensors and wearable sensors; the smartphone-based approach uses sensors that are inbuilt into the device, such as an accelerometer and gyroscope, to identify activity, whereas the wearable sensor-based approach requires the sensors to be attached on the subject body to collect action information.

[7] used wearable and smartphone embedded sensors for detecting six dynamic and six static activities using a machine learning algorithm.

3. Dataset

Open Pose uses the COCO key points detection dataset for the pose estimation task.

UCI HAR database was built from the recordings of 30 subjects performing activities while carrying a smartphone with embedded inertial sensors. Using this embedded accelerometer and gyroscope, we captured 3-axial linear acceleration and 3-axial angular velocity. We divided our dataset into training and testing in the ratio of 70:30

The Video monitoring UR Fall detection dataset has 70 sequences (30 falls + 40 activities in daily life). Fall events are recorded by Kinect Camera and accelerometric data. For real-time action recognition, we use a webcam from the system.

4. Proposed Approach

This work focuses on improving the performance of fall detection and activity recognition using video data, applications such as assisted living and public space monitoring, and the use of wearables and other sensor modalities may be possible. The results can be significantly improved by using articulated bodies(skeletons) extracted from videos.

4.1 Feature Extraction Process

Human Pose detection is done by using Open Pose an open-source method for pose estimator. It uses RGB images as input, then performs pose detection with a pre-trained model (COCO dataset), outputting for each processed image. The extracted pose key points as input with the help of classification algorithms such as logistic regression, SVC, SVM, Decision Tree, and Random Forest.

4.2 Pose Estimation

Before training, the skeleton sequences are pre-processed to remove empty frames. In this work, the COCO dataset has 18 body key points. R_Ankle, R_Knee, R_Wrist, L_Wrist, R_Shoulder, L_Shoulder, L_Ankle, L_Ear, R_Ear, R_Elbow, L_Elbow, L_Knee, L_Eye, R_Eye, R_Hip, L_Hip, Nose, and Neck is used [8].

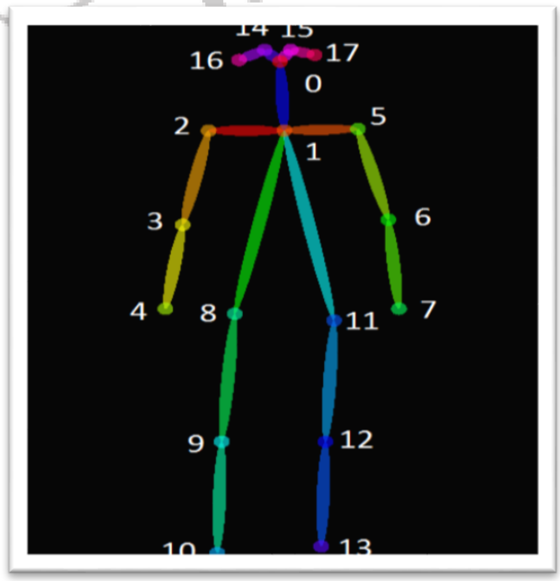


Figure 1: Open Pose Key Points

4.3 Activity Classification

The activity classification problem is a multiclass classification problem, which can be modeled using various machine learning regression and classification algorithms. The classification algorithm takes 18 body key points (x-axis and y-axis values of each point) as input for our model's training and testing. We used a supervised learning approach as our dataset has body key points with an activity label.

5. Evaluation Metrics

For performance evaluation, we use Recall, Precision, and F1-score. We will also show the confusion matrix and classification report of some classifiers.

5.1 Recall

Recall (R) is the ratio of the number of true positives (Tp) to the sum of false negatives (Fn) and true positives. It can also be defined as how many of the correct hits were also found.

R = Tp/(Tp+Fn) [9]

5.2 Precision

Precision (P) is the ratio of the number of true positives (Tp) to the sum of false positives (Fp) and true positives. It can also be defined as how many of the returned hits were true positive.

P = Tp/(Tp+Fp) [10]

5.3 F1-Score

F1 Score is calculated as the harmonic mean of recall and precision.

F1-Score = 2 (P*R)/(P+R) [11]

5.4 Confusion Matrix

A Confusion Matrix is a metric tool for evaluating the results of a classification machine learning model. In the matrix, each row holds the value of the predicted class, and each column holds the value of the actual class. The matrix compares the target with the predicted by the model. So, it gives a better idea of errors by the classifier.

6. Experiment and Results

The following six activities are considered for pose estimation, activity recognition, and classification: walking, walking upstairs, walking downstairs, sitting, standing, and laying. Each Pose frame is proposed independently to detect whether it represents a fall. These algorithms are described below with their confusion matrix. The performance results are provided, which show the recall, precision, f1-score, accuracy, and error of various classifiers used in the proposed approach.

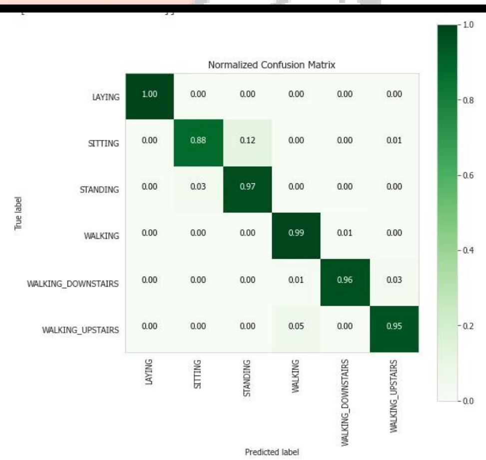
	Accuracy	Error
Logistic Regression	: 95.86%	4.14%
SVC classifier	: 96.71%	3.291%
SVM classifier	: 96.27%	3.733%
DecisionTree	: 87.68%	12.32%
Random Forest	: 92.4%	7.601%

Figure 2: Performance Result

6.1 Classification Algorithm

6.1.1 Logistic Regression

This algorithm is based on supervised learning, and it is used in classification problems [12]. It helps to describe data and explain the relationship between the binary variable and independent variable.



Classifiction Report				
	precision	recall	f1-score	support
LAYING	1.00	1.00	1.00	537
SITTING	0.96	0.88	0.92	491
STANDING	0.90	0.97	0.93	532
WALKING	0.95	0.99	0.97	496
WALKING_DOWNSTAIRS	0.99	0.96	0.97	420
WALKING_UPSTAIRS	0.96	0.95	0.96	471
accuracy			0.96	2947
macro avg	0.96	0.96	0.96	2947
weighted avg	0.96	0.96	0.96	2947

Figure 3: Confusion Matrix (Logistic Regression)

6.1.2 SVC

It also comes under supervised learning algorithms. It is to fit the class you provide, returning a "best fit" hyperplane that divides or categorizes, your class.

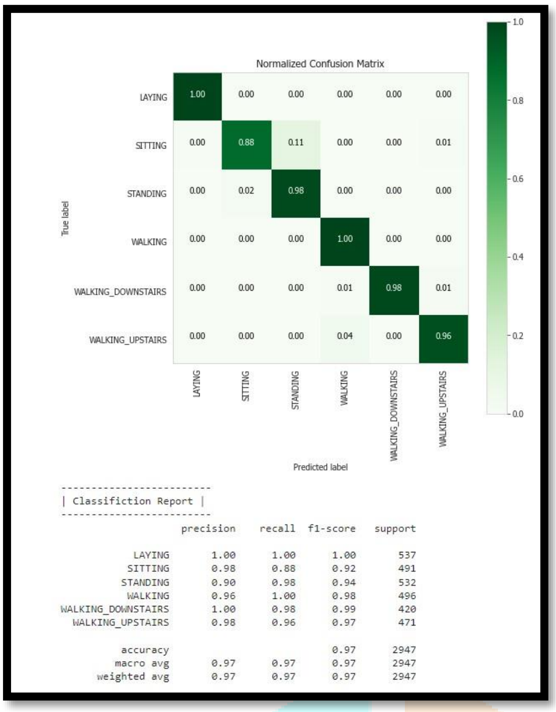


Figure 4: Confusion Matrix (SVC)

6.1.3 SVM

It also comes under supervised learning algorithms and is mainly used in classification and regression problems. The classification is done by finding a hyperplane that supplies similarly different outputs between the two classes.

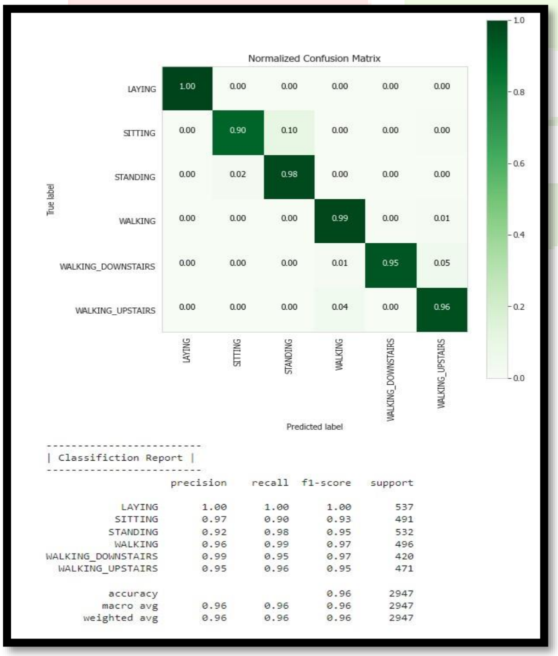


Figure 5: Confusion Matrix (SVM)

6.1.4 Decision Tree

The decision tree comes under supervised learning. It is the most powerful and accepted tool for prediction and classification. This algorithm uses learning to predict a target pose's activity and make decisions from previously trained data. Predictions for activities are made from the root of the tree.

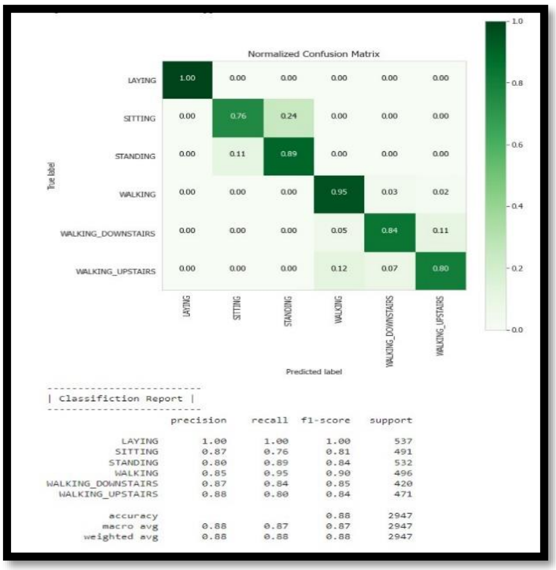


Figure 6: Confusion Matrix (Decision Tree)

6.1.5 Random Forest Classifier

Random forest is a supervised learning algorithm, and it is an ensemble learning classifier. It is also one of the most used and popular algorithms because it gives better results without tuning hyper-parameter. There is accurate decision tree practice of overfitting to their training set.

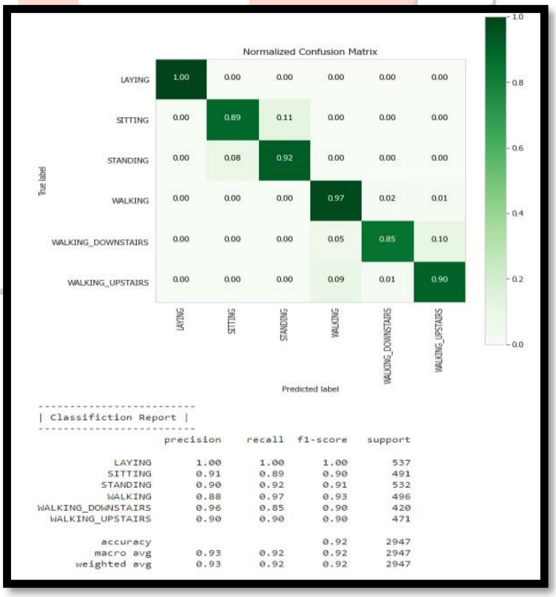


Figure 7: Confusion Matrix (Random Forest Classifier)

7. Conclusion

In this paper, a camera-vision-based fall detection along with human action recognition uses the pose information to classify activity with the help of a supervised machine learning algorithm. There is not yet a standard dataset on which action recognition and fall detection systems can be evaluated and compared. We prepared our dataset for this work which holds six various activities, viz, walking, walking downstairs, walking upstairs, standing, sitting, and laying. We have used five algorithms to find better results for our model.

The main advantage of this work recognizes human activity by using a person's posture on a video image. The posture can be defined by key points being the main joints of the human skeleton. The detection of elderly falls is an example of the potential of health monitoring systems. The focus here was on elderly people, the same and similar systems can apply to mobility problems.

8. Reference

1. Gupta, A., Gupta, K., Gupta, K., & Gupta, K. (2020). A Survey on Human Activity Recognition and Classification. *2020 International Conference on Communication and Signal Processing (ICCSP)*, 0915–0919. https://doi.org/10.1109/ICCSP48568.2020.9182416
2. Sharma, G., Jurie, F., & Schmid, C. (2013). Expanded Parts Model for Human Attribute and Action Recognition in Still Images. *2013 IEEE Conference on Computer Vision and Pattern Recognition*, 652–659. https://doi.org/10.1109/CVPR.2013.90
3. Gupta, A., Kembhavi, A., & Davis, L. S. (2009). Observing human-object Interactions: Using Spatial and Functional Compatibility for Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *31*(10), 1775–1789. https://doi.org/10.1109/TPAMI.2009.83
4. Shotton, J., Fitzgibbon, A., Cook, M., Sharp, T., Finocchio, M., Moore, R., Kipman, A., & Blake, A. (2011). Real-time human pose recognition in parts from single depth images. *CVPR 2011*, 1297–1304. https://doi.org/10.1109/CVPR.2011.5995316
5. de Miguel, K., Brunete, A., Hernando, M., & Gambao, E. (2017). Home camera-based fall detection system for the elderly. *Sensors (Switzerland)*, *17*(12). https://doi.org/10.3390/s17122864
6. Reily, B., Zhu, Q., Reardon, C., & Zhang, H. (2020). Simultaneous Learning from Human Pose and Object Cues for Real-Time Activity Recognition. *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 8006–8012. https://doi.org/10.1109/ICRA40945.2020.9196632
7. Nandy, A., Saha, J., Chowdhury, C., & Singh, K. P. D. (2019). Detailed Human Activity Recognition using Wearable Sensor and Smartphones. *2019 International Conference on Opto-Electronics and Applied Optics (Optronix)*, 1–6. https://doi.org/10.1109/OPTRONIX.2019.8862427
8. Cao, Z., Simon, T., Wei, S.-E., & Sheikh, Y. (2017). Realtime Multi-person 2D Pose Estimation Using Part Affinity Fields. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1302–1310. https://doi.org/10.1109/CVPR.2017.143
9. Gupta, A., Kembhavi, A., & Davis, L. S. (2009). Observing human-object Interactions: Using Spatial and Functional Compatibility for Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *31*(10), 1775–1789. https://doi.org/10.1109/TPAMI.2009.83
10. Yang Wang, Hao Jiang, Drew, M. S., Ze-Nian Li, & Mori, G. (n.d.). Unsupervised Discovery of Action Classes. *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume 2 (CVPR'06)*, 1654–1661. https://doi.org/10.1109/CVPR.2006.321
11. Ramirez, H., Velastin, S. A., Meza, I., Fabregas, E., Makris, D., & Farias, G. (2021). Fall Detection and Activity Recognition Using Human Skeleton Features. *IEEE Access*, *9*, 33532–33542. https://doi.org/10.1109/ACCESS.2021.3061626
12. Gupta, A., Gupta, K., Gupta, K., & Gupta, K. (n.d.). *Human Activity Recognition Using Pose Estimation and Machine Learning Algorithm*.
13. Thomas, M., & Dileep, N. (n.d.). *Review of Fall Detection and Alert Systems for Elderly People*. www.ijert.org
14. Foroughi, H., Aski, B. S., & Pourreza, H. (2008). Intelligent video surveillance for monitoring fall detection of elderly in home environments. *Proceedings of 11th International Conference on Computer and Information Technology, ICCIT 2008*, 219–224. https://doi.org/10.1109/ICCITECHN.2008.4803020