

Human Activity Recognition using Machine Learning

Poornima R, R Jagath, Rishab R, Dhanush N Raju, Niharika P
Kumar

*Department of Information Science and Engineering
RV Institute Of Technology And Management
Bangalore*

Date of Submission: 25-02-2023

Date of Acceptance: 05-03-2023

ABSTRACT— Human Activity Recognition is a dynamic field of research and logical improvement in which different models have been proposed utilizing diverse strategies for recognizable proof and categorization of activities using Machine Learning. The features of picture or video information set are extracted utilizing distinctive kinetic models related with spatial feature leaning. Numerous deep layer trained models have been successfully utilized in this field to reach the basic objective of this model which is recognition and categorization of action taking place. These activities include day to day exercises like running, jogging, eating, sitting, etc. There can be numerous sorts of activities in totally different fields like healthcare, childcare, security or work security. Human Activity Recognition contains an exceptionally noteworthy part in completely different areas like human computer interaction, video surveillance framework, robotics, daily monitoring, wildlife observation, etc. With the use of distinctive datasets like UCF-101, HMDB-51, Hollywood2, Sports-1M and training them this task of recognition of activity can be proficiently done. The execution of Convolutional Neural Network (CNN) model for image recognition with the help of OpenCV helps in effective working of this model. Such application of distinctive datasets on activity recognition model has made a difference in simple categorization of activity based on its nature whether normal or anomalous and suspicious. According to the identified nature a caution is sent through server to the authority concerning the happening of odd movement taking place at real time. Due to such application of this model numerous harmful activities can be dodged or at slightest negative results of such activities can be minimized.

Keywords— Human Activity Recognition;

Machine Learning; Convolutional Neural Network; OpenCV; video surveillance;

I. INTRODUCTION

With fast improvements within the field of activity recognition and proposition of numerous new models based on scientific and innovative improvements monstrous progress in this field can be seen and observed. The improvement in deep learning and OpenCV with highly trained datasets have opened a new entryway of opportunities for upcoming research in this field. Such progress can lead to authentic and valuable application of such models in this digitally prepared world for the well-being of all living creatures. The use of modern and advanced technology in this field by different researchers and engineers have resulted in various applications of these models.

Due to such highly trained models the activities taking place at real time can be observed in exceptionally viable and ideal way. Anomalous or suspicious activities can be treated with convenient strategies guaranteeing peace and concordance within the society of living beings. This can be also very useful in making a smart home environment as well as smart healthcare service with the assistance of regular monitoring. Numerous security issues can be handled carefully and the harm to be caused can be minimized. Such successful application of these models in day-to-day life can also guarantee the psychological well being of individuals without concerns of the harm due to such activities.

The human activity recognition model can be actualized with the use of camera module which captures the crude information that serves as an input to the recognition system. By making distinctive outlines of such input information categorization of activity is done after feature extraction. Such activity is then identified as

normal or suspicious and quick alert is sent to the authority.

II. PROBLEM DEFINITION

After looking over the activities taking place in the town, it was found that numerous sorts of activities are of deceptive nature and are not good for well-being of individuals. Numerous other activities were also observed which led an impact with its consequences due to lack of monitoring and legitimate care. These activities not only cause harm to materialistic things but also take off a mental affect on human beings. Due to such activities much financial misfortune is also done. Due to lack of legitimate monitoring and care people also need to suffer consequences that can be as deplorable as death of their cherished ones particularly children and senior people. Due to numerous such activities the trust on security infrastructure is additionally compromised. To overcome all such problems, this project is going to use accelerometer data collected from multiple users to train the model so that it can predict the human activity. 2D Convolutional Neural Networks is used to build the model.

III. LITERATURE SURVEY

In [1], with the development of the Internet of things (IoT) and wearable devices, the sensor-based human activity recognition (HAR) has attracted more and more attentions from researchers due to its outstanding characteristics of convenience and privacy. Meanwhile, deep learning algorithms can extract high-dimensional features automatically, which makes it possible to achieve the end-to-end learning. Especially the convolutional neural network (CNN) has been widely used in the field of computer vision, while the influence of environmental background, camera shielding, and other factors are the biggest challenges to it. However, the sensor-based

HAR can circumvent these problems well. Two improved HAR methods based on deep CNN are proposed in this paper. Firstly, through the multi-dilated kernel residual (Mdk-Res) module, a new and improved deep CNN network Mdk-ResNet is proposed, which extracts the features among sampling points with different intervals. Furthermore, the Fusion-Mdk-ResNet is adopted to process and fuse data collected by different sensors automatically. The comparative experiments are conducted on three public activity datasets, which are WISDM,

UCI HAR and OPPORTUNITY. The optimal results are obtained by using the indexes such as accuracy, precision, recall and F-measure,

which verifies the effectiveness of the proposed methods.

In [2], the affect of number, situation and sort of the sensors on the exactness of activity/posture recognition are explored. It has been found that combination of diverse sensors accomplished the required precision in complex scenarios indeed when employing a little number of sensors. The objective was to recognize seven activities/postures. With one sensor put on the chest, the classification show was able to recognize between two bunches of stances ({lying and on all fours} and {standing, sitting and sitting on the ground}). This is often since of the contrast within the sensor point introduction. The issues risen whereas recognizing stances inside each gather. That was the case indeed when the gyro and magnetometer information were analyzed.

In [3], human action recognition picked up critical significance in investigate community because it may be a challenging time arrangement classification errand. In human movement expectation, firstly the sensor information is recorded for exercises of particular subjects, at that point a machine learning demonstrate is prepared to generalize the demonstrate for concealed information. There are parts of applications of movement expectations like behavior investigation, well-being and workout checking, stride examination, intuitively gaming, signal acknowledgment, video reconnaissance etc. HAR moreover has numerous applications to progress elderly people's living. The Convolutional Neural Network (CNN) is broadly utilized in picture investigation but it illustrated promising comes about for other sort of signals like discourse acknowledgment, content examination, and human movement forecast. This creates a crossover demonstration with one profound 1D CNN and one SVM to foresee the human movement as appeared. This crossover demonstration could be a great choice since the CNN captures the spatial connection between signals and the SVM captures the spatial-temporal relationship. Together it upgrades the capacity to recognize distinctive exercises that have changed flag dissemination. This demonstration first identifies the inactive and moving movement employing a Arbitrary Timberland (RF) twofold classifier. The RF combines numerous choice trees into a single demonstrate and uses the normal of multiple trees or compute the larger part votes to form a expectation within the terminal leaf. In this inquire about, the UCI-HAR dataset is utilized; the recording of 30 subjects having an age extend from 19 to 48. The dataset comprises of the six action

signals of every day living gotten by a midriff mounted smartphone taking after the movement convention. The information securing employments the smartphone's accelerometer and spinner. The tri-axial (x, y, z) information of exercises are: strolling, walking-upstairs, walking-downstairs, sitting, standing, and laying. In [3], a crossover strategy to viably perform human action recognition is displayed. The strategy to begin with recognizes the unique action by employing a Arbitrary Woodland classifier to recognize the exercises sort as inactive and moving. For inactive activity's particular acknowledgment we utilized bolster vector machine and for moving exercises we outlined a profound 1D CNN. Exactness of 97.71% which is comparable to state-of-the-art execution is accomplished. The future arrange is to convey usage into low-power coordinates circuits to create it well-suited for wearable.

In [4], Human Action Recognition (HAR) points to recognize exercises based on information collected from smartphone sensors. It is one of the developing inquire about zones that meet application requests such as portable computing, surveillance-based security, context-aware computing, and helped living. The proposed CNN show on a Sussex-Huawei Motion- Transportation (SHL) open HAR dataset is evaluated. The SHL dataset was recorded over a period of seven months in 2017 by three individuals (User1, User2 and User3). The three members locked in in eight diverse modes of transportation and movement in real-life settings within the Joined together Kingdom. The objective is to recognize eight modes of movement and transportation (exercises) from the inertial sensor data of a smartphone. This can be done by recognizing the client movement of information coming from the phone of three users. In [4], outfit learning for human action recognition is performed. It presented a CNN demonstrate for classifying the eight strategies of motion and transportation utilizing smartphone IMU sensor information within the SHL Challenge 2020. The information of three sensors, an accelerometer, spinner, and magnetometer are utilized. Exercises which have slower development are way better recognized by the model.

In [5], Human Detection and Activity Recognition (HDAR) in recordings plays an imperative part in different real-life applications. As of late, question discovery strategies have been utilized to identify people in recordings for ensuing decision-making applications. This paper points to address the issue of human location in airborne captured video groupings employing a moving camera connected to an airborne stage with

dynamical occasions such as shifted elevations, light changes, camera jitter, and varieties in perspectives, question sizes and colors. Not at all like conventional datasets that have outlines captured by a inactive ground camera with medium or expansive districts of people in these outlines, the UCF-ARG airborne dataset is more challenging since it contains recordings with huge distances between the people within the outlines and the camera. The execution of human location strategies that have been depicted within the writing are regularly corrupted when input video outlines are misshaped by commotion, obscure, light changes, and the like. To address these restrictions, the question location strategies utilized in this consider were prepared on the COCO dataset and assessed on the freely accessible UCF-ARG dataset. The comparison between these locators was tired terms of discovery exactness. The execution assessment considers five human activities like burrowing, waving, tossing, strolling, and running. Exploratory comes about illustrated that EfficientDetD7 was able to beat other locators with 92.9% normal precision in recognizing all exercises and different conditions counting obscuring, expansion of Gaussian commotion, helping, and obscuring. Moreover, trained Convolutional Neural Networks (CNNs) such as ResNet and EfficientNet were utilized to extricate exceedingly instructive highlights from the identified and edited human patches. The extricated spatial highlights were utilized by Long Short-Term Memory (LSTM) to consider worldly relations between highlights for human movement acknowledgment (HAR). Test results found that the EfficientNetB7-LSTM was able to outflank existing HAR strategies in terms of normal precision (80%), and normal F1 score (80%). The result may be a strong HAR framework which combines EfficientDetD7, EfficientNetB7, and LSTM for human location and action classification.

In [6], movement recognition is an exceptionally capable inquire about point in PC vision with different essential applications, counting human PC interfacing, content-based video asking, video discernment, furthermore, apply self sufficiency, among others. Really, visual activity affirmation has been allocated into sub-subjects, for occasion, flag seen for human PC interfacing, outward appearance affirmation, and improvement conduct affirmation for video recognition. In the proposed strategy, protest discovery is done from the video by applying the channels to upgrade the quality of outlines extricated from the video, at that point extricating the boundary of the outline to distinguish the

question. Form choice is done for partitioning the picture into distinctive segments like upper half and lower half. Presently the concept is how upper and lower half discovery of the protest demonstrates itself good. In this work, protest location system utilizing boundary extraction and form choice is shown. It is not possible to recognize the various exercises to be performed by a human body but we are able too consider the diverse stances to be made by them. With the assistance of discovery highlights of upper body(head) and lower body(legs). Able to become able to discover the specific heading of introductions of those highlights. Location of different postures demonstrates itself much simpler in the event that it can be done as it were on the premise of this division of upper and lower parcels of the total. This work does its part clearly and able recognize more than 90% precision which combines the head location and lower/upper body division and acknowledgment as a entirety.

In [7], the papers that were analyzed in profundity had confinements that were exceptionally particular to the one of a kind engineering of the models. Common impediments with human movement acknowledgment with profound learning, whereas being to some degree insignificant, is that frequently datasets must be exceptionally huge some time recently a demonstrate can gotten to be proficient with recognizing complex exercises. Moreover, models that bargain with complex exercises with numerous focuses have numerous complex steps of information pre-processing to adjust out the information or move weights to permit for generally exactness of acknowledgment for distinctive exercises. Most model's impediments don't come from the profound learning classifiers but utilizing highlight extraction and information preprocessing. Most state-of-the art models indeed actualize whole profound learning based engineering layers include extraction and preprocessing handle. In [7], Human Activity Recognition the de facto strategy for persistently checking not as it were what human creatures are up to but moreover in checking the exercises of gadgets, machine parts, pets, and others. This has made HAR based on IMU sensors a hot region for inquire about. Not to specify that these keep up tall levels of protection and consolation for the client. Much as a few approaches to accurately classifying the exercises of a client with IMU sensor information have been proposed, numerous of them make it out to be an awfully troublesome errand. In this paper, we show a CNN-LSTM classifier for human action acknowledgment. While

both CNN and LSTM systems have been broadly investigated within the past, they've been considered in segregation. Our paper looks for to use the qualities of combining the two systems particularly in as distant as human action acknowledgment is concerned. CNN-LSTM approach to human activity recognition that looks for to progress the exactness of action acknowledgment by leveraging the vigor in highlight extraction of a CNN organize whereas taking advantage of the work an LSTM demonstrate does for time arrangement estimating and classification is proposed. This CNN-LSTM demonstrate is both spatially and transiently profound and accomplished way better execution when it was compared with other profound learning approaches that utilize crude flag information as input. This model with respect to prescient precision and Softmax misfortune on both an inside (iSPL) and a freely accessible (UCI HAR) dataset is assessed. In both cases, it beats the other models particularly on the iSPL dataset with over 1% more exactness that its closest equal and near to 2% less Softmax misfortune.

In [8], the reliance on smartphones for a normal human is expanding day by day. A social test conducted by Sumathi Etal., with scholarly understudies as subjects, found that 34% of the respondents went through 5 to 7 hours of their day on smartphones on scholastic and social communication itself. Human Activity Recognition (HAR) is by and large assessed based on precision and computational fetched. In arrange to successfully perceive diverse exercises, past works have attempted to extricate hand-crafted highlights from the signals of the accelerometer and whirlygig, and connected diverse classical machine learning approaches like Bolster Vector Machine, Irregular Woodland etc. for classification. The dataset has 561 highlights and 7352 perceptions, the aggregate of which is utilized for our introductory investigation for producing conclusions utilizing K-Nearest Neighbor (KNN) and Irregular Timberland classifier. KNN runs through the whole dataset finding the separate "d" between the imperceptibly point and each point of the preparing set. The techniques and test setup to gather information are expounded in critical detail in . Anguita et al. chosen 30 volunteers aged from 19 to 48 a long time and inquired them to perform six essential exercises: strolling, strolling upstairs, strolling ground floor, sitting, standing and laying while wearing a "waist-mounted Samsung Universe S II smartphone". In this think about, information extricated from the accelerometer and spinners of smartphones is utilized to infer designs and define

comes about relating to the fundamental physical action of an person. The proposed learning technique gives superior outcomes compared to the other strategies, when computed with the same information, and beneath the same environment of computation.

(HAR) based on inertial measurement unit (IMU) has ended up

In [9], HAR is the capacity of a framework to identify and recognize particular human exercises with the information collected through a sensor or camera. Deep Learning methods have been goal examined and connected to achieve sensor- based HAR. CNNs have demonstrated to work for HAR due to their capability of capturing nearby conditions on flag information, as well as their conservation of include scale in variance when completing include extraction. CNNs are profound learning models that utilize convolution, pooling, completely associated layers, and covered up layers to achieve classification and recognition. Input information is bolstered into convolutional layers, which apply channels to the input information and recognize nearby correlations. The dataset we utilized to test our models is the open space UCI dataset for HAR. It could be a dataset containing inertial information from the inserted accelerometer and spinner in a Samsung System S II smartphone. In this paper, we show a multi-layer, crossover and lightweight CNN-LSTM design that beat already displayed conventional machine learning and profound learning strategies. Our 2-layer CNN and 1-layer LSTM half breed model. outperforms all other multi-layer cross breed models. In addition, the lightweight half breed demonstrate not as it were incorporates a tall execution precision, but too incorporates a quicker demonstrate preparing time. Future work will utilize bigger, more complicated datasets to confirm the execution of lightweight cross breed demonstrate.

IV. PROPOSED SYSTEM

With the development of the Internet of things (IoT) and wearable devices, the sensor-based human activity recognition (HAR) has attracted more and more attentions from researchers due to its outstanding characteristics of convenience and privacy. Meanwhile, deep learning algorithms can extract high- dimensional features automatically, which makes it possible to achieve the end-to-end learning. Especially the convolutional neural network (CNN) has been widely used in the field of computer vision, while the influence of environmental background, camera shielding, and

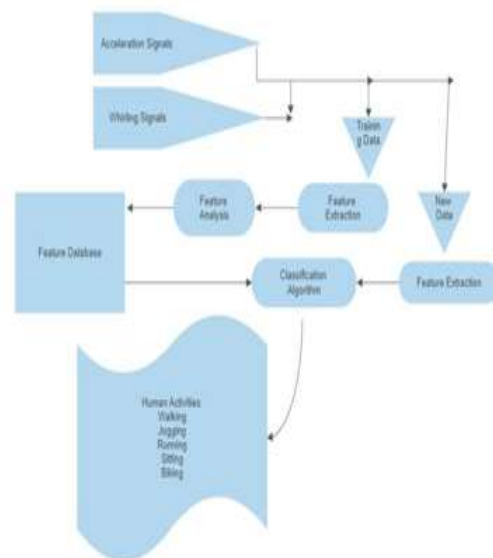
other factors are the biggest challenges to it. However, the sensor-based HAR can circumvent these problems well. Two improved HAR methods based on deep CNN are proposed in this paper. Firstly, through the multi-dilated kernel residual (Mdk-Res) module, a new improved deep CNN network Mdk-ResNet is proposed, which extracts the features among sampling points with different intervals. Furthermore, the Fusion-Mdk-ResNet is adopted to process and fuse data collected by different sensors automatically. The comparative experiments are conducted on three public activity datasets, which are WISDM, UCI HAR and OPPORTUNITY.

The objectives of the proposed system are as follows :

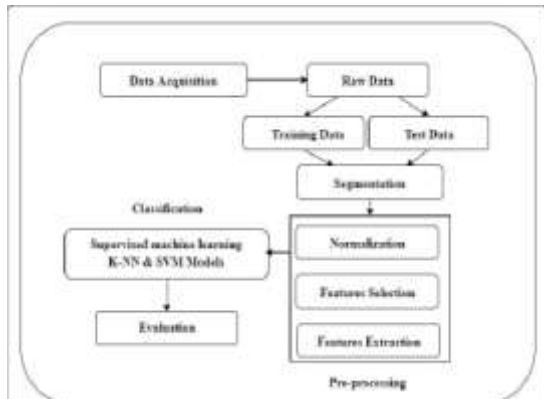
- Load and prepare human activity recognition time series classification data.
- Explore and visualize time series classification data in order to generate ideas for modeling.
- Predict the correct human activity among the 7 activities (Downstairs, Jogging, Sitting, Standing, Upstairs, Walking, Biking) and displaying its probability on an Android App.

V. DIAGRAMS

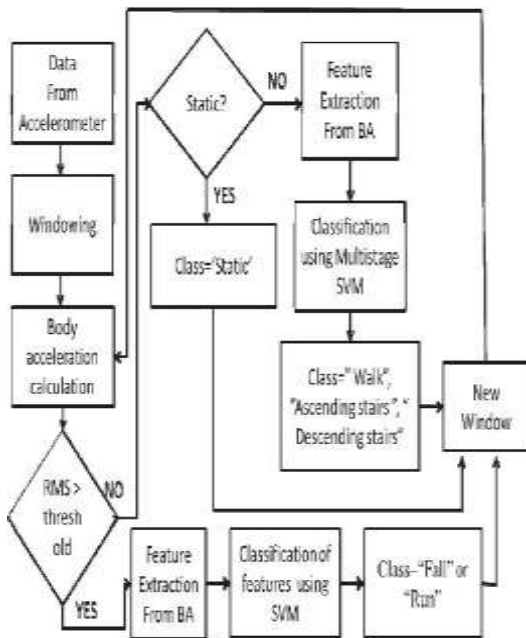
A System Architecture



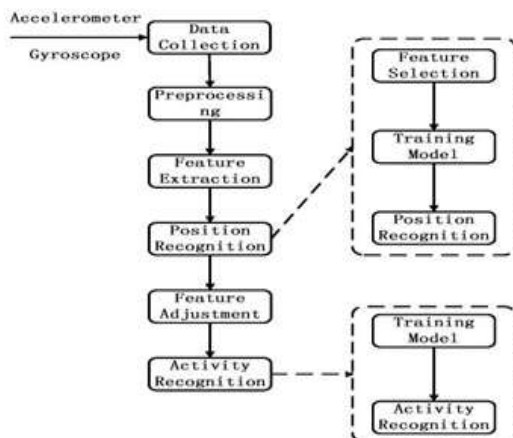
B DFD Level-0



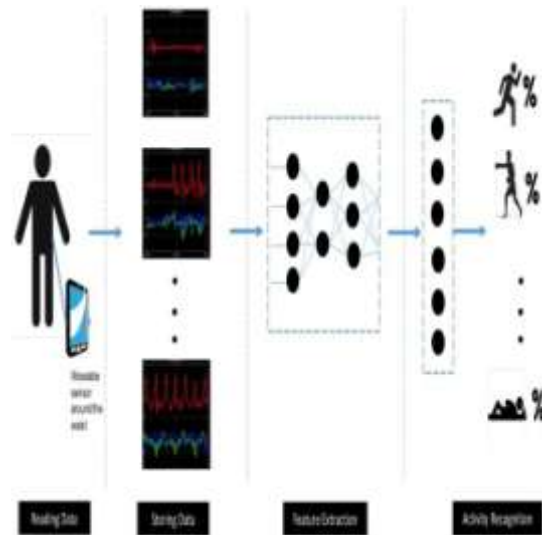
C DFD Level-1



D DFD Level-2



E Sample Class Diagram



VI. FUTURE SCOPE

Human Activity Recognition can benefit various applications in fields like smart home monitoring, healthcare services, security surveillance, childcare etc. In future this application can be updated by using object activity recognition in which activities performed by objects can also be tracked and analyzed. Application of integrated large datasets can be done to identify the activity taking place as slower rate of time. Even very subtle or minute variations should be recognized by the system. The data of actor performing the anomalous activity can be stored and identification of actor can be done if not caught in the first place. Activities that are of reoccurring manner should be stored to save time and space during recognition process. Implementation of such model can also be done in Government authority section. Much more developments for improvisation in accuracy and dealing with issues related to optical identity and background clutter of image can be done.

VII. CONCLUSION

The accelerometer data collected from multiple users is used to train the model so that it can predict the human activity. 2D Convolutional Neural Networks is used to build the model. Load and prepare human activity recognition time series classification data. Explore and visualize time series classification data in order to generate ideas for modeling. Predict the correct human activity among the 7 activities (Downstairs, Jogging, Sitting, Standing, Upstairs, Walking, Biking) and displaying its probability on an Android App.

REFERENCES

- [1] Human Activity Recognition Based on Graman Angular Field and Deep Convolutional Neural Network (Hongji Xu, Juan Li, Hui Yuan, Qiang Liu) - 2021
- [2] Activity/Posture Recognition using Wearable Sensors Placed on Different Body Locations (Hristijan Gjoreski, Matjaz Gams) - 2020
- [3] A Hybrid Approach for Human Activity Recognition with Support Vector Machine and 1D Convolutional Neural Network (Md Maruf Hossain Shuvo, Nafis Ahmed, Koundinya Nouduri, Kannappan Palaniappan) - 2022
- [4] Human Activity Recognition Using Convolutional Neural Networks (Gulustan Dogan, Sinem Sena Ertas, İremnaz Cay)
- [5] Human Body Posture Recognition Using Artificial Neural Networks . (Manu Bali, Devendran V) - 2020
- [6] A Close Look into Human Activity Recognition Models using Deep Learning . (Wei Zhong Tee, Rushit Dave, Jim Seliya, Mounika Vanamala) - 2021
- [7] A CNN-LSTM Approach to Human Activity Recognition . (Ronald Mutegeki, Dong Seog Han) - 2022
- [8] Human Activity Recognition using Deep Neural Network . (Piyush Mishra, Sourankana Dey, Suvro Shankar Ghosh, Dibyendu Bikash Seal, Saptarsi Goswami) - 2020
- [9] Improved Sensor Based Human Activity Recognition via Hybrid Convolutional and Recurrent Neural Networks . (Sonia Perez-Gamboa, Qingquan Sun, Yan Zhang) - 2021