International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Elderly People's Abnormal Behavior Detection Using HAR and CNN Algorithms.

Sneha Bajirao Paymal[1], Mahadev S. Patil[2]

Submitted: 28/11/2023 **Revised**: 09/01/2024 **Accepted**: 20/01/2024

Abstract: Artificial intelligence machine learning systems are developing rapidly and very fast with the period with increasing demand and dependencies on them. New developments have been made for Artificial Intelligence and machine learning has made artificial brains for detection, identification, and decision-making abilities available for computer machines. The paper proposes the recognition of human health prediction by recognizing abnormal actions, or signs given by humans using OPENCV CNN, a tensor flow platform comparing live actions with different action datasets stored. The system detects falls of a person, sleep, heart pain, stomach pain, shoulder pain, dizziness, and different actions related to daily routine such as exercise, reading, writing, playing, makeup, etc recognize actions are sent to the Firebase cloud platform to be monitored by user or user recommended physician. Abnormal action will provide a warning message for help or raise an alarm for help. The system can detect action using surveillance cameras, or Pi camera, or a webcam.

Keywords: *Convolutional Neural network (CNN), Human Activity Recognition (HAR), Deep Learning (DL), firebase, IOT, OpenCV*

1. Introduction

The machine learning Python platform provides vital support for artificial intelligence-based systems providing perfect detection, recognition, and control over many applications in various fields such as medicine, industries, defense, security, area monitoring, weather forecast, atomized applications, etc. The main purpose of this paper is to implement a video surveillance system for the detection of human health or elderly people's health by detecting abnormal action gestures or abnormal signs. After detection of those actions help notification will be given to the neighborhood or doctor over the Firebase cloud platform to save lives by getting help. Human-to-human interactions, human-to-machine, human actions or activity monitoring such as embracing and handshaking, and also human-to-object interactions, such as playing guitar or ball tossing, human falling, dizziness, different body pains are the two categories in which various human actions are classified. This classification is based on human behavior, actions captured by the camera from the live feed and processed by OpenCV deep learning, tensor flow, and HAR, CNN algorithms recognize and differentiate and separate different normal actions and abnormal actions. The proposed system in this paper is implemented with HAR human action recognition with kinetics data set having accuracy of more than 90 percent which can detect 400 actions set compared with pre-trained 400 actions stored. Tensor flow CNN platform is preferred at the backend to detect health-related issues from live feeds from the camera such as heart pain, brain stress, shoulder pain, headache, stomach pain, etc. And provides notifications to doctors via the Internet Firebase cloud platform, every action after recognition is sent to the cloud platform, and saves time to monitor overtime when required. By leveraging advanced technologies, we can improve the lives of the aging population and promote healthy aging in our society

2. Related Work

For the recognition of human activity, a variety of algorithms and methodologies have been developed, such as the two-phase recognition system paradigm with probabilistic generative models that classify activities using deep belief networks. This deep belief network is in charge of reconstructing data, creating features, and categorizing human conduct or activities. It implies several Restricted Boltzmann Machines (RBMs). Methods used in these models are CNN and LCRN which are the latest and most effective methods for recognition of actions from video sequences. Here model uses pre-trained dataset kinetics-400 implies three steps training, testing, and validation. Achieved accuracy is up to 90%. [1]. Human actions also can be recognized with TELNET temporal action recognition compiled with faster r-CNN object detection framework technology from live feed and can able to recognize actions with faster speed. TAL-Net corrects three major flaws in existing methods: (1) technique increases receptive field alignment using a multi-scale architecture that can accommodate extreme variation in action durations; (2) technique appropriately extends receptive fields to take advantage of the temporal context of actions for proposal generation and action classification, and (3) technique explicitly considers multi-stream feature fusion and

[1]Research Scholar, RIT, Islampur, 415409, India
[2]Peofessor, HOD of the E&TC Department, RIT Islampur, City, 415409, India

shows that fusing motion late is important. On the THUMOS'14 detection benchmark, we achieve cutting-edge performance for both action suggestion and localization, as well as competitive performance on the Activity-Net challenge. [2]. The kinetics 400 dataset is used for action recognition by slow fast networks, which compile with various sampling rates and achieve a maximum accuracy rate of 82% when evaluated on reset-50 and resnet-101 trained datasets.

A Fast pathway that runs at high frame rates to record motion with precise temporal resolution and (ii) a Slow pathway that operates at low frame rates to collect spatial semantics. By lowering its channel capacity, the Fast route may be made very light while still learning relevant temporal information for video identification. Our Slow Fast idea is credited with making significant improvements to our models' performance in motion categorization and detection in video. Modern accuracy on the key video recognition benchmarks, Kinetics, Charades, and AVA, is reported in the paper.

Action recognition can also be useful for detecting the worst activities of violence, theft, etc. Nowadays surveillance and security issues are increasing day by day. Detection of public violence and theft can be recognized with CNN and LSTM long short-term memory python machine learning algorithm from live video feed. This paper proposes a system for hockey sport rule violation detection while playing [4]

3. Proposed Method

Technology implemented detects a person's live activities normal and abnormal falling activity, and different body pains such as heart pain, stomach, shoulder, head pain, heart rate detection, brain stress detection, etc. with machine learning CNN architecture and with HAR algorithm 400 different activity recognized such as walking, jogging, playing, eating, etc. For the recognition of human activity, a variety of algorithms and methodologies have been developed, such as the two-phase recognition system paradigm with probabilistic generative models that classify activities using deep belief networks. This deep belief network is in charge of data reconstruction, feature building, and the categorization of human conduct or activities [1]. It implies several Restricted Boltzmann Machines (RBMs). Human actions also can be recognized with TELNET temporal action recognition compiled with faster r-CNN object detection framework technology from live feed and can able to recognize actions with faster speed [2]. Slow fast network-based action recognition uses kinetics 400 datasets for recognition, slow fast network compiles with variable sampling rate with a highest accuracy rate of 82% approximately tested with reset-50 and resnet-101 trained dataset [3]. Action recognition can also be useful for detecting the worst activities of violence, theft, etc. Nowadays surveillance and security issues are increasing day by day. Detection of public violence and theft can be recognized with CNN and LSTM long short-term memory Python machine learning algorithm from live video feed [4]. This paper implies the design and implementation of elderly people's abnormal health conditions monitoring and alarm system using machine learning technology from live video feed. Live video feed is given to AI-based machine learning python code then images are grabbed and features are extracted from it using CNN TensorFlow platform for faster recognition. These features are compared with pre-trained models that use HAR, kinetics-400 action dataset to detect actions with 90% accuracy useful in the detection of normal actions and abnormal actions. As abnormal actions such as cough, head pain, shoulder pain, abnormal heart rate, chest pain, dizziness, fall, etc. are detected soon message with an action image is grabbed and sent to Firebase which is an IOT platform, and raises an alarm to get help to victim. Live feed recognition can be monitored by doctors continuously and can be implemented in public places, hospitals, homes, etc. This can prove a lifesaving gadget in the future. Processing feature extraction of images using Convolutional Neural Networks (CNN) and comparing them with a dataset for activity recognition involves several steps. Here's an overview of the process:

1. Dataset Preparation: Gather a labelled dataset that consists of images or video frames representing different activities or actions. Each image/frame should be labelled with the corresponding activity or action it represents. The dataset should cover a diverse range of activities to train the CNN effectively

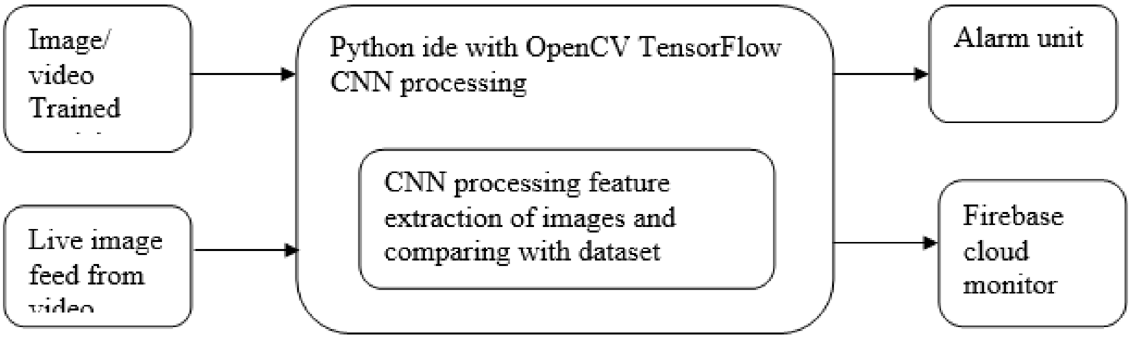

Fig 1. Proposed Elderly People Action Recognition

2. Convolutional Neural Network Architecture: Design a CNN architecture suitable for image processing tasks. CNNs are particularly effective in extracting meaningful features from images due to their ability to capture spatial relationships. Convolutional layers, pooling layers, and

fully connected layers are common components of CNN designs.

3. Training Phase: a. Data Pre-processing: Pre-process the dataset by resizing the images to a consistent size, normalizing pixel values, and augmenting the data if

necessary (e.g., applying random rotations, flips, or crops) to increase the diversity of training samples. b. Dataset Division: Separate the dataset into training and validation sets. The validation set aids in performance monitoring and overfitting prevention while the training set is utilized to train the CNN model. c. Model Training: Train the CNN using the labeled training data. The CNN learns to extract relevant features from the input images through the convolutional layers, applying filters and feature maps to capture patterns and edges at different scales. d. Feature Extraction: Extract features from the CNN by feeding the labeled training data through the trained network up to a specific layer before the fully connected layers. This layer acts as a feature extractor, transforming the input images into a vector representation of learned features.

4. Feature Comparison: a. Feature Vector Representation: Convert the extracted features of each image into a vector representation. b. Similarity Metrics: Utilize similarity metrics, such as Euclidean distance or cosine similarity, to compare the feature vectors between the query image (or video frame) and the feature vectors from the dataset. Activity Recognition: Identify the closest matches based on the similarity scores. The labeled activities corresponding to the closest matches are considered potential activity recognition results.

5. Evaluation and Testing: Evaluate the performance of the activity recognition system using a separate testing dataset. Measure measures like accuracy, precision, recall, or F1 score to evaluate the model's efficacy in correctly identifying activities.

6. Optimization and fine-tuning: If the initial performance is unsatisfactory, optimize the CNN model by using transfer learning techniques to take advantage of previously trained CNN models, changing the architecture, expanding the training data, or fine-tuning the hyperparameters.

By following these steps, the CNN can learn to extract relevant features from images, and the feature vectors can be compared with a dataset to recognize activities accurately. This approach is particularly effective in applications such as action recognition in videos, surveillance systems, human-computer interaction, and activity monitoring in various domains.

Firebase Cloud Monitoring is a service provided by Google Firebase that allows you to monitor and gain insights into the performance and availability of your Firebase applications. It provides real-time monitoring and alerting capabilities to help you proactively identify and resolve issues.

Here are some key features and components of Firebase Cloud Monitoring:

1. Metrics: Firebase Cloud Monitoring collects and analyses various metrics related to your Firebase applications, such as response latency, error rates, throughput, and resource utilization. These metrics give you a detailed view of how your app is performing and help you identify areas that require attention.

2. Dashboards: Firebase Cloud Monitoring provides customizable dashboards where you can visualize and track your application's key metrics.

Features Selection

Some common approaches for feature selection in CNNs are:

1. Pre-trained models: Using pre-trained CNN models like VGG16, ResNet, or Inception, which have already learned discriminative features from large datasets like ImageNet.
2. Fine-tuning: Fine-tuning a pre-trained model on a smaller dataset specific to elderly people's health monitoring can help adapt the learned features to the target domain.
3. Transfer learning: Extracting features from one or more layers of a pre-trained CNN and using those features as input to a classifier to detect abnormal behaviors.

In CNNs, the feature selection process is typically integrated with the model architecture, where convolutional layers learn to extract hierarchical features automatically. The selected features are then passed through fully connected layers for classification.

In both HAR and CNN algorithms, feature selection is important for reducing the dimensionality of the input data, eliminating irrelevant or redundant features, and improving the models' accuracy and efficiency. The specific feature selection techniques and algorithms used may vary depending on the dataset, target behaviors, and the desired performance of the models.

4. Experimentation

The system is implemented with two algorithm techniques first one is HAR human action recognition uses kinetics-400 pre-trained dataset to detect different 400 actions in real time. The second one is with tensor flow CNN to detect elder people's health or show different body pains through actions recognized via Python and sent to the Firebase cloud platform. Accuracy for the first technique is shown in Table 1. The accuracy for the second technique is shown in Table 2.

The accuracy score for different actions refers to the performance of a model in correctly classifying each action or activity in a classification task. It represents the percentage of correctly classified instance

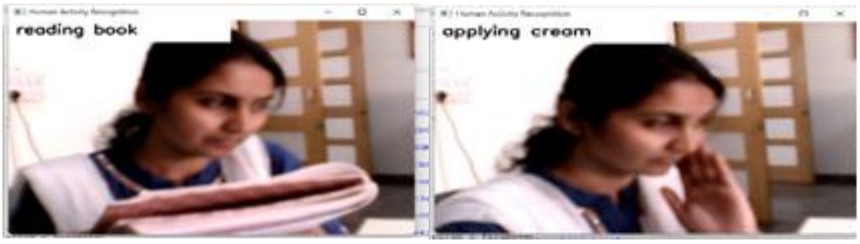

Figure 2. Reading Book, Applying Cream Actions Detection.

for each action out of the total instances of that action in the dataset. Here's an explanation of how to interpret accuracy scores for different actions:

1. Action Labels: Each action or activity in the dataset is assigned a specific label or class. For example, in a human action recognition task, action labels can include activities like walking, running, jumping, or sitting.

2. Accuracy Score Calculation: The accuracy score is calculated separately for each action by comparing the predicted labels assigned by the model to the ground truth labels in the dataset.

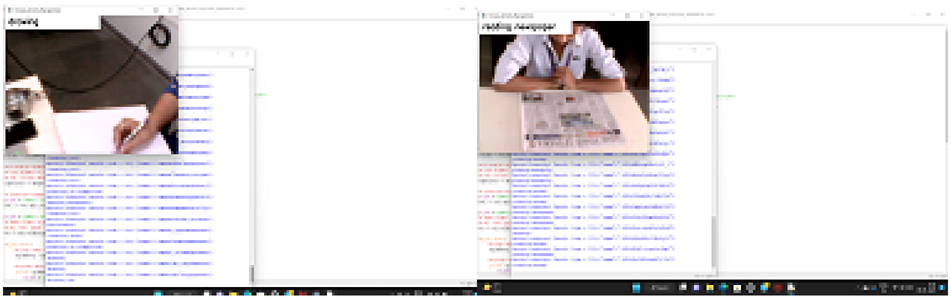

Fig 3. Drawing and Reading Newspapers sent actions to Firebase.

Monitoring heart pain, cough, and head pain for elderly people can help in assessing their health condition and detecting any potential issues or symptoms that require

attention. Here's an overview of how these symptoms can be monitored:

1. Heart Pain (Chest Pain): ECG Monitoring: Electrocardiogram (ECG) monitoring can be used to monitor the electrical activity of the heart. It can help detect any abnormal rhythms or signs of heart-related issues that may cause chest pain. Heart Rate Monitoring: Continuous heart rate monitoring can provide insights

into the heart's activity and detect any irregularities or fluctuations that could be associated with heart pain. Wearable Devices: Wearable devices such as heart rate monitors or smartwatches can track heart rate, heart rate variability, and other relevant metrics. They can provide real-time alerts or notifications if any abnormalities are detected.

2. Cough Monitoring: Audio Monitoring: Monitoring audio recordings or using microphones can help detect and analyze cough sounds. Cough detection algorithms can be applied to identify the frequency and intensity of coughing episodes. Wearable Sensors: Wearable sensors placed on the chest or throat can detect coughing vibrations or movements associated with coughing. These sensors can transmit data wirelessly to a monitoring system for analysis. Environmental Sensors: Air quality sensors or particulate matter sensors can help monitor the air quality in the surroundings. Poor air quality or high levels of pollutants can trigger or worsen coughing episodes.

3. Head Pain Monitoring:

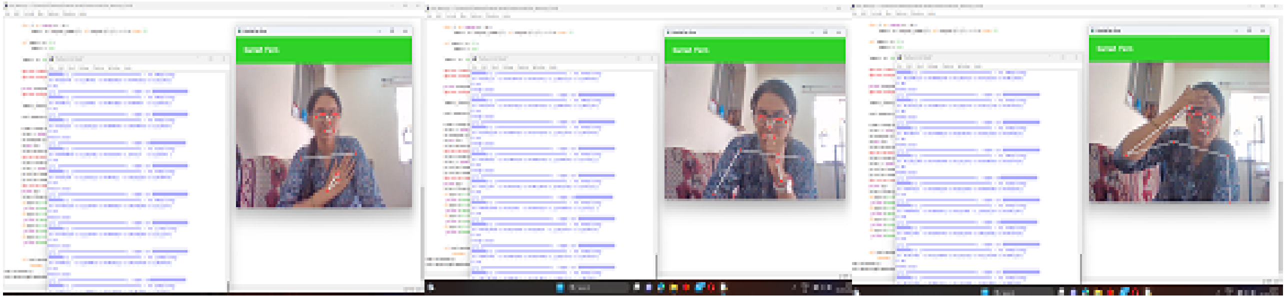

Fig 4. Showing Heart Pain, Cough, and Head Pain for Elderly People Monitors.

Self-Reporting: Elderly individuals can report their head pain levels using a pain scale or rating system. They can note the intensity, frequency, and duration of their head pain episodes. Mobile Applications: Mobile apps can provide a platform for individuals to record and track their head pain episodes, allowing them to input details such as pain level, triggers, associated symptoms, and medications taken. Wearable Devices: Some wearable devices, such as smart headbands or EEG sensors, can monitor brain activity and detect abnormal patterns or signals that may be related to head pain. These devices can transmit data to a monitoring system for analysis.In all these monitoring approaches, it is crucial to have a centralized system or platform that receives and analyses the collected data. This system can employ data analysis techniques, machine learning algorithms, or rule-based systems to identify patterns, triggers, and potential correlations between symptoms and health conditions. Real-time alerts can be generated for caregivers or healthcare professionals if any concerning patterns or changes are detected, allowing for timely intervention and appropriate medical care.

Accuracy = True Negative +True Positive / (True Positive + False Positive + False Negative + True Negative) *(1)*

Alternatively, can calculate accuracy using the number of correctly classified instances (TP) and the total number of instances (N):

Accuracy = (Number of True Positives) / (Total Number of Instances) *(2)*

5. Result and Discussion

To calculate accuracy with a mathematical expression needs two values: the number of correctly classified instances (TP: true positives) and the total number of instances (TP + TN: true positives + true negatives).

Here is the formula:

To calculate the average accuracy score, we sum up all the individual accuracy scores and divide by the total number of actions:

Average Accuracy = (Sum of all Accuracy scores) / (Number of actions) *(3)*

Table 1: Accuracy Score for different actions given below

Action	Fixing hair	Reading	Drawing
Accuracy (%)	87.48	88.30	85.06
Action	Boxing	Hand waving	Handclapping
Accuracy (%)	83.36	89.12	90.83
Average (%)	87.36		

Table 2: Accuracy Score by different action

Action	Walking	Jogging	Running
Accuracy (%)	92.20	93.53	90.04
Action	Boxing	Hand waving	Hand clapping
Accuracy (%)	89.27	94.41	95.50
Average (%)	92.49		

This average accuracy score represents the overall performance of the model across all actions. It is calculated by averaging the accuracy scores obtained for each action individually. In this case, the model achieves accuracy scores of 87.48% for fixing hair, 88.30% for reading, 85.06% for drawing, 83.36% for boxing, 89.12% for hand waving, and 90.83% for handclapping. The average accuracy score provides an overall measure of how well the model performs across the different actions. The accuracy score, in the context of action recognition using pre-trained CNNs, represents the performance of the model in correctly classifying different actions or activities in a dataset. When a pre-trained CNN is used for action recognition, it typically involves the following steps:

1. Pre-training CNN: A CNN (Convolutional Neural Network) is trained on a large-scale image classification task such as ImageNet. This pre-training phase allows the CNN to learn general features and patterns from a vast number of labeled images.

2. Transfer Learning: After pre-training, CNN's learned weights and features are utilized as a starting point for action recognition. The last few layers of the CNN are often modified or replaced with new layers to adapt the network to the specific action recognition task.

3. Training on Action Recognition Dataset: The modified CNN is then trained on a labeled dataset specifically created for action recognition. This dataset consists of video sequences or frames where each action is associated with a specific label.

4. Evaluation and Accuracy Score: Once the CNN model is trained, it is evaluated on a separate test dataset that contains video sequences or frames with known labels. The accuracy score is computed by comparing the predicted labels by the model with the ground truth labels. I T is the proportion of properly identified actions in the test dataset relative to all activities. How successfully the pre-trained CNN model generalizes to the action recognition test may be inferred from the accuracy score. Higher accuracy scores show that the model can effectively categorize activities more precisely

It is crucial to remember that the accuracy score needs to be evaluated in light of the particular action recognition dataset and its features. Different actions may have varying levels of complexity, intra-class variations, or similarities, which can influence the accuracy score.

The accuracy scores for different actions are in Table 2 and the average accuracy score using mathematical expressions:

Using equation (3), the Average Accuracy will be

Average Accuracy = (92.20 + 93.53 + 90.04 + 89.27 + 94.41 + 95.50) / 6

Average Accuracy = 555.95 / 6

Average Accuracy ≈ 92.49%

Therefore, the average accuracy score is approximately 92.49%.

This average accuracy score represents the overall performance of the model across all actions. It is calculated by averaging the accuracy scores obtained for each action individually. In this case, the model achieves accuracy scores of 92.20% for walking, 93.53% for jogging, 90.04% for running, 89.27% for boxing, 94.41% for hand waving, and 95.50% for hand clapping. The average accuracy score provides an overall measure of how well the model performs across the different actions. Table 2 and Table 3 show the accuracy achieved between the two models HAR without the pre-trained model and tensor flow with CNN trained model. As shown in the results system accuracy results increased by 6% with the trained model as compared with without a trained model detecting reading, drawing, fixing hair, heart pain, cough, and head pain as real-time monitored on an elder person's life

6. Conclusion

OpenCV python platform-based system designed for elderly people, or human health monitoring and abnormal behavior detection and raising alarm for help to give medication in time and save a life will prove as life-saving gadget over some time as after medical emergency suffering person does not get immediate medical help results in the death of a person as time extends. The system will provide immediate notification to doctors via IOT after abnormal actions are detected. Through ML python platform-based system uses a kinetic dataset which detects 400 actions; the tensor flow CNN-based platform detects various health issues such as stress on the brain, stomach pain, heart pain, shoulder pain, and many more with 90% accuracy. As the system detects any of this abnormal activity immediately it sends a notification for help to doctors to save a life without wasting time. The system can monitor human health using live camera feeds to doctors over the internet. A real-time system is implemented for live health detection of patients or elderly people by using computer vision to generate auto calls to nearby persons or a doctor to save their lives.

References

[1] Abdellaoui, M., Douik, A., "Human action recognition in video sequences using deep belief networks", Traitement du Signal, Vol. 37, No. 1, https://doi.org/10.18280/ts.370105, pp. 37-44,2020

[2] Yu-Wei Chao, Sudheendra Vijayanarasimhan, Bryan Seybold, David A. Ross, Jia Deng, Rahul Sukthankar; "Rethinking the Faster R-CNN Architecture for Temporal Action Localization", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 1130-1139.

[3] Feichtenhofer, Christoph & Fan, Haoqi & Malik, Jitendra & He, Kaiming," Slow Fast Networks for Video Recognition" 2018.

[4] Ding, Chunhui & Fan, Shouke & Zhu, Ming & Weiguo, Feng & Jia, Baozhi.," Violence Detection in Video by Using 3D Convolutional Neural Networks", 8888. 10.1007/978-3-319-14364-4_53. Pp. 551-558.2014.

[5] Karpathy, Andrej & Toderici, George & Shetty, Sanketh & Leung, Thomas & Sukthankar, Rahul & Fei-Fei, Li. "Large-Scale Video Classification with Convolutional Neural Networks. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition", pp. 1725-1732. 101109/CVPR.2014.223.

[6] S. Z. Gurbuz and M. G. Amin, ''Radar-based human-motion recognition with deep learning: Promising applications for indoor monitoring,'' IEEE Signal Process. Mag., vol. 36, no. 4, pp. 16–28, Jul. 2019.

[7] Y. Kim and H. Ling, "Human activity classification based on micro doppler signatures using an artificial neural network,'' in Proc. IEEE Antennas Propag. Soc.Int. Symp., Jul. 2008.

[8] Jalal A, Kamal S, Kim D. A Depth "Video Sensor-Based Life-Logging Human Activity Recognition System for Elderly Care in Smart Indoor Environments". *Sensors*. 2014;14(7): 1173511759.https://doi.org/10.3390/s140711735.

[9] J. Li, S. L. Phung, F. H. C. Tivive, and A. Bouzerdoum," Automatic classification of human motions using Doppler radar," in Proc. Int. Joint Conf. Neural Netw. (IJCNN), pp.1– 6, Jun. 2012,

[10] MW. Li, B. Xiong, and G. Kuang, "Target classification and recognition based on micro-Doppler radar signatures," in Proc. Progr. Electromagn. Res. Symp. -FALL (PIERS-FALL), pp. 1679–1684, Nov. 2017,

[11] https://www.deepmind.com/open-source/kinetics

[12] Ullah, J. Ahmad, K. Muhammad, M. Sajjad, and S. W. Baik, "Action Recognition in Video Sequences using Deep Bi-Directional LSTM With CNN Features," in IEEE Access, vol. 6, pp. 1155-1166, 2018, doi: 10.1109/ACCESS.2017.2778011

Authors Profile

Sneha B. Paymal received a BE degree in Electronics and Telecommunication engineering from Shivaji University, Kolhapur, India, in 2013 and, an ME degree in Electronics engineering from Shivaji University, Kolhapur, India, in 2016. she has 6 years of experience and currently working as a Junior Research Fellow at RIT, Islampur. Maharashtra, India. She is pursuing PhD at the research center Rajarambapu Institute of Technology, Rajaramnagar affiliated with Shivaji University Kolhapur. her research interests include Deep learning, machine learning artificial intelligence.

Mahadev S. Patil received a BE degree in Electronics and Telecommunication engineering from Karnataka University Dharwar in 1995, M. Tech degree in Power Electronics from Indian Institute of Technology Bombay in 2002, and a Ph.D. degree in Electronics and Telecommunication engineering from Shivaji University, Kolhapur, India in 2014. He has 23 years of experience and currently working as a Professor and Head

of the Electronics and Telecommunication Engineering Department at Rajarambapu Institute of Technology, Islampur, Maharashtra.