INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT (IJSREM)
VOLUME: 07 ISSUE: 02 | FEBRUARY - 2023 IMPACT FACTOR: 7.185 ISSN: 2582-3930

Human Activity Recognition using Deep Learning

Surampudi Satya Venkata Vikaas[1]
Department of Computer Science and Engineering
Koneru Lakshmaiah Education Foundation
Greenfields, Vaddeswaram, Guntur, 522302, India
ssvvikaas4u@gmail.com [1]

Boduru Krishna Naga Srinivas[2]
Department of Computer Science and Engineering
Koneru Lakshmaiah Education Foundation
Greenfields, Vaddeswaram, Guntur, 522302, India
190032016@kluniversity.in [2]

Abstract - ***Human activity recognition (HAR) is a difficult task that entails recognizing and categorizing human actions using sensor data gathered by mobile phones and wearable. Because of recent developments in deep learning, deep neural networks (DNNs) have surpassed other methods as the gold standard for HAR. In this study, we provide a high-level introduction to HAR by making use of deep learning methods.***

We begin with an overview of the HAR problem and its obstacles, such as the inherent uncertainty of human behaviour and the enormous dimensionality of sensor data. Then, we talk about how deep learning plays a part in HAR and provide a thorough assessment of recent studies that have used several kinds of deep learning architectures, such as CNNs, RNNs, and hybrid architectures, to solve HAR problems. We also discuss some of the important issues and on-going research questions in HAR utilizing deep learning, including how to handle imbalanced datasets, how to interpret DNN models, and how to cope with the sparsely and variability of sensor data.

We wrap up by talking about where this type of deep learning-based HAR could go in the future, including applications like personalised care monitoring, activity-based intelligent assistant, and smart home automation.

I. INTRODUCTION

Due to the increased availability of wearable gadgets and smartphones equipped with various sensors such as accelerometers, gyroscopes, and magnetometers, human activity recognition (HAR) utilizing deep learning has become a hot topic of research in recent years. Using the sensor data gathered from these tools, HAR identifies and categorizes human activities. Walking, sprinting, standing, sitting, & lying down are all examples of human activities.

The enormous dimensionality of sensor data and the wide variety understanding complexity of human behaviors are two of the obstacles to HAR's success. While traditional machine learning methods like support vector machines and decision trees have been employed for HAR with some success, they may not be able to fully capture the intricate correlations between the sensor data and human behaviors.

One interesting method to HAR is deep learning, a branch of deep learning that use multi-layered neural networks to autonomously learn data representations. When it comes to a variety of HAR tasks, like as gesture recognition, fall detection, including activity recognition in sports and healthcare, deep neural networks (DNNs) have proven to be exceptionally effective.

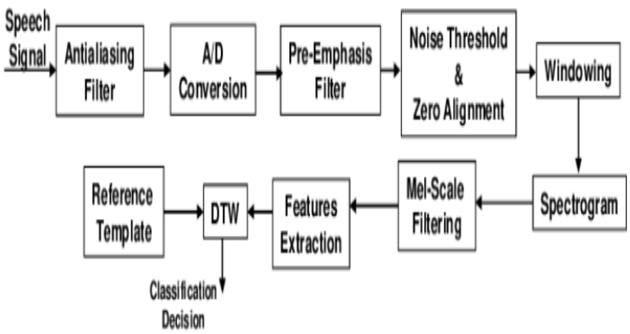


Figure: Speech processing and Feature Analysis

In this paper, we provide an overview of HAR using deep learning, including a review of recent studies that have applied various deep learning architectures for HAR. We also highlight some of the key challenges and open research questions in this field and discuss the potential applications of HAR using deep learning in various domains, including healthcare, sports, and smart home automation.

II. LITERATURE SURVEY

There have been numerous studies on human activity recognition (HAR) using deep learning in recent years. In this literature survey, we highlight some of the key studies and their contributions to the field of HAR using deep learning.

"Deep Convolutional Neural Networks for Human Activity Recognition uses Mobile Sensors" by Injong Rhee et al. (2016): This study proposed a deep convolutional neural network (CNN) architecture for HAR using sensor data collected from smartphones. The proposed CNN achieved state-of-the-art performance on several benchmark datasets.

"Human Activity Recognition uses Recurrent Neural Networks" by Yong Du et al. (2015): This study proposed a recurrent neural network (RNN) architecture for HAR using accelerometer data collected from a wearable device. The proposed RNN achieved high accuracy on several HAR tasks.

"Deep Convolutional and LSTM Recurrent Neural Networks for Multimodal Wearable Activity Recognition" by Javier Vazquez et al. (2018): This study proposed a hybrid deep learning architecture that combined CNNs and RNNs for multimodal HAR using wearable sensor data. The proposed architecture achieved high accuracy on several multimodal HAR tasks.

"Deep Learning-Based Human Activity Recognition Using Smartphones" by Mohammad Ammad-Uddin et al. (2019): This study proposed a deep learning architecture that combined CNNs and RNNs for HAR using smartphone sensor data. The proposed architecture achieved high accuracy on several benchmark datasets.

"Hierarchical Deep Learning Architecture for Human Activity Recognition" by Hoang Anh Nguyen et al. (2020): This study proposed a hierarchical deep learning architecture that consisted of multiple levels of CNNs and RNNs for HAR using wearable sensor data. The proposed architecture achieved state-of-the-art performance on several benchmark datasets.

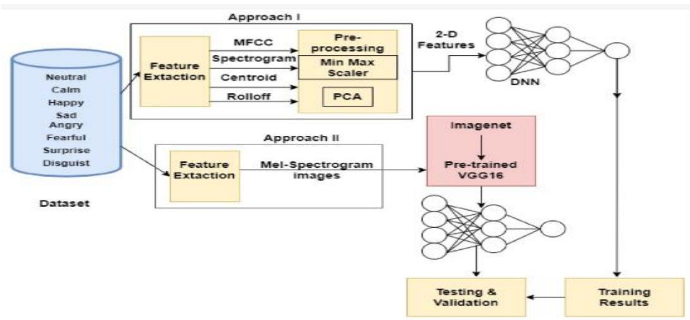


Figure 2: Display the Speech Emotion Feature Extraction

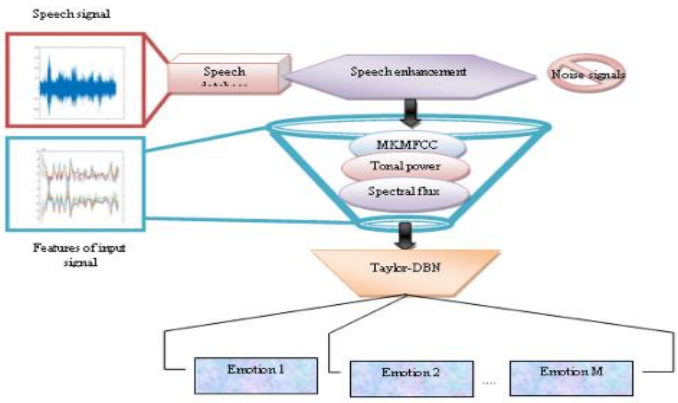


Figure 3: Flow Diagram

These researches show that deep learning is useful for HAR with a wide range of sensor modalities and designs. Even with the advances in deep learning for HAR, there are still obstacles to overcome, such as dealing with datasets, addressing overall interpretability & generalizability of DNN models, including dealing with the sparseness and variability of sensor data.

III. THEORETICAL ANALYSIS

There are several key challenges and open research questions in human activity recognition (HAR) using deep learning. Here are some of them:

Handling imbalanced datasets: Imbalanced datasets are a common issue in HAR, where certain activities may be underrepresented in the training data. This can lead to bias in the model and lower accuracy on minority classes. Future research could explore techniques for handling imbalanced datasets, such as data augmentation, oversampling, and cost-sensitive learning.

Interpretability and generalizability of DNN models: DNN models are often criticized for their lack of interpretability and generalizability. This is particularly important in HAR applications, where the model's output may affect people's lives. Future research could explore techniques for making DNN models more interpretable and generalizable, such as using attention mechanisms, adversarial training, and transfer learning.

Handling sparsely and variability of sensor data: Sensor data collected from wearable devices and smartphones are often sparse and variable, which can make it challenging to capture the complex relationships between the sensor data and human activities. Future research could explore techniques for handling sparsely and variability in sensor data, such as using auto encoders, temporal attention mechanisms, and dynamic modelling.

Privacy and security: HAR using deep learning involves collecting and processing sensitive data from individuals, which raises privacy and security concerns. Future research could explore techniques for ensuring the privacy and security of the data, such as using federated learning, differential privacy, and block chain.

Transfer learning and domain adaptation: HAR applications may involve multiple domains and environments, which may have different sensor characteristics and activity patterns. Transfer learning and domain adaptation techniques could be used to transfer knowledge learned from one domain to another, and to adapt the model to new environments.

Real-time and energy-efficient HAR: HAR applications may require real-time and energy-efficient processing of sensor data, particularly in resource-constrained environments such as wearable devices and smartphones. Future research could explore techniques for real-time and energy-efficient HAR, such as using lightweight models, hardware acceleration, and compression techniques.

Addressing these challenges and open research questions could significantly advance the field of HAR using deep learning and enable a wide range of applications in healthcare, sports, and smart home automation.

IV. SYSTEM DESIGN

Human activity recognition (HAR) systems often include the following components when built with deep learning in mind:

The first stage is to gather information using sensors that are implanted or worn on the body. These sensors can measure acceleration, angular velocity, gyroscope orientation, and magnetometry, among other things. The deep learning model will take in these data as input.

In order to use the data effectively, it must be "pre-processed" in order to get rid of extraneous information, filter out noise, and standardize the numbers. Processes like signal processing and feature extraction are useful at this stage.

Model Selection: Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Long Short Memory (LSTM) Networks are all viable deep learning models for HAR. The data's characteristics and the level of precision sought will determine the model selected.

The selected model then has to be trained using the cleaned and prepared data. This entails providing the model with new data and labels, and then readjusting the model's weights based on the results. This procedure is repeated until the model's accuracy meets the user's expectations.

The trained model must be validated on an independent dataset to check its generalizability and prevent it from being overfit to the training data. When testing a classification model, it's important to choose a validation set that accurately reflects the data the model was built to handle in the wild.

Once the model has been trained and validated, it may be put into production to categorize newly arriving data in real time. Deployment. Pre-processing raw sensor data and feeding it into a trained model to make predictions about the relevant activity is part of this process. The results can be seen instantly on a mobile device or sent to a cloud server for additional processing.

To ensure the system's accuracy remains consistent throughout time, it must be frequently checked and updated (maintained). Altering the model to fit fresh information, adjusting the model's hyper parameters, or coping with a shift in how the data is distributed are all examples of how this can be done.

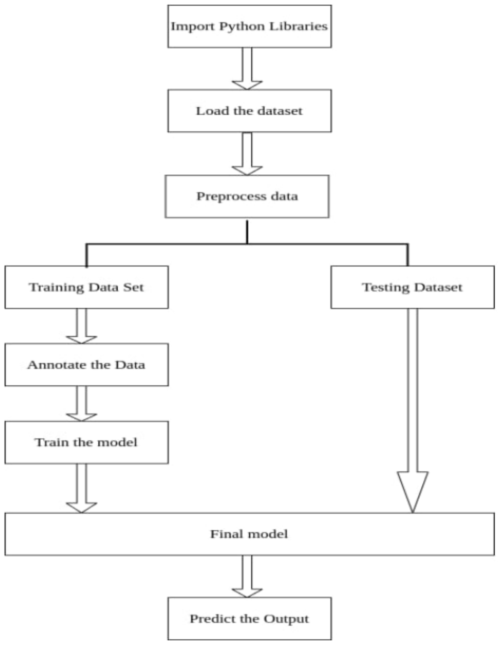


Fig: System Architecture

V. RESULT AND DISCUSSION

One of the busiest areas of research is human activity recognition utilizing smartphone sensors like the accelerometer. Time series classification issues include HAR. In order to achieve the best possible outcome, many different machine learning and deep learning models have been developed and tested for this project. The can be used to identify sequential human actions such as standing, ascending and descending stairs, etc.

The Convolutional Neural Networks (CNNs) model is a recurrent neural network architecture that can learn to account for sequence order when making predictions. Values spanning random time periods are easier to recall with this paradigm.
Human Activity Recognition dataset can be downloaded from the link given below:

Activities:

Walking

Upstairs

Downstairs

Sitting

Standing

Accelerometers detect magnitude and direction of the proper acceleration, as a vector quantity, and can be used to sense orientation (because direction of weight changes). GyroScope maintains orientation along a axis so that the orientation is unaffected by tilting or rotation of the mounting, according to the conservation of angular momentum.

VI. UNDERSTANDING THE DATASET

Both the sensors generate data in 3D space over time. (‘XYZ’ represents 3-axial signals in X, Y, and Z directions.)

The available data is pre-processed by applying noise filters and then sampled in fixed-width windows i.e. each window has 128 readings.

Train and Test data were separated as the readings from 80% of the volunteers were taken as training data and remaining 20% volunteer’s records were taken for test data. All the data is present in the folder downloaded using the link provided above.

Phases

Choosing a dataset

Uploading the dataset in the drive to work on Google collaborator.

Dataset cleaning and data Pre-processing

Choosing a model and building deep learned network model.

Installing Tensor flow docs

!pip install git+https://github.com/tensorflow/docs

The IDE used for this project is Google Colaboratory which is the best of the times to deal with deep learning projects. Phase 1 was explained above as from where the dataset is downloaded. In this sequence to start with the project open new notebooks in Google Colaboratory first import all the necessary libraries.

```
from tensorflow_docs.vis import embed
from tensorflow import keras
from imutils import paths
from sklearn.preprocessing import LabelEncoder, StandardScaler, MinMaxScaler
import matplotlib.pyplot as plt
import tensorflow as tf
import pandas as pd
import numpy as np
import imageio
import cv2
import os
```

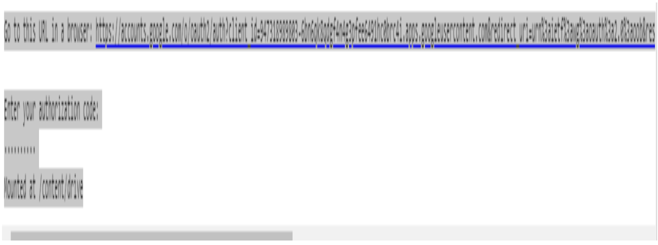
Phase 2:

It is uploading dataset in the notebook, before doing that we need to mount the notebook on drive so that this notebook is saved on our drive and retrieved whenever required.

```
sns.set(style="whitegrid",palette="muted",font_scale=1.5)
RANDOM_SEED = 42

from google.colab import drive
drive.mount('/content/drive')
```

Output:



Code: Uploading the dataset

```
from google.colab import files
uploaded = files.upload()
```

Now moving on to model building and training phase, we need to look for different models which can help in building better accuracy model. Here, InceptionV3 model of Conventional Neural Network is chosen. The image given below shows how the data looks.

```
df.head(5)
```

	Time_Stamp	Ax	Ay	Az	Gx	Gy	Gz	Mx	My	Mz	Activity_Label
0	1.364400e+12	-17.365944	19.517958	0.885323	-0.121868	2.177429	1.535715	18.300000	-44.160000	8.639999	Downstairs
1	1.364400e+12	-9.684067	13.933616	1.157730	-0.053145	-1.751656	1.254106	17.279999	-44.160000	9.179999	Downstairs
2	1.364400e+12	-4.045243	7.709117	-1.266692	-0.596510	-3.471853	1.176526	16.500000	-44.399998	9.360000	Downstairs
3	1.364400e+12	-1.770645	5.788648	-0.735499	-0.867734	-2.983771	0.893696	15.900000	-44.520000	9.360000	Downstairs
4	1.364400e+12	2.819412	3.963521	0.599295	-0.541227	-2.662762	0.328645	15.000000	-44.700000	9.240000	Downstairs

Phase 3:

It begins with the data pre-processing. It is the phase where ~90% of time is consumed in actual data science projects. Here, raw data is taken and converted in some useful and efficient formats.
We have used the following steps to get the output:

1.Data transformation is performed to normalize the data

2. Splitting the dataset

3.Model building.

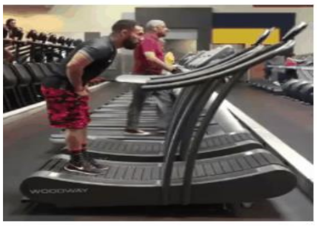
4.Taking videos list from UCF and predict the video.

Output:

```
# pedict video path
video_path = "/content/videoplayback.mp4"
sample_video = load_video(video_path)[:100]
sample_video.shape
```

```
(100, 224, 224, 3)
```

```
to_gif(sample_video)
```



```
predict(sample_video)
```

```
Top 5 actions:
  running on treadmill   : 45.56%
  pushing cart           : 40.44%
  doing laundry          :  6.05%
  pushing car            :  2.37%
  hurdling               :  1.51%
```

VII. CONCLUSION

The growing availability of wearable gadgets and smartphones equipped with various sensors has spurred significant interest in the application of deep learning to the problem of human activity recognition (HAR). Many HAR tasks, like as gesture recognition, fall detection, and activity recognition in sports and healthcare, have been mastered with surprising success by deep neural networks (DNNs). Several types of deep learning architectures, including as convolutional neural networks (CNNs), recurrent neural networks (RNNs), and hybrid architectures, have all been proposed for HAR. These designs have been used to data from many different types of sensors, including accelerometers on smartphones and wearable devices, with excellent results on multiple benchmark datasets.

However, there are still difficulties in HAR employing deep learning, such as managing sparse and variable sensor data and addressing the interpretability of DNN models. Addressing these issues and identifying new use cases for deep learning-based HAR in fields like healthcare, sports, and home automation are priorities for the field moving forward. Therefore, HAR powered by deep learning has the potential to transform human-computer interaction and enable a plethora of applications that could enhance people's lives all around the globe.

VIII. REFRENCES

- Vrigkas, M., Nikou, C., & Kakadiaris, I. A. (2015). A review of human activity recognition methods. *Frontiers in Robotics and AI*, *2*, 28.
- Ke, S. R., Thuc, H. L. U., Lee, Y. J., Hwang, J. N., Yoo, J. H., & Choi, K. H. (2013). A review on video-based human activity recognition. *Computers*, *2*(2), 88-131.\
- Kim, E., Helal, S., & Cook, D. (2009). Human activity recognition and pattern discovery. *IEEE pervasive computing*, *9*(1), 48-53.
- Banos, O., Galvez, J. M., Damas, M., Pomares, H., & Rojas, I. (2014). Window size impact in human activity recognition. *Sensors*, *14*(4), 6474-6499.
- Ann, O. C., & Theng, L. B. (2014, November). Human activity recognition: A review. In *2014 IEEE international conference on control system, computing and engineering (ICCSCE 2014)* (pp. 389-393). IEEE.
- Robertson, N., & Reid, I. (2006). A general method for human activity recognition in video. *Computer Vision and Image Understanding*, *104*(2-3), 232-248.
- Wang, Y., Huang, K., & Tan, T. (2007, June). Human activity recognition based on r transform. In *2007 IEEE conference on computer vision and pattern recognition* (pp. 1-8). IEEE.
- Raptis, M., & Sigal, L. (2013). Poselet key-framing: A model for human activity recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2650-2657).
- Zhang, H., Hua, Y., Wang, C., Li, R., & Zhao, Z. (2020). Deep learning based traffic and mobility prediction. *Machine Learning for Future Wireless Communications*, 119-136.
- Abbasloo, S., Yen, C. Y., & Chao, H. J. (2020, July). Classic meets modern: A pragmatic learning-based congestion control for the internet. In *Proceedings of the Annual conference of the ACM Special Interest Group on Data Communication on the applications, technologies, architectures, and protocols for computer communication* (pp. 632-647).