Coimbatore Institute of Technology
Government Aided Autonomous Institution Approved by AICTE
New Delhi and Affllated to Anna University Chennai-600025

CAT 2 PROJECT

Human Activity Recognition

NAME MAHIMA A [1932024]

VANDANA A [1932055]

DATE 17 / 04 / 2023

SUBJECT Image Processing Lab

[16MDSEL3]

FACULTY Dr. P. Velvadivu

INTRODUCTION

Human activity recognition (HAR) is determining the activities performed by a person. Movements are often activities like standing , Walking , Running ,Falling and so on. Human activity recognition is basically determining the activities performed by the person under study with the help of 3D landmarks.

AIM:

The aim of the project is to build a model that will extract the coordinates of the input image and predict the activity performed by the subject and use parallel computing to make this process easier and less time consuming.

KEYWORDS:

Activity recognition, Pose detection, Human Activity Prediction, XGBoost model, Computer Vision, Parallel computing

ABSTRACT:

The idea is that once the subject's activity is recognized and known, an intelligent computer system can then offer assistance.

Two levels of parallelism are adopted to perform the operations in a faster manner. First level is at Image annotation and extracting landmarks. Second at the recognition stage which recognises the type of activity. The features are extracted from the given image using mediapipe and classify it with the XGBoost model.

Human activity recognition (HAR) is determining the activities performed by a person by inferring 33 landmarks and background segmentation masks on the whole body from RGB video frames utilizing our Blaze Pose detection.

PROCESS

The task is divided into three segments that are executed parallelly as follows :

1. Generating dataset and reading
2. RGB to Grayscale conversion
3. Image annotation & landmark extraction
4. Prediction of activity using XGBoost model

1. Generating dataset and reading:

Images of different activities performed by different persons are recorded and loaded. The images collected are random samples generated from google, Youtube Videos and personal data. These images are made of the same size and converted into B&W format using the CV2 package.

2.RGB to Grayscale conversion

The snippet cv2.cvtColor (img,cv2.COLOR_BGR2GRAY) can be used to do the processing. Using parallel processing the images are converted and stored in a new folder which is then used to extract landmarks.

The landmarks extracted from the BW images are fed into a csv file as 66 different attributes with x and y coordinates for each landmark.

3. Image annotation & landmark extraction:

The activities performed by a person is determined by inferring 33 landmarks and background segmentation masks on the whole body from RGB image frames utilizing our BlazePose detection. Draw segmentation on the image. To improve segmentation around boundaries, consider applying a joint bilateral filter to "results.segmentation_mask" with "image".Next draw pose landmarks on the images of the respective areas.

A list is created to append these values and used to train the model.

4. Prediction of activity using XGBoost model:

The final part is to detect the main landmarks and predict the activity being performed by the subject. Using the XGBoost algorithm, the model is trained and activity is recognised. .

MODELS USED:

FRAMEWORK WITH TECHNIQUES :

- CV2 - *Computer vision [Image reading]*
- PIL - *To load Images*
- PICKLE - *To dump model's trained data and load*
- XGBoost - *Model used for prediction*
- Image processing techniques
    - RGB to Grayscale conversion
    - Image segmentation
    - Image annotation
    - Edge Detection
    - ML Prediction

DATASET USED:

The HAR dataset contains images of humans performing 5 different actions. Each landmark activity consists of two attributes X and Y

The target column takes values from 0 to 4 based on the corresponding activity

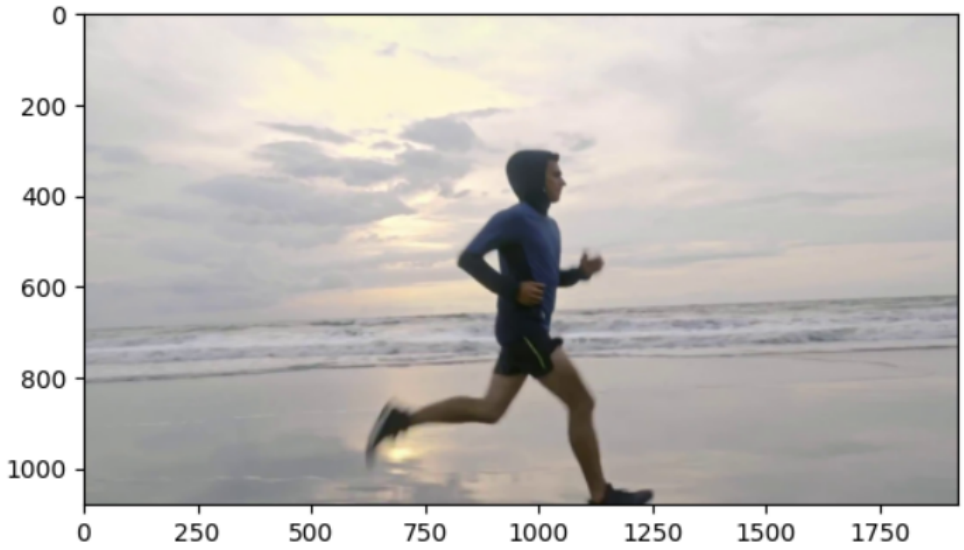
	0	1	2	3	4	5	6	7	8	9	...	57	58	59	60	61	62	63	64	65	66
0	1328	765	1339	774	1339	778	1337	781	1343	764	...	693	1289	835	1272	698	1369	810	1321	688	0
1	1309	720	1326	734	1327	737	1327	740	1328	730	...	644	1131	627	1178	640	1189	561	1233	631	0
2	1331	771	1340	772	1341	769	1341	767	1340	772	...	937	1193	607	1269	943	1248	560	1336	971	0
3	1200	696	1206	693	1207	694	1211	694	1204	690	...	720	1140	719	1115	726	1176	692	1144	712	1
4	1190	699	1195	692	1197	692	1199	692	1193	693	...	895	1161	964	1205	902	1172	987	1227	925	1

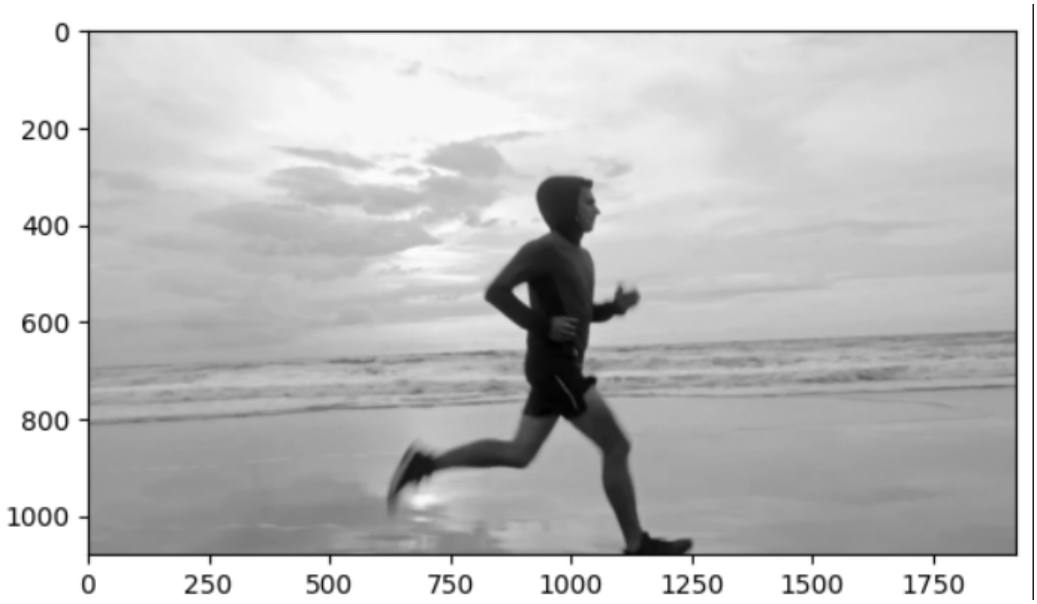


OUTPUT SCREENSHOTS:

1.IMAGE PROCESSING

a. RGB to Grayscale conversion





b. Edge Detection:

Original Image

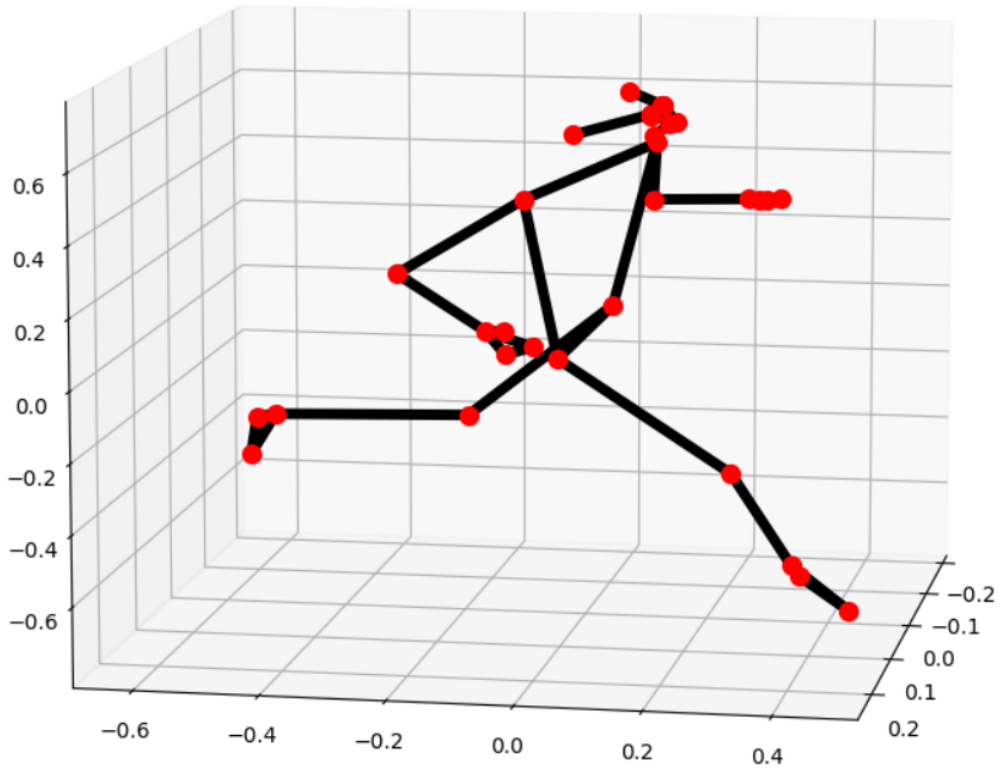
Sobel Edge Image

Original Image

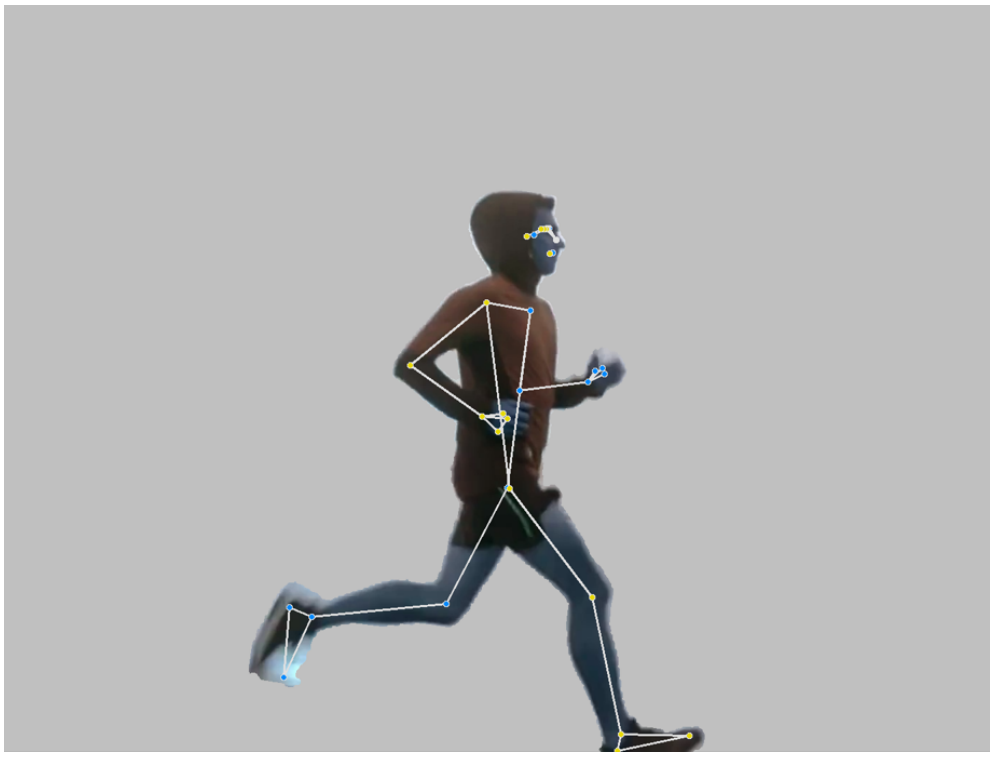
Prewithh Edge Image

Original Image

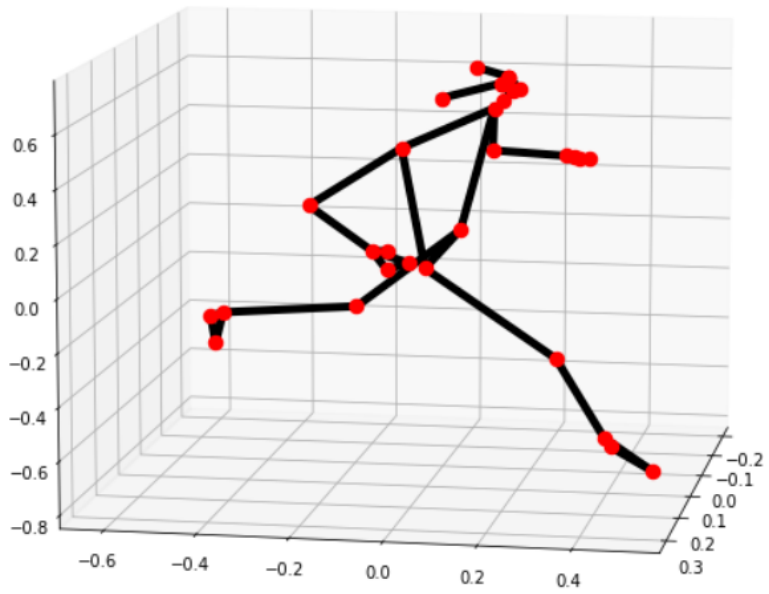
Canny Edge Image



c. Segmentation:



2.ACTIVITY RECOGNITION:



RESULTS AND CONCLUSION:

- Using XGBoost model the activities performed by the person is successfully predicted and recognised using various image processing techniques
- Further, Activity recognition can be done using video clips and also predict images recorded from webcam.

INFERENCE:

We can see that canny edge detection has the best performance as compared to the other two methods as we can clearly detect the outline of the person whose activity we has to recognize.

REFERENCE:

PAPER:

https://www.researchgate.net/publication/282298953_Human_activity_recognition_A_review

MEDIA PIPE: https://google.github.io/mediapipe/

IMAGE PROCESSING TECHNIQUES: https://towardsdatascience.com/common-image-processing-techniques-in-python-e768d32813a8