JAYPEE INSTITUTE OF INFORMATION TECHNOLOGY

ML Based Human Identification using Gait Pattern and Computer Vision

Submitted By:

Rachit Goel (17103146)

Shagun Agarwal (17803003)

Astitva Singhal (17103136)

Submitted To:

Dr. Ankit Vidyarthi

CONTENTS

ABSTRACT

Our project involves generation and analysis of gait pattern from raw videos of individuals. The main aim is to differentiate among individuals on the basis of their walking pattern with the help of Machine Learning Algorithms. A dataset of around 10 individuals was generated from different angles and then processed to obtain frames that are divided into different Gait cycles. Pose Estimation Algorithms are then applied to the processed frames for extraction of positions of different parts of individuals. The positions of parts thus obtained are used to differentiate people based on two approaches, Gait Parameters and Template Matching. The results from the approaches are then intersected and conclusions are derived.

INTRODUCTION

People's identification is one of the most important topics of concern nowadays. Biometric systems employ methods to uniquely identify humans based upon one or more intrinsic physical or behavioral traits such as the face, fingerprint, retina, iris, etc. Currently used face recognition systems are prone to failures when images are not sharp or the face is obscured due to some reason. Thus there is a need for a new system that is robust to such failures and can strengthen currently employed mechanisms. Gait (walk) pattern has several attractive properties as a soft biometric trait. From the surveillance perspective, gait pattern biometrics is appealing for its possibility of being performed at a distance and without body-invasive equipment or subject cooperation.

Human gait refers to locomotion achieved through the movement of human limbs. Gait evaluation helps to understand the way humans move. In order to analyze and quantify how someone walks, it is necessary to isolate the shortest, unique, repeatable task during gait. This task is called the gait cycle. A single gait cycle is measured from a one-foot strike to the subsequent foot strike of the same foot.

The gait cycle can be broken down into two primary phases, the stance and swing phases:-

- Stance phase-The stance phase of gait begins when the foot first touches the ground and ends when the same foot leaves the ground. The stance phase makes up approximately 60% of the gait cycle.
- Swing phase-The swing phase of gait begins when the foot first leaves the ground and ends when the same foot touches the ground again. The swing phase makes up the other 40% of the gait cycle.

In our project, we have considered 3 phases:

- Heel Strike-The stage in gait at which the heel of the foot or shoe first makes contact with the walking surface.
- Terminal stance-The phase of gait cycle is defined as the time from heel rise until the other limb makes contact with the floor.
- Terminal swing-The final phase of the gait cycle where the lower leg is vertical and ends at initial contact.

For collecting the data we have chosen two datasets. Firstly the CASIA-B datasets and our own dataset.

CASIA-B Dataset: This dataset contains 124 subjects but is Multiview: records are captured from 11 different viewpoints at angles ranging from 0 to 180 degrees. The videos have lengths of 2–3s a piece at a frame rate of 30 fps; we used all of them in our experiments.

All videos are recordings of people of full height walking, captured from the side view. There are several video sequences per subject taken under various conditions.

Overall, there are 10 videos per subject: six normal walks, two walks in coating shoes, and two walks with a backpack.

Our Own Dataset : We have created videos of different people in varied backgrounds to test the robustness of our algorithms. These videos have many gait cycles that are later extracted and separated. These videos are broken down into frames and then individual gait cycles are separated which are further used for extraction of parameters. The videos were recorded from two different angles, lateral and front.

Pose Estimation refers to computer vision techniques that detect Human Figures in images and videos so that one could determine precisely where different body parts of an individual show up on an image. Traditional models could not provide estimations for multiple people in an image directly and had to employ person detectors to detect people in an image and then apply a single person Pose Estimation on each person identified. This was computationally expensive as the runtime was directly proportional to the number of people in the image which used to span in the order of hours. The model used in this report brings it down to seconds by taking a bottoms-up approach employing Part Affinity Fields. Part Affinity Fields are a set of vectors that encode the orientation between points in an image domain. The model detects various potential points using confidence maps and connects them using Part Affinity Fields. Thus the model provides an estimated position of each part for each individual in an image simultaneously in a matter of seconds which is crucial for calculation of gait parameters.

Gait parameters are measures that are used to quantitatively analyze the Gait pattern of a person. Looking at gait from a spatial perspective allows us to measure gait patterns related to the distance between steps and stride lengths. Typical Gait parameters include Cadence, Step Width, Step Length, Stride Length, Stride Velocity, etc. Several studies have shown that each person has a distinctive way of walking and these parameters can be used to analyze the gait pattern of different individuals. Among several gait parameters, we have chosen 5 parameters that have been described below:-

- Step Length - Distance between heel contact of the opposite foot.
- Stride Length - Distance between heel contact of the same foot
- Leg Length - The length of the legs from hip to ankle
- Step Angle - The angle between the axis of both legs when both legs touch the ground
- Stride Velocity - Distance covered by the body in unit time

Template Matching is another feature that we have added to our project to get more precise results. We process the videos and divide them into image frames. In this, we take the first frame of the test and the input video and merge them into one frame for comparing them side by side. Then we apply pose estimation on this image frame. Pose estimation plots the heat points over the image and then we draw horizontal lines to compare the similarity between the input image frame and the test frame. We calculate the percentage match between the test and the input image by checking the deviation of the horizontal lines.

Problem Statement

To develop ML based human identification system using Gait Biometric Parameters and Computer Vision based Template Matching

LITERATURE SURVEY

1. Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields:

We present an approach to efficiently detect the 2D pose of multiple people in an image. The approach uses a nonparametric representation, which we refer to as Part Affinity Fields (PAFs), to learn to associate body parts with individuals in the image. The architecture encodes global context, allowing a greedy bottom-up parsing step that maintains high accuracy while achieving real time performance, irrespective of the number of people in the image. The architecture is designed to jointly learn part locations and their association via two branches of the same sequential prediction process.

(a) Input Image (c) Part Affinity Fields (d) Bipartite Matching (e) Parsing Results

The system takes, as input, a color image of size w×h and produces, as output, the 2D locations of anatomical key-points for each person in the image . First, a feed-forward network simultaneously predicts a set of 2D con-fidence maps S of body part locations and a set of 2D vector fieldsLof part affinities, which encode the degree of association between parts .

The set $S=(S1,S2,...,S_J)$ has J confidence maps, one per part, where $S_j \in \mathbb{R}^{w\times h}, j \in \{1...J\}$.

The set $L= (L_1,L_2,...,L_C)$ has C vector fields, one per limb1, where $L_c \in \mathbb{R}^{w\times h\times 2}$, $c \in \{1...C\}$, each image location encodes a 2D vector. Finally, the confidence maps andthe affinity fields are parsed by greedy inference to output the 2D key points for all people in the image.

2.Uncooperative MoCap Gait Identification For Video Surveillance with Incomplete and Noisy Data

This work offers a design of a video surveillance system based on a soft biometric -- gait identification from MoCap data. The main focus is on two substantial issues of the video surveillance scenario: (1) the walkers do not cooperate in providing learning data to establish their identities and (2) the data are often noisy or incomplete. We show that only a few examples of human gait cycles are required to learn a projection of raw MoCap data onto a low-dimensional subspace where the identities are well separable. Latent features learned by Maximum Margin Criterion (MMC) method discriminate better than any collection of geometric features. The MMC method is also highly robust to noisy data and works properly even with only a fraction of joints tracked. The overall workflow of the design is directly applicable for a day-to-day operation based on the available MoCap technology and algorithms for gait analysis. In the concept we introduce, a walker's identity is represented by a cluster of gait data collected at their incidents within the surveillance system: They are how they walk.

3.GaitSet: Regarding Gait as a Set for Cross-View Gait Recognition

We begin with formulating our concept of regarding gait as a set. Given a dataset of *N* people with identities y_i , $i \in 1, 2, ..., N$, we assume the gait silhouettes of a certain person subject to a distribution P_i which is only related to its identity. Therefore, all silhouettes in one or more sequences of a person can be regarded as a set of *n* silhouettes $Xi = \{x^j_i \mid j = 1, 2, ..., n\}$, where $x^j_i \sim P_i$. Under this assumption, we tackle the gait recognition task through 3 steps, formulated as

$$f_i = H(G(F(X_i)))$$

where *F* is a convolutional network aims to extract frame level features from each gait silhouette. The function *G* is a permutation invariant function used to map a set of frame level features to a set-level feature . It is implemented by an operation called Set Pooling (SP). The function *H* is used to learn the discriminative representation of *Pi* from the set level feature. This function is implemented by a structure called Horizontal Pyramid Mapping (HMP). The input *Xi* is a tensor with four dimensions, i.e. set dimension, image channel dimension, image height dimension, and image width dimension.

4. A Survey of Gait Recognition Based on Skeleton Model for Human Identification

Biometric means the identification of persons by their traits or characteristics. We study a simple survey and a general review of gait recognition based on a skeleton model for human identification of recent gait progresses. Model-based approaches obtain a series of dynamic or static body parameters via tracking or modeling body components such as legs, limbs, thighs and arms. Gait signatures derived from these parameters are then employed for recognizing or identifying a person. They have the advantage of view invariant and scale-independent recognition. Though, model-based approaches have sensitivity to the quality of gait sequences and also suffer high computing costs due to parameters calculation. Model-based methods have an advantage of insensitivity to carrying variation and clothing (Li and Chen, 2013). Research examples of this approach are static body parameters, thigh joint trajectories, dual oscillator, articulated model and 2D stick figure.

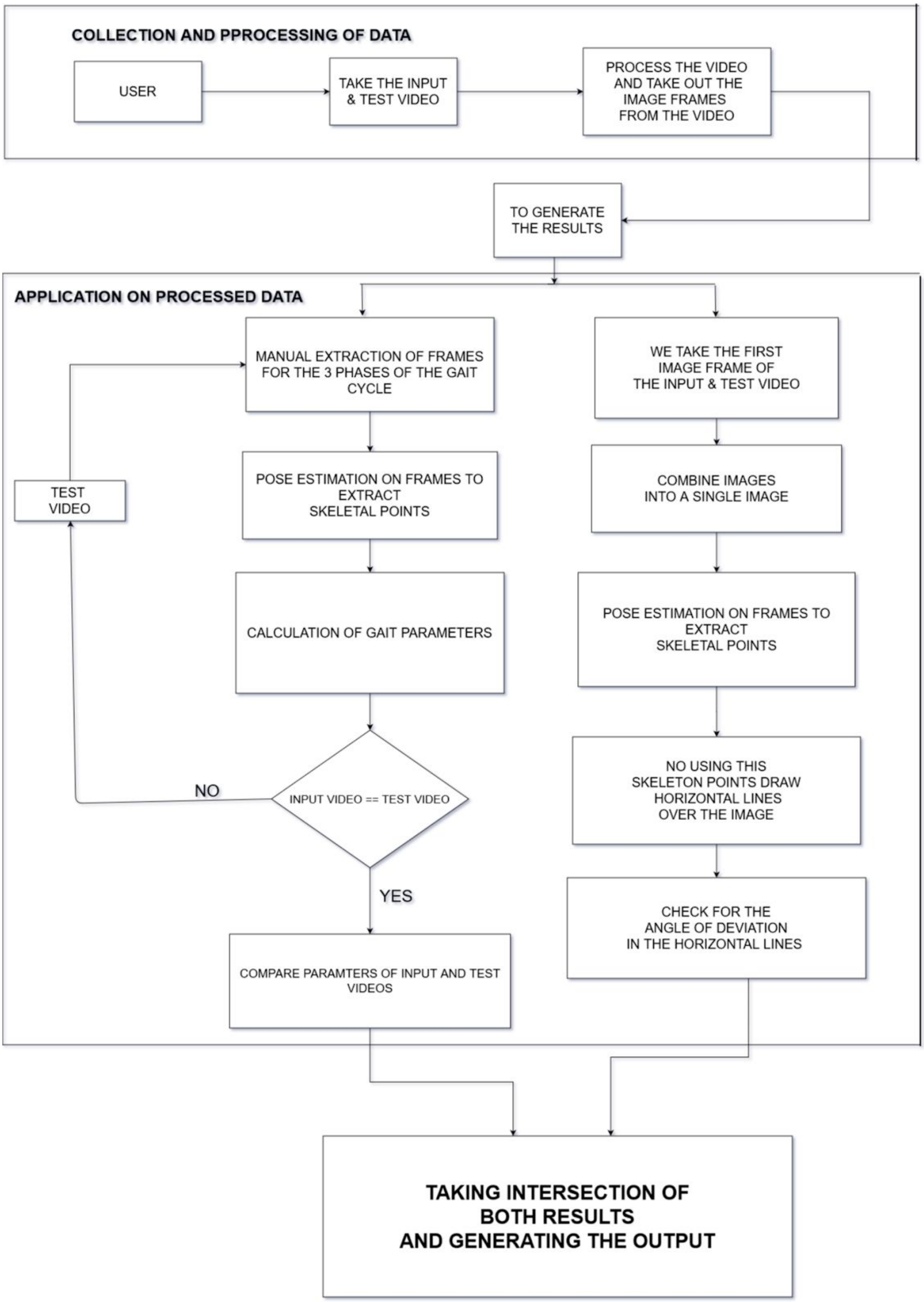
5. GlidarCo: gait recognition by 3D skeleton estimation and biometric feature correction of flash lidar data

A model-based gait recognition method, using sequences recorded by a single flash lidar. Existing state-of-the-art model-based approaches that exploit features from high quality skeletal data collected by Kinect and Mocap are limited to controlled laboratory environments. The performance of conventional research efforts is negatively affected by poor data quality. Here they address the problem of gait recognition under challenging scenarios, such as lower quality and noisy imaging process of lidar, that degrades the performance of state-of-the-art skeleton-based systems. We present GlidarCo to attain high accuracy on gait recognition under the described conditions. A filtering mechanism corrects faulty skeleton joint measurements, and robust statistics are integrated to conventional feature moments to encode the dynamic of the motion. As a comparison, length-based and vector-based features extracted from the noisy skeletons are investigated for outlier removal.

6) Gait Recognition from Motion Capture Data

Gait recognition from motion capture data, as a pattern classification discipline, can be improved by the use of machine learning. This article contributes to the state of the art with a statistical approach for extracting robust gait features directly from raw data by a modification of Linear Discriminant Analysis with Maximum Margin Criterion. Experiments on the CMU MoCap database show that the suggested method outperforms 13 relevant methods based on geometric features and a method to learn the features by a combination of Principal Component Analysis and Linear Discriminant Analysis. The methods are evaluated in terms of the distribution of biometric templates in respective feature spaces expressed in a number of class separability coefficients and classification metrics. Results also indicate a high portability of learned features, which means that we can learn what aspects of walk people generally differ in and extract those as general gait features. Recognizing people without needing group-specific features is convenient, as particular people might not always provide annotated learning data.

PROPOSED METHODOLOGY



We have first obtained the dataset from the internet of around 124 people.It is the CASIA-B dataset.It is mulviewed and has been captured from 11 different viewpoints at different angles ranging from 0 to 180 degrees.These videos has a definite length of 2-3s at a frame rate of 30fps.This dataset consists of around 10 videos per subject captured from a side view. Per subject: six normal walks, two walks in coating shoes, and two walks with a backpack.

We also tried to generate our small dataset for verification of results and also to explore a way of generating dataset .These videos have been captured in varied backgrounds and are being made from two different angles, lateral and front.

Now the videos are taken as input video and test video.

The video is then processed and broken down into image frames using OpenCV to obtain different phases of a single gait cycle.

Now different frames are obtained and pose estimation is applied on each frame . Then the skeleton is drawn on it .

From this step we divide our work going through two different approaches where one of the approaches is identification using gait parameters and the second one using template matching.

For the first approach we first have to manually find 3 phases of a gait cycle namely heel strike,terminal stance and terminal swing from the generated frames .These phases will be used for calculating gait parameters.The frames consisting these phases are put into seperate folders comprising of one gait cycle.

The Pose Estimation Model draws a heat map from the image to identify potential point of interests and then draws part affinity fields to map the flow of edges between these points and thus marks the poses’ of various individuals and provides coordinates of the points.We have taken into account mainly 6 points out of the 18 points provided by the model. We use Left Ankle, Right Ankle, Left Knee, Right Knee, Left Hip, Right Hip for calculating the gait parameters.

5 Gait Parameters namely Leg Length, Step Length, Stride Length, Stride Velocity and Angle between both Legs, which can be collectively used to uniquely identify an individual’s Gait Pattern are calculated this step.

On the other hand while going through the second approach we took two frames , one is the input frame and another is the test frame.We take the first image of the test video and the input video as at the start of the video every person is in the position to move forward and that is the best time to identify the person. These two frames are used for template matching at a time.

These two frames are combined into a single frame and put side by side using OpenCv for further processing.

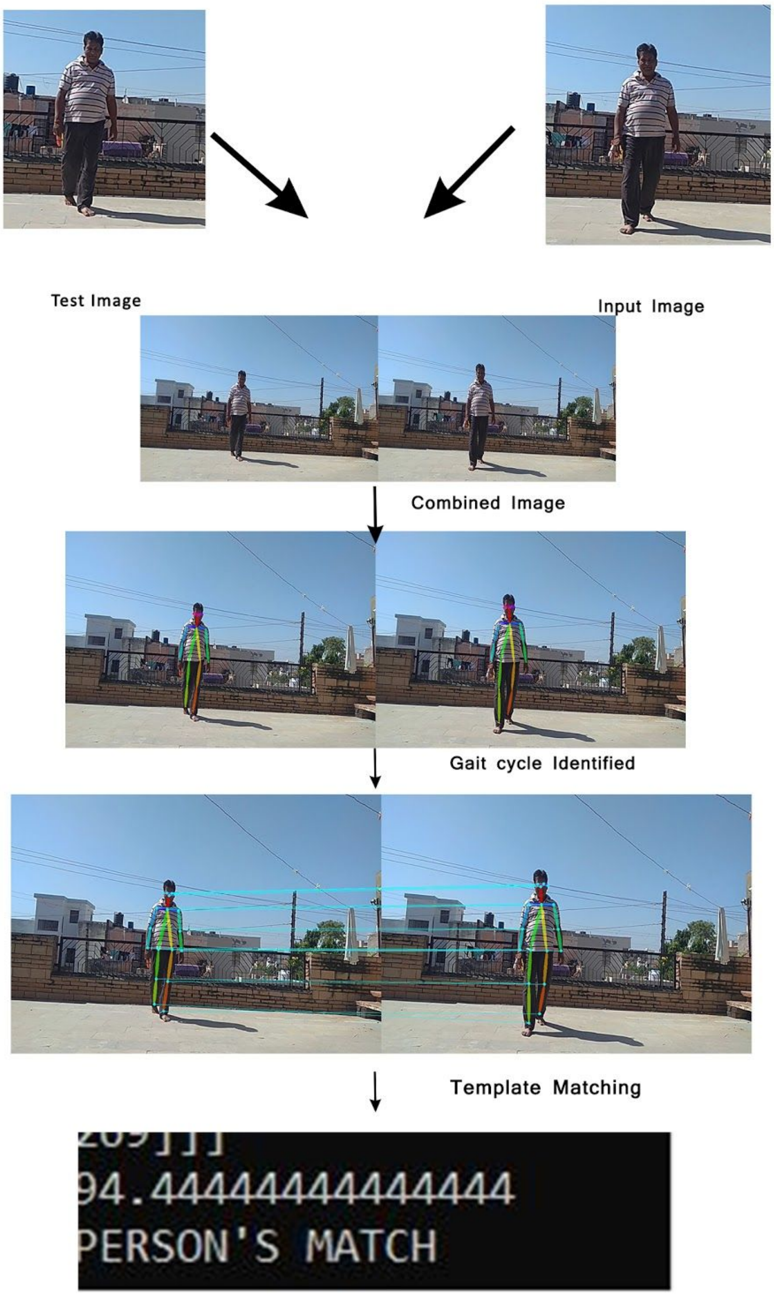
Here also this frame is processed to find the skeleton points as we did on our first approach.The coordinates are plotted through the pose estimation model over the body of the test image and input image.

Now corresponding skeletal points on both the images are joined by lines and angular deviation from horizontal axis is measured which if exceeds a certain threshold is indicative that both the images represent different individuals.

Now the results from both the above mentioned approaches are combined and represented below.

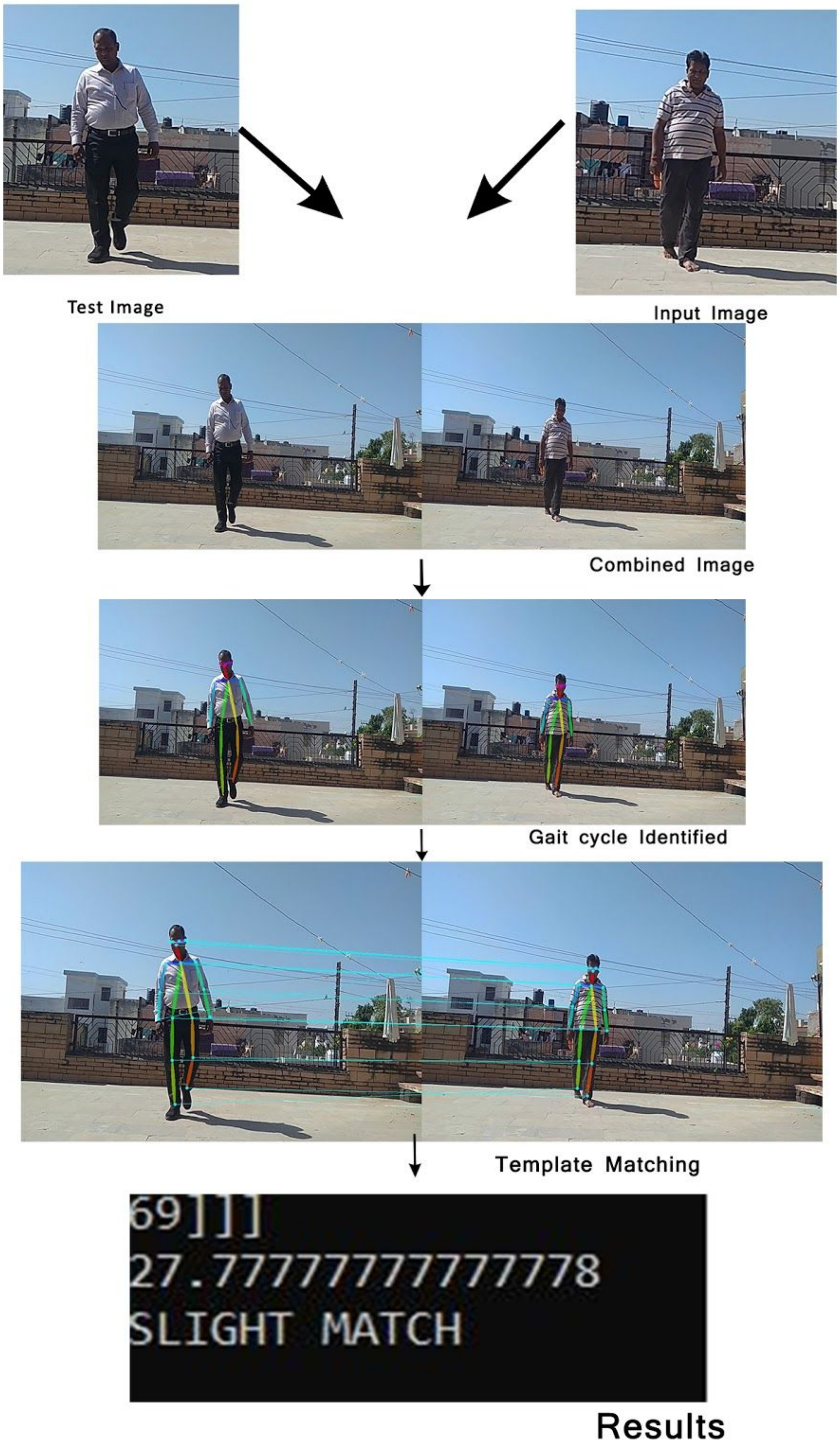
RESULT

- <u>Template Matching Taking Same Person</u>



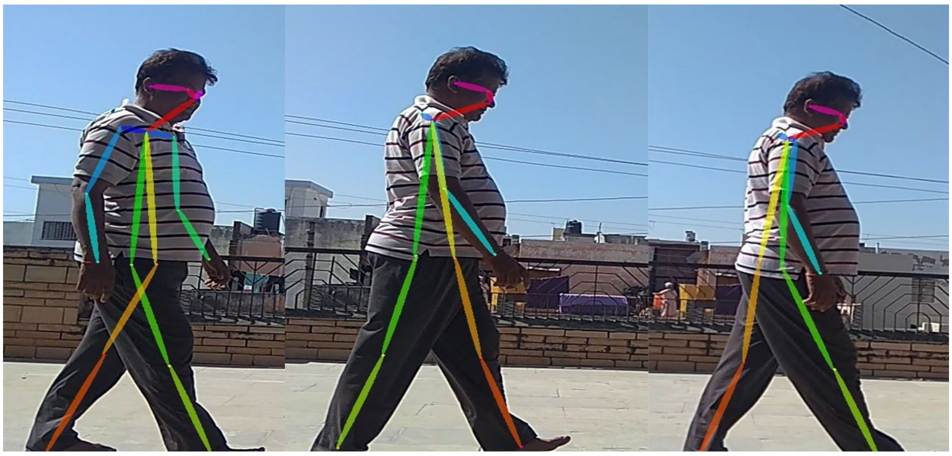
Results

- <u>Template MatchingTaking Different Persons</u>



- <u>Comparison of Gait Parameters of Different persons</u>

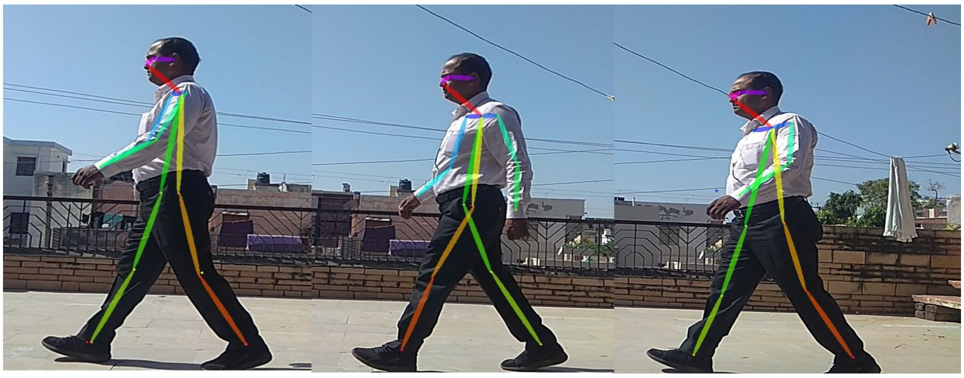
Subject 1



Sample of Parameters

Step Angle	Leg Length	Step Length	Stride Length	Stride Velocity
38	318	215	422	639

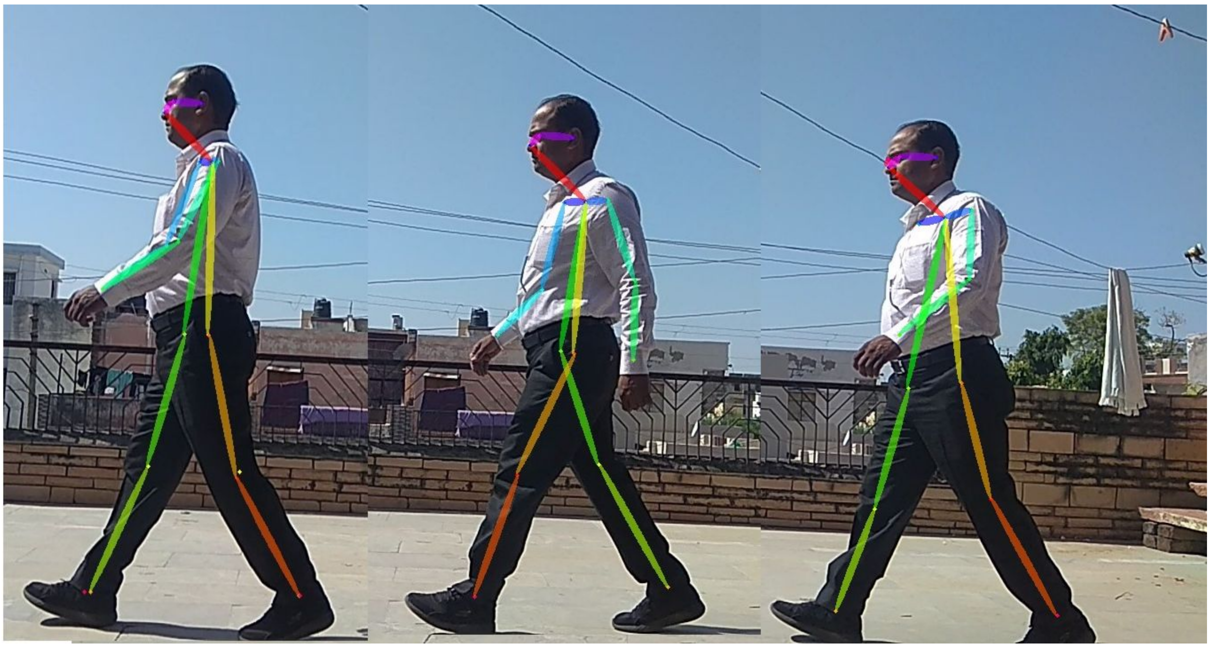
Subject 2



Sample of Parameters

Step Angle	Leg Length	Step Length	Stride Length	Stride Velocity
45	300	238	484	733.3

Subject 2 (2nd)



Sample of Parameters

Step Angle	Leg Length	Step Length	Stride Length	Stride Velocity
47	289	234	492	745

Cosine Similarity for Subject 1 and Subject 2 = 0.6

Cosine Similarity for Subject 2 and Subject 2 (2nd) = 0.99

Conclusion

We were successfully able to differentiate our subjects in the dataset. The Template Matching provided high similarity for the same subject and far lower for different subjects. The Cosine Similarity in Gait Parameters was also quite indicative.

References

1. Zhe Cao, Tomas Simon, Shih-En Wei, Yaser Sheikh ;" Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields"
2. Michal Balazia, Petr Sojka; "You Are How You Walk: Uncooperative MoCap Gait Identification for Video Surveillance with Incomplete and Noisy Data"
3. Hanqing Chao, Yiwei He, Junping Zhang, JianFeng Feng; "GaitSet: Regarding Gait as a Set for Cross-View Gait Recognition"
4. M.D. Jan Nordin and Ali Saadoon Center for Artificial Intelligence Technology;" A Survey of Gait Recognition Based on Skeleton Model for Human Identification"; Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, Selangor, Malaysia
5. GlidarCo: gait recognition by 3D skeleton estimation and biometric feature correction of flash lidar data
6. MICHAL BALAZIA and PETR SOJKA, Masaryk University; "Gait Recognition from Motion Capture Data"