PES UNIVERSITY
(Established under Karnataka Act No. 16 of 2013)
100 Feet Ring Road, Bengaluru, Karnataka, India – 560 085

Project report on

'Biometrics Using Gait Analysis'

Submitted by

Prabhu Shankar GV (01FB16EEC192)

Ritesh S (01FB16EEC230)

Rithvik C (01FB16EEC231)

Jan – May 2020

under the guidance of

Prof. Raghavendra MJ
Designation- Assistant professor
Department of Electronics and Communication Engineering
PES University
Bengaluru – 560085

FACULTY OF ENGINEERING

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

PROGRAM B. TECH.

CERTIFICATE

This is to certify that the Report entitled

'Biometrics Using Gait Analysis'

is a bonafide work carried out by

Prabhu Shankar GV (01FB16EEC192)

Ritesh S (01FB16EEC230)

Rithvik C (01FB16EEC231)

In partial fulfillment for the completion of 8th semester course work in the Program of Study B. Tech. in Electronics and Communication Engineering, under rules and regulations of PES University, Bengaluru during the period Jan – Apr 2020. It is certified that all corrections/suggestions indicated for internal assessment have been incorporated in the report. The dissertation has been approved as it satisfies the 8th semester academic requirements in respect of project work.

Signature with date & Seal	*Signature with date & Seal*	*Signature with date & Seal*
Prof. Raghavendra MJ	*Dr. Anuradha M*	*Dr. Keshavan BK*
Internal Guide	*Chairperson*	*Dean of Faculty of Engineering*

Name and signature of the examiners:

1.
2.

DECLARATION

We, **Prabhu Shankar G V**, **Ritesh S** and **Rithvik C**, hereby declare that the project report entitled, '**Biometrics Using Gait Analysis**', is an original work done by us under the guidance of **Prof. Raghavendra MJ**, Designation, Dept. of ECE and is being submitted in partial fulfillment of the requirements for completion of 8th Semester course work in the Program of Study B. Tech. in Electronics and Communication Engineering.

PLACE:
DATE:

NAME AND SIGNATURES OF THE CANDIDATES

1. Prabhu Shankar GV
2. Ritesh S
3. Rithvik C

ACKNOWLEDGEMENT

We would like to sincerely thank our guide and mentor **Mr. Raghavendra MJ, Assistant Professor, Electronics and Communication Department**, PES University for his invaluable guidance, constant assistance, support, remarkable mentoring, endurance and constructive suggestions for the betterment of the project. We would like to express our deep sense of gratitude to **Prof. M.Rajasekhar,chief project coordinator , Dr.Anuradha M , HOD , Electronics and Communication Department , Dr. B.K.Keshavan , Dean of Faculty , Dr. J Surya Prasad, Vice Chancellor , Prof. Jawahar Doreswamy , Pro Chancellor ,** and **Dr. Doreswamy , Chancellor** and **PES University** for giving us the opportunity to embark upon this topic and work on this project. We would also like to thank all the teaching and non-teaching staff, family and friends for providing us the moral backing in realizing this project along with the facilities and conducive environment required for this project.

Thank you all

ABSTRACT

In the current world scenario there has been a drastic increase in the criminal wrongdoings. The perpetrators in order to get away from such crimes are using masks, gloves, sunglasses. Which would indeed help them cover their identity. In this project , a person's identity is recognized using his gait patterns. This model can be used for various other situations instead of fingerprint biometrics .

The first part consists of creating the parameters of the dataset of a person walking, which was done by recording the videos of us walking at different speeds and varying step widths . Then using image segmentation techniques, videos are converted into silhouettes. From the segmented images, required features such as persons body length, body width, step width are extracted during each time frame.

The second part of this project is training the neural network using the extracted gait patterns. Which upon training and analyzing the given gait patterns appropriate results are obtained. This project is done using Matlab and various toolboxes available on Matlab.

Table of Contents

List of figures

List of tables

Chapter 1: Introduction

Biometric technology analyzes human characteristics to verify a person using the person's fingerprints, eyes , face, gait, heart beat , and voice recordings. Gait is defined as locomotion of human beings, more importantly, walking style of human beings. In the process of walking, a foot from the front, reaching the back and coming back to the front is called one Gait Cycle.

This profile of a person is unique and differentiable from person to person. Hence it can be used for biometrics. It can be used for various purposes such as chiropractic and osteopathic, but in this project, it is mainly focused on biometrics usage of gait. Gait in biometrics can be carried out at farther distances or at low resolutions.The main parameters in any gait analysis is dynamic base, step length, cadence, stride length, speed. Since the gait here is adopted for biometrics, only novel features: body width, leg width(step length), leg length and body length are required. This study emphasizes the quantification of these novel features and used for biometrics.

1.1 Problem statement

To analyse and recognize the walking style of a person for biometric operations

In today's world the perpetrators are trying to evade the current biometrics using various gadgets and are getting away from their criminal activities. It is very important to control such activities. So implementing a gait recognition model which would help in recognizing the person with the help of the gait patterns along with the neural network for training the patterns.

The project has been done on the following parameters:

- To Intervene Biometrics using Gait Analysis
- Database Creation for Gait analysis
- Smoothening of Boundaries in each frame using Gaussian and Median Filters
- Conversion of Frames into Grayscale images
- Extracting Binary Silhouette from Grayscale images
- To Use Features which cannot be faked and lead to malicious way
- Training the Neural network to get optimised output
- To use Neural Network for minute Variations in Gait analysis
- Identification of a person

1.2 Motivation

Security and Privacy go hand in hand in this current digital age, we have to make sure that an individual's privacy is not interceded. Biometrics is a way that a person's identity cannot be copied by anyone and it can be used for many security and privacy purposes.

Biometrics are Fingerprint scanners, Iris Scanners, Retinal Scanners, Facial Recognition, Voice Recognition. These Biometrics can be faked and used for malicious purposes.

Gait as mentioned, is a walking style of a person, this cannot be faked and used in an harmful way. Gait may change over a course of 5-6 years, but not in the span of 12-24 months (Traumatic Experiences are not taken into consideration).

Since this Biometrics has lesser limitations than all the other forms of biometrics, we have considered to work on this particular project.

Apart from applying gait studying for security purposes such as crime avoidance, it can be used for medical motives such as chiropractic and osteopathic. Moreover, gait biometrics can be used for public surveillance as well. In places such as airports and shopping malls.

Chapter 2: Literature Survey

[1] Z. Liu and S. Sarkar, "Improved gait recognition by gait dynamics normalization," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 6, pp. 863-876, June 2006.

This paper mainly proposes about the different approaches to the gait recognition. The three gait approaches mentioned in this paper are 1) Temporal alignment 2) Static parameter 3) Silhouette approach.

Consideration of a Silhouette based approach has been done in our project. This approach considers the normalisation technique and linear discriminant analysis in order to maximise the differentiation between persons while reducing the variation of the same subject, thus helping in better classification. The factors like walking surface, walking speed and shadow affect the gait recognition. This method partially helps in eliminating certain factors like walking speed by normalisation. Here normalisation mainly helps in emphasising the shape of the silhouette. Thus upon emphasising the shape feature extraction becomes more easier in the further development of the project.

[2] S. Benbakreti and M. Benyettou, "Gait recognition based on leg motion and contour of silhouette," *2012 International Conference on Information Technology and e-Services*, Sousse, 2012, pp. 1-5.

This paper presented an algorithm which recognizes the gait of the person. This algorithm recognizes people by their style of walking.

In this algorithm initially, the video is converted into a frame. In each frame the background is subtracted to segment out the person from the background. This process extracts the moving silhouettes of the subject. This is one of the methods tried in our project(frame difference method). This silhouette is categorised by computing the parameters such as contour of silhouettes and the area of the silhouettes. This paper proposes a combination of appearance based approach and model based approach to extract proper feature parameters. Feature parameters such as height, stride length, cadence and width can be derived by dividing the silhouette into different body parts. From this the features to be extracted in the project and method to extract them has been considered.

[3] P. Nithyakani, A. Shanthini and G. Ponsam, "Human Gait Recognition using Deep Convolutional Neural Network," ***2019 3rd International Conference on Computing and Communications Technologies (ICCCT)*****, Chennai, India, 2019, pp. 208-211.**

Artificial Neural Network has an exceptional quality of deriving complicated, raw data into meaningful results. In this paper, consideration of DCNN (deep convolutional Neural network) has been done. In this paper, GEI (gait energy image) are recorded and fed to a neural network. According to the image matching and frame matching. The output is obtained and results are published. Since GEI extraction and feature vectors cannot be copied from another individual. Main reason to consider this paper is to add deep learning by backward propagation for feature extraction. Doing so, the minimal abnormalities in gait would be curved into the output, giving a optimal output

[4] Han Su, Zhi-Wu Liao and G. Chen, "A gait recognition method using L1-PCA and LDA," *2009 International Conference on Machine Learning and Cybernetics*, Hebei, 2009, pp. 3198-3203.

This paper proposes gait analysis method using PCA (principal component analysis) base L1 normalisation and LDA (linear discriminant analysis) method. L1-PCA represents the features and LDA to analyze and classify these features. The importance of the width feature is stressed in this paper. L1-PCA is used for representation of data, dimension reduction and analysis. LDA helps in maximising the discrimination between features of different persons. This was taken into consideration during the early stages of our project.

[5] K. Moustakas, D. Tzovaras and G. Stavropoulos, "Gait Recognition Using Geometric Features and Soft Biometrics," *in IEEE Signal Processing Letters*, vol. 17, no. 4, pp. 367-370, April 2010.

This paper proposes the use of different features extracted for biometrics. Particularly the features include stride length and height. The use of these features is used to increase the performance of the output obtained by feature extraction for better biometric recognition system. This paper explains about how posterior features can be used for biometrics recognition.

Initially gait recognition process is done to extract silhouette image for feature extraction. Then by using a probabilistic framework, gait score obtained by gait recognition and values of features extracted are combined to generate a final score.

Chapter 3: System Model

3.1 Video segmentation

Video segmentation is a processing method in which video is divided into multiple frames upon which image segmentation happens.

3.2 Image Segmentation

Segmentation of the image is the basic step to analyze images and extract required data from them. It is the process of division of the digital image into many segments. Segmentation is the process to simplify the representation of the image into something more meaningful. It consists of converting an image into a collection of pixels. That is represented by a mask or a labeled image. By converting an image into segments, only the required segments of the image can be processed instead of processing the whole image.

3.3 Testing of Different methods

There are different segmentation techniques available. The two methods tried in the project are :

1. **Frame difference method [1],[2]**
2. **Gaussian and median filter method**

3.3.1 Frame difference method

This method is the simplest form of background subtraction.This method can also be referred to as temporal difference. Uses the frame at time t-1 as the background model for the frame at time t . This technique is sensitive to noise and variations in illumination, and does not consider local consistency properties of the change mask. [1],[2]

3.3.2 Gaussian and median filter method

A Gaussian filter being a linear filter is used to blur the image and reduce unwanted noise. The Gaussian filter will blur edges and reduce contrast in each frame. Whereas a median filter known as a non-linear filter is used in reducing salt and pepper noise. The noise caused by sudden and sharp disturbances in the image signal is the salt and pepper noise. This method has an enhanced edge detection and increased segmentation performance compared to frame difference method.

Chapter 3.4 Matlab

3.4.1 Introduction to Matlab

Matlab is a high-performance language for technical computing. It integrates visualization, computation and programming in an easy-to-use environment. The problems and answers are expressed in similar mathematical expressions. It consists of a family of application-specific solutions called toolboxes. They allow to learn and apply specialised technology. Toolboxes are comprehensive collections of MATLAB functions (M-files) that extend the MATLAB environment to solve particular classes of problems. Studies for which toolboxes are available include wavelets,signal processing, neural networks,control systems, fuzzy logic, simulation and many others.

3.4.2 Tools used in Matlab

1. Computer Vision Toolbox

It involves algorithms, functions and applications for testing and designing computer vision. 3D vision, and video processing systems. Object detection and tracking, as well as feature detection, extraction, and matching can be performed.

2. Database Toolbox

It provides functions and an app for exchanging data with relational and non-relational databases. It enables this exchange by automatically converting between database and Matlab data types.

3. Image Processing Toolbox

It provides a comprehensive set of reference-standard algorithms and workflow applications for image processing, analysis, visualization, and algorithm development. Image segmentation, image enhancement, noise reduction, geometric transformations, image registration, and 3D image processing can be performed. It allows to automate common image processing workflows. It helps to interactively segment image data, compare image registration techniques, and batch-process large data sets.

4. Deep Learning Toolbox

It presents a framework for designing and enforcing deep neural networks with algorithms, pre-skilled fashions. With the help of this toolbox class and regression on photograph, time-collection, and textual content facts may be accomplished.With this, the community can be designed, analyzed, and skilled graphically. It helps to control multiple deep learning experiments, maintain the required training parameters, examine results, and examine code from exceptional experiments. Layer output can be visualized.

5. Statistics and Machine Learning Toolbox

It provides features to explain, analyze the facts. The toolbox gives supervised and unsupervised devices gaining knowledge of algorithms, which includes aid vector machines, k-nearest neighbour, k-manner, hierarchical clustering, Gaussian aggregate models, and hidden Markov Models. Many of the facts and system learning algorithms can be used for computations on records units that are too huge to be saved in memory.

6. Vision HDL Toolbox

It provides pixel-streaming algorithms for the layout and implementation of imaginative and prescient structures on FPGAs and ASICs. It provides a layout framework that supports a diverse set of interface sorts, body sizes, and frame rates. The photo processing, video, and computer vision algorithms inside the toolbox use an architecture appropriate for HDL implementations. The generated HDL code is FPGA-confirmed for frame sizes as much as 8k decision and for high frame video .

3.5 Neural Network

A neural network is a circuit consisting of artificial neurons which can also be called as nodes (referred to in a neural network). An artificial neuron is a mathematical function used as a version of neurons. The neurons are modelled as weights. All inputs are modified for the use of weights and summed. This network can examine from the facts, so it is able to study to realise patterns . Neural network breaks the input into layers of abstraction. Then it's miles educated on the usage of exclusive examples to recognize photos, speech and lots of different matters. The weights are adjusted at the same time as schooling in step with a specific rule with the intention to acquire the required outcomes.[3]

3.5.1 Feed Forward Back Propagation

In the neural network a feed forward back propagation algorithm is used to train. Here the information moves in a single direction, from the input nodes, through the hidden nodes after which to the output nodes. This does not have loops or cycles in the network. In this feed forward neural network , back propagation algorithm helps in efficiently adjusting the weights in order to get the required result.This performance makes it possible to use gradient techniques for using multilayer networks, updating weights to minimise loss. This is how feed forward back propagation works.

3.6 Block diagram of gist of working

3.6.1 Structural silhouette and feature extraction (SSFE)

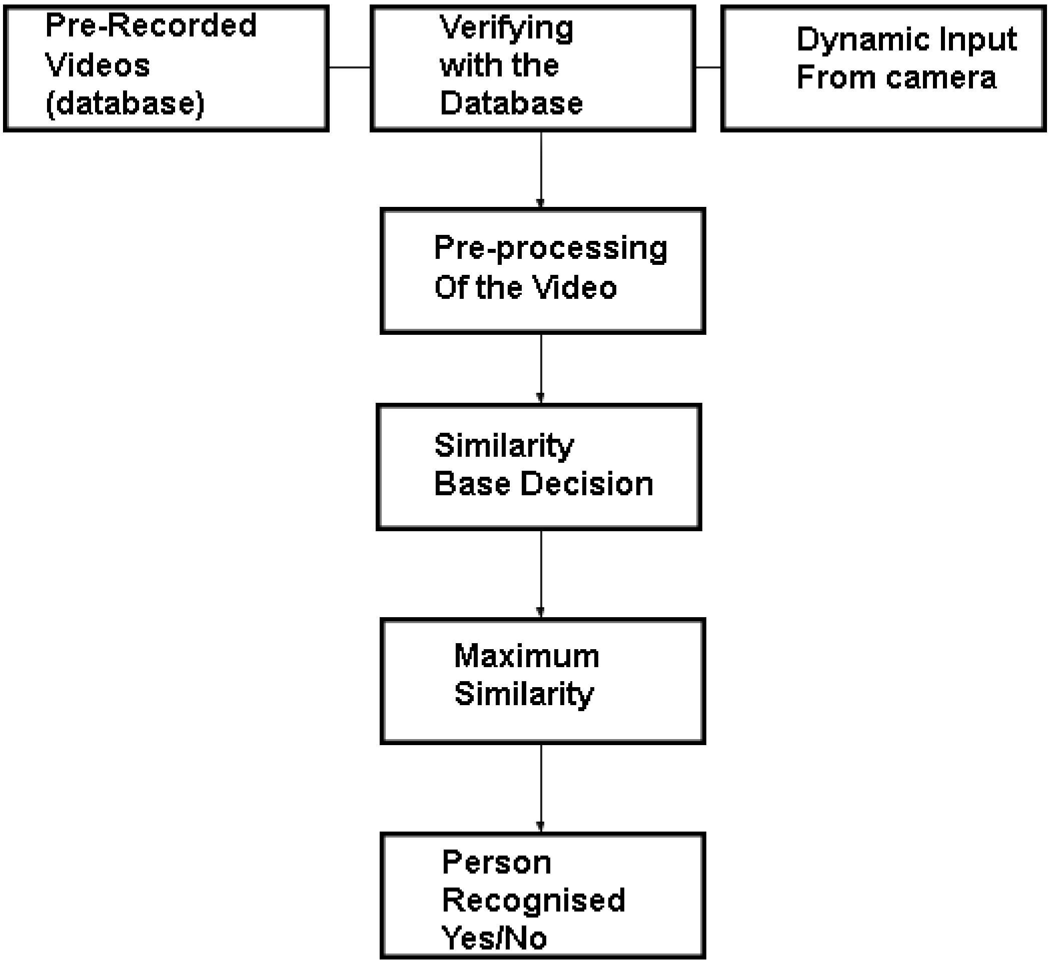


Figure-3.1 Block Diagram of the working of the project

An analog video is recorded using a 25 megapixels camera and is converted into digital format. The standardized video format is about 25 frames per second. Each frame is captured by the camera and played continuously to give video. All the recorded videos are saved in a database, which will be used to examine the video. Dynamic input from the camera is taken and recorded. This video is examined with the pre-recorded database created earlier. Code written, converts the video into desirable output by smoothing the frames, converting it into grayscale image, cropping the required binary silhouette for feature extraction. Silhouette is a black and white image of a person with the boundary either filled with black or white.From the above process different features are extracted which will be used to train the neural network. All the features from the silhouette frames are compared with the trained database to provide a result. This result is then checked for maximum similarity. If there is maximum similarity the person will be recognized.

3.6.2 Neural networks

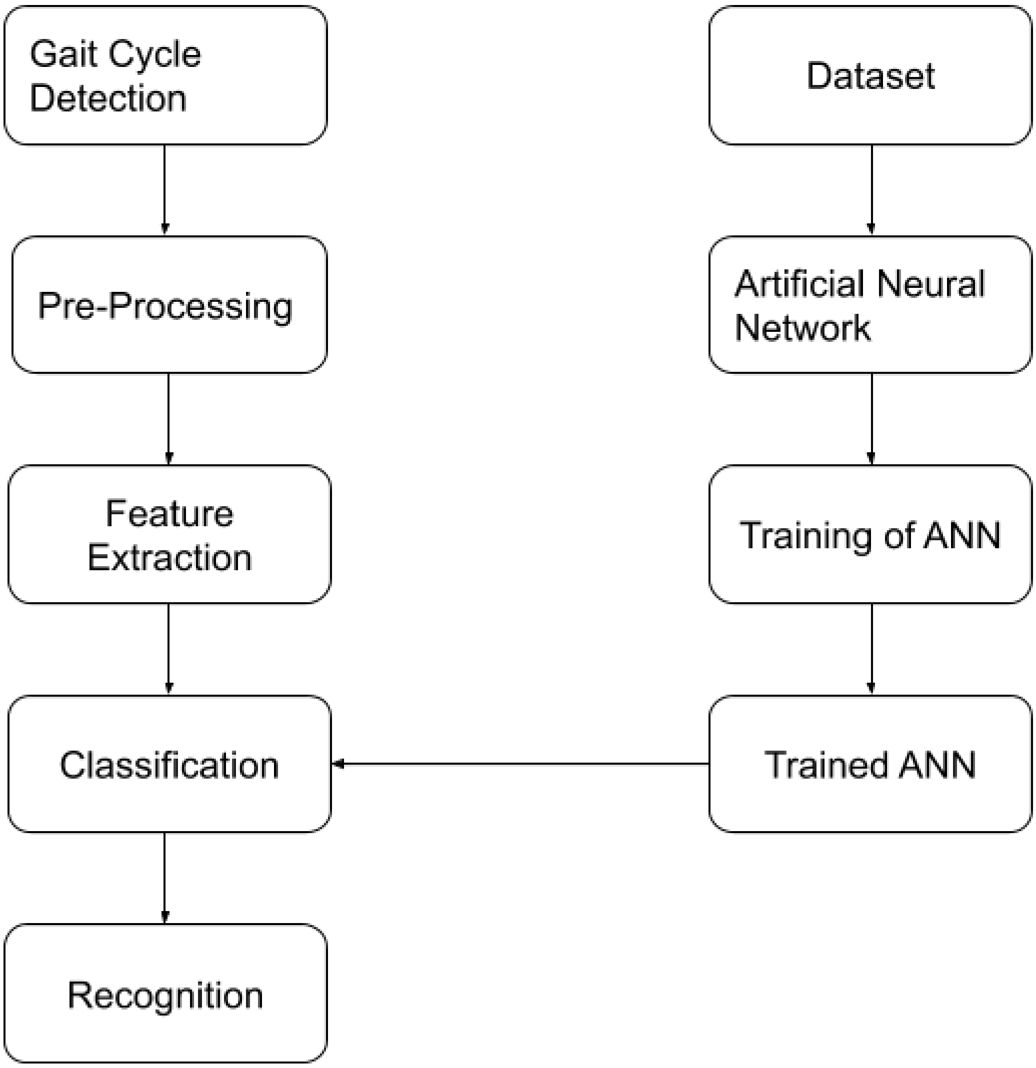


Figure-3.2 Block Diagram of the working of the Neural Network

After Gait Detection has been done, smoothening of the image is done, the extracted features are then used for training the neural network. From the dataset created, the neural network alters the weights and trains the nodes in order to obtain the output. According to the trained weights and nodes, the neural network takes the features from the ongoing video and gives the classification of a person. Concluding with the recognition, leading to biometrics.

Chapter 4: Database creation

4.1 Introduction, what is a database?

Database is the basic information required in a project to get a more efficient result and also a base for training any neural network. The word itself has the meaning in it, 'Data'base, meaning it has some data which can be used for calculations. [1]

Database is an organised collection of data which is accessible by users for calculations and proper training of the neural network.

Database in the digital world is a crucial part of networking, without databases; it would fail to keep a check on abnormal activities, program computers in such a way they are being used now or create Artificial Intelligence.

4.2 Database created

Databases that have been used in the project are our own videos recorded from a 25-megapixel camera having the recording speed of 25 frames per second. The main reason for considering this database is because of the easy access. Initial stage where the algorithm had to be developed and a neural network for the proper training of the weights, this was carried out on the basis of trial and error method until properly segmented video was obtained for better feature extraction. The accurately segmented video is considered as the standard video for all the future references. The videos considered in the database have variations in the gait patterns with respect to the subject i.e.speed, arm-swing, step width. These variations were considered because of the changes in surroundings of the subject at a particular time. These variations are not supposed to intervene in the calculations of the results. Hence a wide range of variations are considered and almost 35 videos of each person has been used to train the Neural network. These videos cover almost all types of gait, including Slow steps, Fast steps, Broad steps, Tapered steps and so on.

For the better weighted average in the Neural Network, more the variations should be considered in the database. It is always known, “The more the merrier”; the same way more videos are pre recorded and saved in the database, the more accurate will be the output.

Each video should be a minimum of 5 seconds and max of 15 seconds i.e minimum of 125 - 375 frames. This is required for the least gait analysis to be performed by the algorithm. Recording lesser than the threshold would be insufficient for the algorithm to perform the segmentation for each frame and extract accurate features. Anything above 375 frames (15 seconds) is just wasting time and data by recording. Hence the ideal database video for the project is between 125 - 375 frames.

Name	Size	Kind
20200421_182334.mp4	12.9 MB	MPEG-4 File
20200421_182839.mp4	10.1 MB	MPEG-4 File
20200421_182901.mp4	8.1 MB	MPEG-4 File
20200421_183051.mp4	11.6 MB	MPEG-4 File
20200421_183113.mp4	12.2 MB	MPEG-4 File
20200421_183133.mp4	10.2 MB	MPEG-4 File
20200421_183156.mp4	11.5 MB	MPEG-4 File
20200421_183214.mp4	11.7 MB	MPEG-4 File
database1.mp4	4.8 MB	MPEG-4 File
database2.mp4	4.4 MB	MPEG-4 File
database3.mp4	3.9 MB	MPEG-4 File
database4.mp4	3.8 MB	MPEG-4 File
database5.mp4	4 MB	MPEG-4 File
database6.mp4	4.1 MB	MPEG-4 File
IMG_6367.mp4	4.3 MB	MPEG-4 File
IMG_6368.mp4	3.1 MB	MPEG-4 File
IMG_6371 (1).mp4	4.3 MB	MPEG-4 File
IMG_6371.mp4	4.3 MB	MPEG-4 File
IMG_6373.mp4	3 MB	MPEG-4 File

Figure-4.1 A screenshot of the database created

Chapter 5: Human detection

5.1 Introduction

Human detection is a type of object detection. Since the object considered in the project is a person, it is referred to as human detection. Human detection(object detection) is a computer vision technique used for identification of the person's movement from the image or video.

5.2 Vision.Foregrounddetector

In the project vision.foregrounddetector function is used to detect the person. This function enables in computing and returning the foreground masks with the use of the gaussian mixture model. This function compares a color or grayscale video frame to a background to determine whether the subject's pixels are part of the background or the foreground. It then finds the value of the foreground mask. These can be in real time from a video camera or from photographs.

The detected images are enhanced using gaussian filter and median filter for the removal of noise from them. Later the RGB image is transformed into a grayscale image which is the best way to extract the edges and other features from it. As a result of this silhouette image is obtained.

Figure-5.1 An example of how the median filter works

Figure-5.2 An example of how the Gaussian Filter works

Chapter 6: Silhouette Extraction

6.1 Silhouette Meaning Explained

A silhouette is a type of image of a person, represented as a shape of a single colour, with its edge of the shape matching the outline of the subject. Usually the solid color will be black or white, with white or black white background simultaneously.

6.2 Usage of Silhouette

When the extraction of the frames from the video was done, each frame will have an image of a person with a different stance of walking. These stances are extracted from the frames and converted into silhouette images.

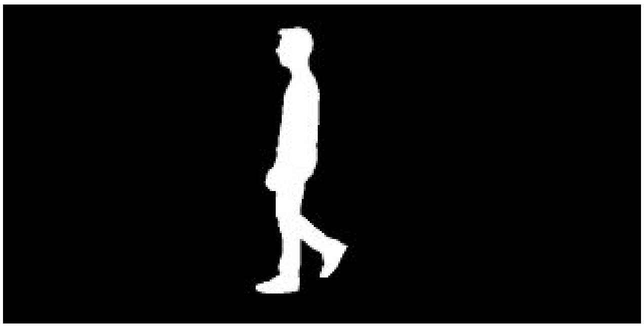


Figure-6.1 Gait silhouette

From these Silhouettes the features are then extracted and used for person detection.

6.3 Silhouette storage.

Each silhouette is stored as a binary image in a matrix form. This is done by converting the RGB to grayscale, and then to binary image. A sample of a matrix is given below.

91×22 logical array

0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0
0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0
0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	0	0
0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0
0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0
0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0
0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0
0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0
0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0
0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0
0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0
0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0
0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0
0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0
0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0
0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0
0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0
0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0
0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0
0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0
0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0
0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0
0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0
0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0

Figure-6.2 An example of the matrix obtained after the extraction of silhouette

The binary image is then used for further calculation of feature vectors to be given for the neural network.

6.4 Advantages of Silhouette imaging

Silhouette Extraction of a subject leads to a better edging and proper boundary detection. When the silhouette is extracted from a grayscale image, the boundaries extracted from the grayscale would be accurate to 20x30 pixel level. With such crisp boundaries, the feature will be seamless, more accurate features will be extracted. The Gaussian mixture model using silhouette extraction gives a better output as shown in the table below. Hence Silhouette extraction is significantly suitable for feature extraction.

6.5 Parameters used for silhouette extraction efficiency.

1. Mean:

$$\mu = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} P_{ij}$$

Where,

M= Number of columns in the matrix
N= Number of rows in the matrix
Pij=Pixel value or scalar component of the binary image.

2. Entropy:

$$-\sum_{1}^{N} (P_i .* \, log_2(P_i))$$

Where,

p= probability count value returned by the function imhist()
N= number of frames.

3. Standard deviation: $$\sigma = \sqrt{\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(P_{ij}-\mu\right)^2}$$

Where,
M= Number of columns in the matrix
N= Number of rows in the matrix
Pij=Scalar component or the pixel value of the binary image
μ = mean

4. Variance: σ^2

where,
σ = Standard deviation

5. Root mean square: $$x_{rms} = \sqrt{\frac{1}{MN}\sum_{i=1}^{M}\sum_{i=1}^{N}\left|P_{ij}\right|^2}$$

where,
M= Number of columns in the matrix
N= Number of rows in the matrix
Pij=Scalar component or the pixel value of the binary image

6. Skewness :

$$\theta = \frac{\sum_{i=1}^{M} \sum_{j=1}^{N} \left(P_{ij} - \mu\right)^3}{MN\sigma^3}$$

Where,
M= Number of columns in the matrix
N= Number of rows in the matrix
Pij=Scalar component or the pixel value of the binary image
μ = mean
σ = Standard deviation

7. Kurtosis :

$$\gamma = \frac{\sum_{i=1}^{M} \sum_{j=1}^{N} \left(P_{ij} - \mu\right)^4}{MN\sigma^4}$$

Where,

M= Number of columns in the matrix
N= Number of rows in the matrix
Pij=Scalar component or the pixel value of the binary image
μ = mean
σ = Standard deviation

Chapter 7: Feature Extraction

7.1 Information on Feature Extraction

Feature Extraction is one of the main stages of the project. The features extracted are used to identify the person in the finale stage of the project. The initial dataset will have many information which is not required for our recognition of a person using their walking style. Hence the feature extraction is a process of dimensionality reduction in the dataset to exclude all the non essential information, and extract the necessary information using various algorithms.

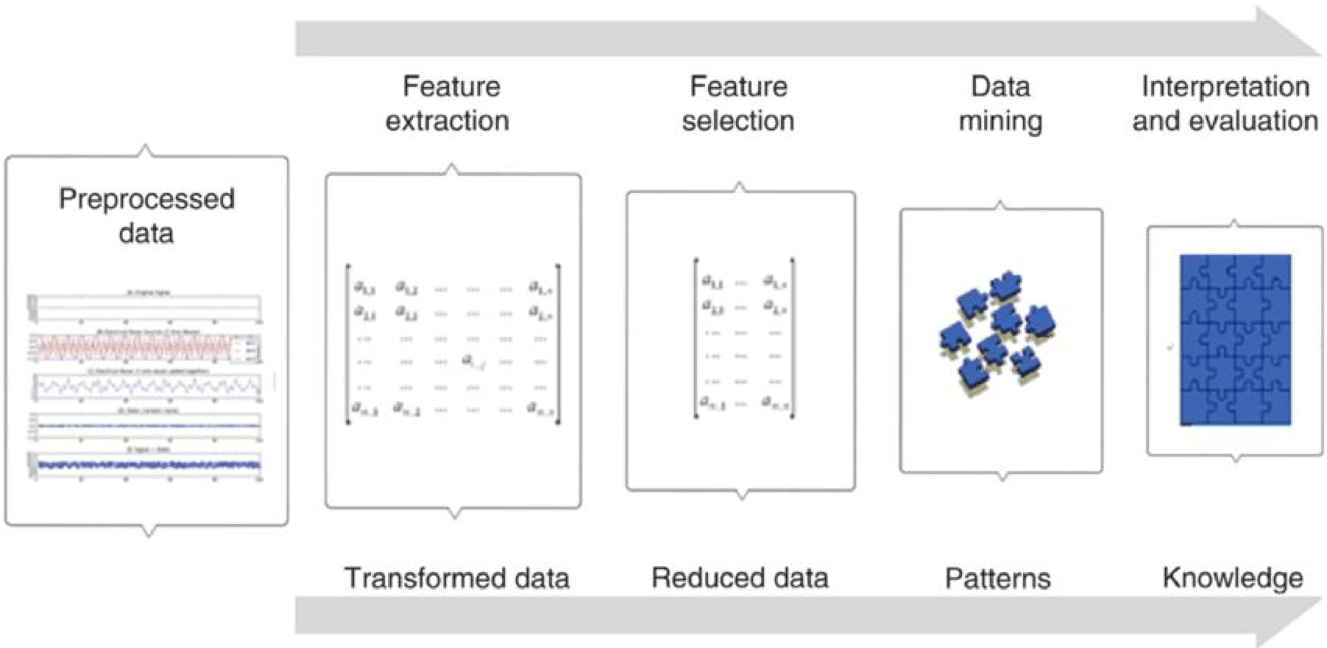


Figure-7.1 Image of neural network usage

7.2 Features considered for calculations

The Following Features are considered for the identification of a person.[2],[4]

1. Height of a person
2. Arm swing (body width)
3. Leg movement (leg width)
4. Upper body length (waist and above)
5. Lower body length (waist and below)

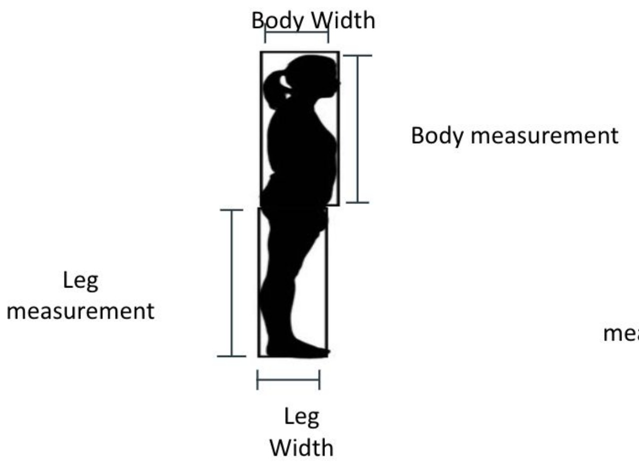


Figure-7.2 Feature consideration while standing

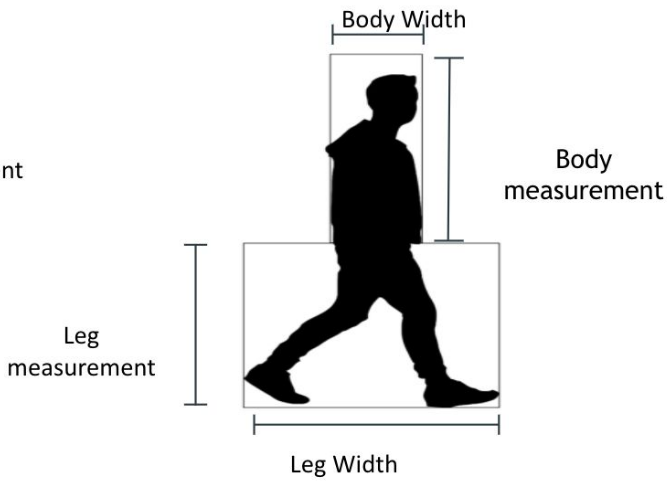


Figure-7.3 Feature consideration while walking

7.3 Method of extracting features

In the binary matrix, the top most pixel having the value 1 is taking the tip of the head, likewise the bottom most part of the matrix is taken as the leg tip. The leftmost and rightmost pixel having a value 1 will be taken for body edge measurements. The below image gives a better understanding of the feature extraction measurement in the binary image.

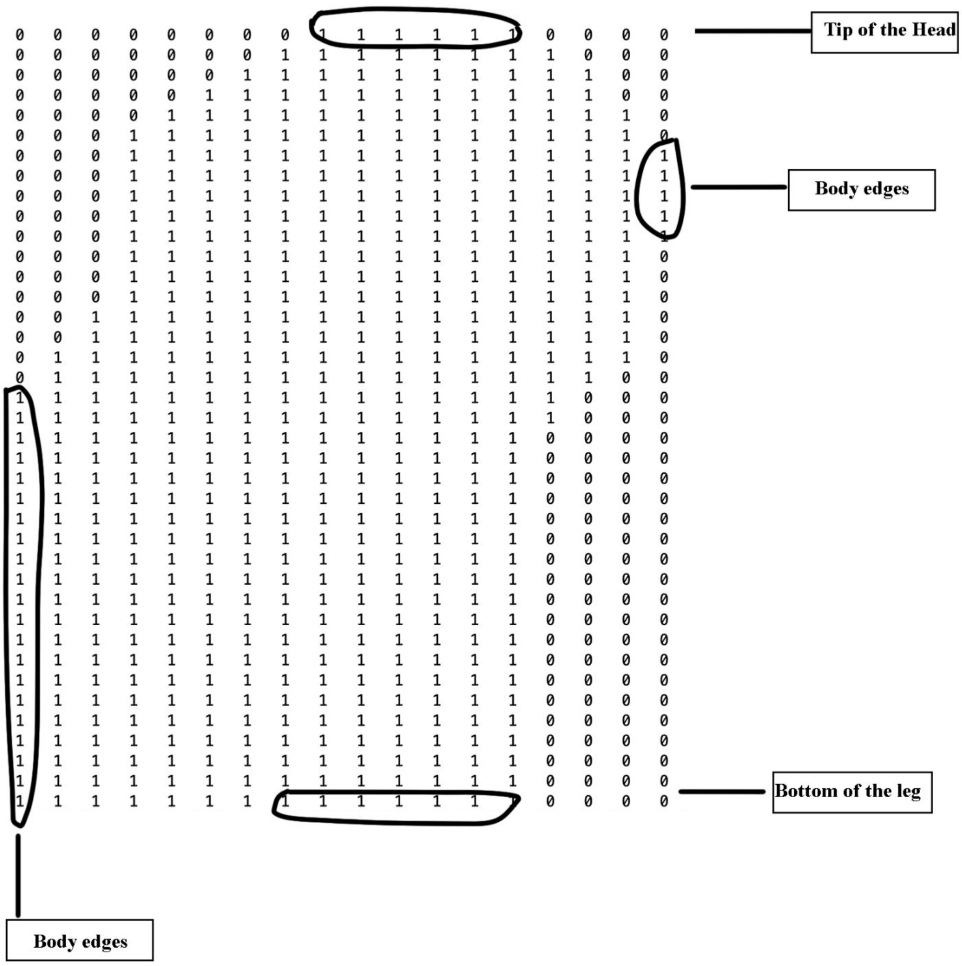


Figure-7.4 How the features are extracted from the matrix

The top most and bottom most pixel values having 1's indices are then subtracted to get the length and height. The upper body is taken to be 60% of the total height of a person. From the research conducted, the majority of people have their upper body 60% of their total height.

7.4 Advantages of using Feature extraction

The main advantage is that features extracted are not the same for different persons. For example, two persons even having the same height, will not have the same arm swing and leg movement at a time.

The features being used in this project cannot be easily replaced by someone with the same physique. Rest are as follows

1. Accuracy improves, as unwanted data is reduced.
2. Overfitting risk reduction.
3. As data is less, the speed of the algorithm to function increases.
4. Improved Data Visualization.
5. Increase in explainability of our model.

Chapter 8: Neural network

8.1 Introduction

Artificial neural networks (ANN) are computing systems which are modelled after the biological neurons that constitute the human brains. This system learns to perform specific tasks from examples (known as a learning set). It generally does this without programming with a set of rules. They do this without any prior knowledge. The neural network automatically generates identifying characteristics from the learning datasets that it processes. Every neuron can transmit a signal to another neuron via connections to each other. A neuron that receives a signal can process it and then signal additional neurons connected to it for further processing.

8.2 Neural network being used

In this project Feed forward backpropagation neural networks has been used. Feedforward neural network is a type of artificial neural network in which the connected nodes do not form a cycle and data flow happens in a forward direction. It is the simplest type of artificial network present. There are three layers present that are input nodes, hidden nodes and output nodes. Information travels from the input node through hidden nodes to the output node. There are no loops present. Backpropagation is a type of algorithm used in feed forward neural networks used for supervised learning(it maps input to output). Backpropagation computes the gradient of the loss function by using weights of an input output mapped network. It is a very efficient network which makes it feasible for training a multi layered neural network. It uses chain rule to compute the gradient of loss function with respect to each weight. It computes one layer at a time iterating backward from the last layer to avoid redundant calculations intermediate terms in the chain rule.

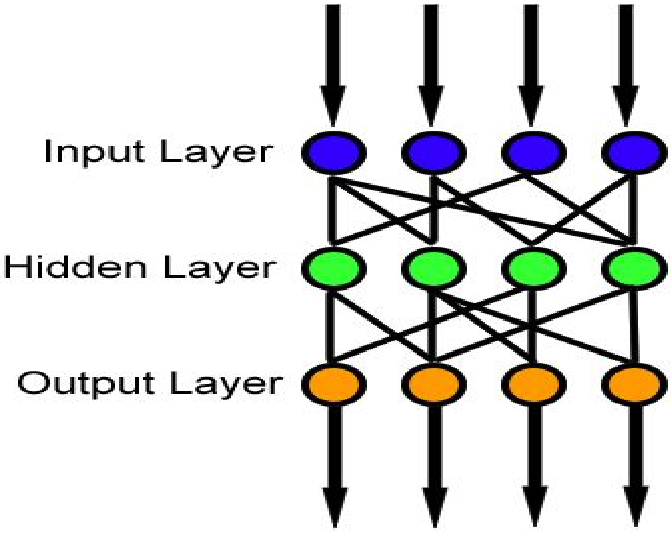


Figure-8.1   Layered Neural Network

8.3 Importance of neural network in our project

After all the features are extracted from the videos the mathematical output extracted after feature extraction is fed to the neural network for training the network. Training the network to get the best possible values so that when the biometric calculations are done, getting the best possible result. There are 14 hidden layers present in the neural network (this can be changed according to the performance of a graphics processing unit(GPU), more the hidden layer, better is the result).

Hidden layer is present between input layer and output layer. It consists of artificial neurons. These neurons take input and produce outputs using an activation function. When using a backpropagation method, all the hidden layers can be calibrated and fine tuned according to our needs.

14 hidden layers are used because when processing the data set each video has a different subject and calibrates the hidden layer according to the subject in our dataset, therefore 14 hidden layers.

After all the databases are trained dynamic input from the camera and the subject is recognized however he walks in the given condition.

8.4 Advantages of neural networks

- Neural networks have the ability to learn by themselves and produce output, inputs need not be provided all the time.
- No need of creating a seperate database to store inputs. It stores all the data in its network, so loss of data is minimized.
- Ability to learn by themselves and apply them when a similar event arises therefore making it useful for real-time events.
- Even if there is a piece of information missing they can produce the desired outputs.

8.5 Advantages of feedforward backpropagation

- It is the simplest type of neural network which is easy to use.
- Many numbers of hidden layers which can be fine tuned or calibrated according to our needs can be used.
- Output is efficient, therefore is feasible for multi layered training.

Chapter 9: Algorithms Developed

9.1 Frame difference Method

In Frame difference Method, each frame is subtracted with the previous frame. It is one of the easiest methods which can be used for silhouette extraction.In this method detection of object is done by sequence of frames from a static camera. Its objective is to detect the moving object from the difference of existing frame and the reference frame. It is the most commonly used method of object detection. It uses pixel based difference to detect the moving objects.

9.1.1 Implementation of frame difference method

In the beginning , the function “uigetfile('*.mp4;*.avi','Select Video')” is used to select the video from the dataset which was created. This selected video is then assigned to a variable to be used for calculations using the function VideoReader([path,file]).

For the background to be subtracted, the algorithm should be fed with an image of the background. Initially when the video has begun, the average value of the first 10 frames is taken as a background. This frame of averaged value, is then used as (n-1)th for frame subtraction. Then later the number of frames in the recorded video is found by using the function “obj.reader.NumberOfFrames”.

A loop is created to read all the frames (20:n-20 frames , where n - number of frames in the video taken). Each frame is extracted from the recorded video which has a 20:n-20 number of frames using the function “read(obj.reader,i” (where i varies from 20:n-20). n value is considered from 20 to n-20, to avoid the initial errors in the gaussian filter application on the frames. A morphological operation called “strel” is used as a structural element which is used for dilation or erosion. This function is used to create a disk shaped structural element as the video is in RGB format(which is a model in 3-dimension).

Gaussian filters are used to smoothen the object in the frames to extract an almost perfect silhouette. This filter is used for all the extracted frames. The function for smoothening frames is “imgaussfilt(frame1,1)”, this filter smoothens each frame with a Standard Deviation of 1. After the image is smoothed the frames extracted are subtracted to the previous frame, for example, if Frames extracted are frame 1 and frame 2, subtract frame 2 from frame 1 to obtain a subtracted frame. If the pixel value of the object in a subtracted frame is less than 30 then it is neglected. The function used is “diffrenceimge = (frame2-frame1)>30”.

Later resizing of frames to pixels is done with a 300x300 matrix. Resizing is done, as the applied Gaussian filter causes delay if used for higher magnitude image size, disturbing the silhouette extraction. Hence for the gaussian filter to be applied smoothly and without much delay it is resized to the standard value of 300x300.

After the subtraction of frames is done we resize the subtracted frame and create a mask with the function “mask = imresize(diffrenceimge,[300 300)”. To make the output frame more enhanced and to remove small and unwanted objects ``bwareaopen” function is used. It removes all the unwanted small objects with pixel value of p from the binary image bw to give bw2. Therefore this function is used to remove unwanted objects with pixel value less than 100 in subtracted and resized frames. Then dilation of the frame is done to get a perfect silhouette image of the desired object using the function “imdilate(mask,se)”. It dilates the surrounding pixels using strel operation and returned by the strel function. Final stage the object then obtained from all the frame subtraction is then converted into a silhouette for feature extraction. The mean, entropy, Standard Deviation and RMS values are calculated to verify our accuracy of the outputs with the Structural Silhouette and Feature Extraction (SSFE) method.

9.1.2 Advantages of SSFE method

As the frame difference subtraction was completed, the below parameters are used in checking the accuracy of the algorithm.

Mean of the Silhouette extracted was **0.0684.**

Standard Deviation extracted was **0.2524.**

Entropy Extracted was **0.3599.**

RMS was **0.1617**.

The standard Deviation and entropy Extracted was disruptive and discontinuous.

Figure-9.1 Silhouette extraction from frame Subtraction method with ideal background

Here is a snapshot of the extracted Silhouette from a completely non disturbed background.

This Silhouette would be difficult to calculate the features considered for the recognition of a person. Hence a new algorithm called Structural Silhouette and Feature Extraction (SSFE) is developed in this project..

9.2 SSFE Method

In Structural Silhouette and Feature Extraction (SSFE), extraction of the foreground (human) by using a matlab function of vision.ForegroundDetector. Then Gaussian filters for boundary clearance. The function needs to be specified as how many frames it should consider for training the function in differentiating the foreground and background. The minimum background ratio is a threshold to determine the background model, stated as a numerical scalar. The extracted foreground in RGB is then converted into grayscale. The human silhouette is extracted in terms of a binary image. Hence by this the feature extraction is done in the similar way as has been discussed in frame subtraction method. The feature extraction parameters remain the same.

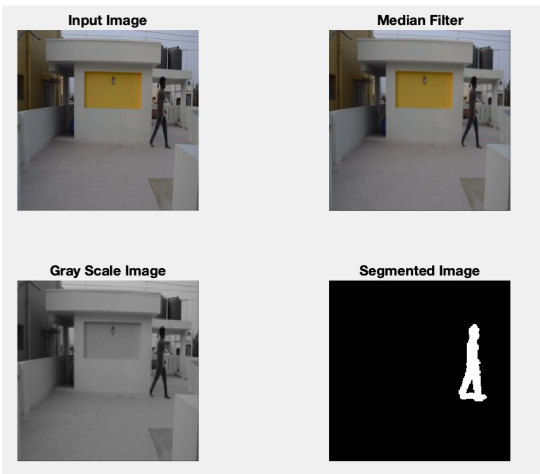


Figure-9.2 Silhouette extraction from SSFE method with busy background

The above image is one of the examples of the silhouette extraction process.

Due to better silhouette extraction in SSFE method, feature extraction is done by using the same algorithm.

9.3 Algorithm used in neural network

In the neural network the backpropagation algorithm is used in order to train the network. This algorithm trains the network based on the chain rule method. In this chain rule method, after each forward pass through a network, backpropagation performs a backward pass while adjusting the model's parameters.

1. **Input x:** Set the corresponding activation a^1 for the input layer.
2. **Feedforward:** For each $l = 2, 3, \ldots, L$ compute $z^l = w^l a^{l-1} + b^l$ and $a^l = \sigma(z^l)$.
3. **Output error δ^L:** Compute the vector $\delta^L = \nabla_a C \odot \sigma'(z^L)$.
4. **Backpropagate the error:** For each $l = L-1, L-2, \ldots, 2$ compute $\delta^l = ((w^{l+1})^T \delta^{l+1}) \odot \sigma'(z^l)$.
5. **Output:** The gradient of the cost function is given by $\frac{\partial C}{\partial w^l_{jk}} = a^{l-1}_k \delta^l_j$ and $\frac{\partial C}{\partial b^l_j} = \delta^l_j$.

The above algorithm shows a mathematical explanation of how our algorithm works. Since it computes the error vector backwards, it is called backward propagation. And 14 hidden layers are considered .

9.3.1 Advantages of backpropagation algorithm:

- Relatively simple implementation.
- Mathematical formulas in the algorithm can be used in any network for training.
- Computing time is less comparatively
- It does not require any prior knowledge about the inputs and is flexible

Chapter 10: Results and Discussions

In this report, presentation and discussion have been done of two methods of Gait Analysis which can be used for biometrics; one is Frame Difference Method and the other method is SSFE (Structural silhouette and Feature Extraction). Neural networks also have been adopted to avoid the minor changes in Gait. Both methods are based on silhouette extraction of a walking human.

The first method that is Frame subtraction, silhouette extracted by subtracting each frame was discussed and the results are displayed. Since the method of frame subtraction is a traditional method used in all the algorithms of video encoding, silhouette extraction was improper in this method.

In SSFE method, the silhouette extraction is better and more stable. Hence feature extraction like body width, leg width, leg height and body height (explained in a detailed way in feature extraction chapter) is more feasible in this method. The results of both the methods have been displayed in the below diagrams and mathematical values.

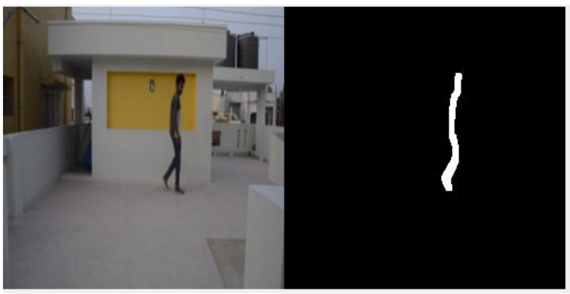


Figure-10.1 Frame Subtraction Method

Figure-10.2 SSFE Method

The Fig 10.1 and 10.1 are of the same video, 10.1 is from frame subtraction and 10.2 is from SSFE. It can be clearly seen from the image, segmentation of the silhouette is not enough to consider feature extraction to be done on the silhouette. Hence SSFE would be a better method to extract features of a human walking. The following table shows the comparison of frame difference method and SSFE method. The experimentation is done with a data set of 105 videos, each video of 5 seconds,with a frame rate of 25 frames per seconds.

Table 10.1 Difference Between Frame difference and SSFE

Methods	**Mean of the matrix**	**Standard deviation of the matrix**	**Entropy of the matrix**
Frame difference method	0.221	0.431	0.657
SSFE method	0.753	0.239	0.230

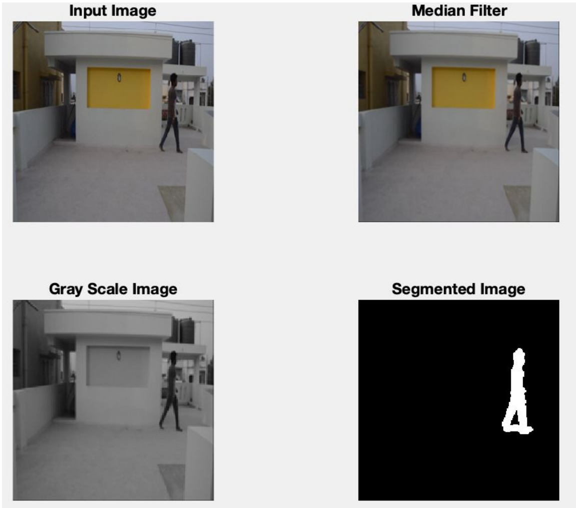
The values in the above table are obtained by performing operations such as Mean, Standard Deviation and Entropy on the binary matrix of the silhouette extracted.

In a binary image, the pixel has a value of either 0/1. Hence taking a mean for the matrix, it is known that the unwanted information is not being extracted, as the value is closer to 1 in SSFE method.

Standard Deviation and Entropy is lesser than the Frame Difference Method, Hence the disturbance in the matrix and deviation is less. Better silhouette extraction is being done.

A perfect Silhouette extracted video is taken as a standard video from our database and compares all values to this standardised video.

The same subject is taken in a different background and the silhouette is extracted. The values and results are recorded. The imagery results of the many have been displayed below.



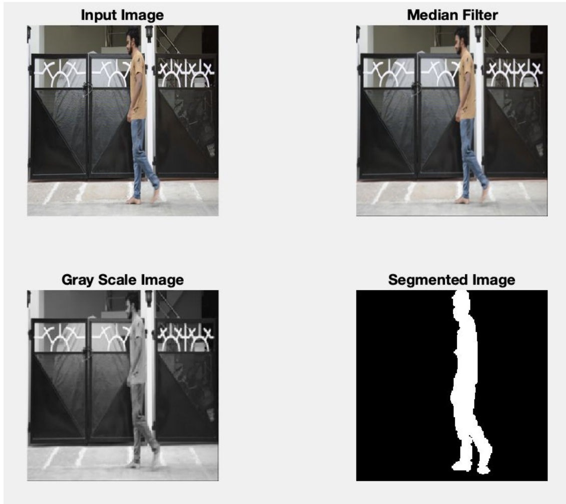
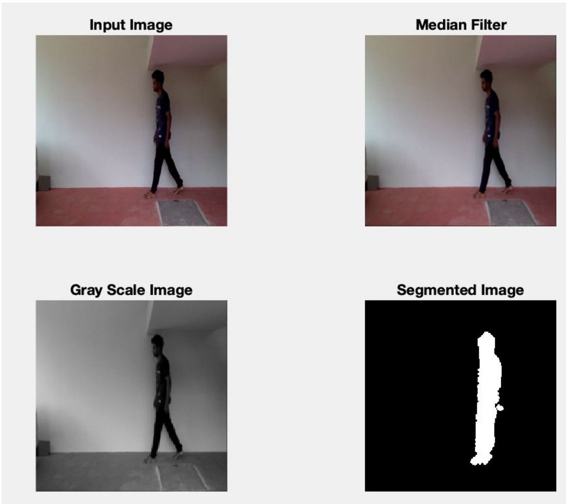


Figure-10.3 Subject_1, Background_1,
Standard Video for person 1

Figure-10.4 Subject_1, Background_2
Experimental Video 1 for person 1



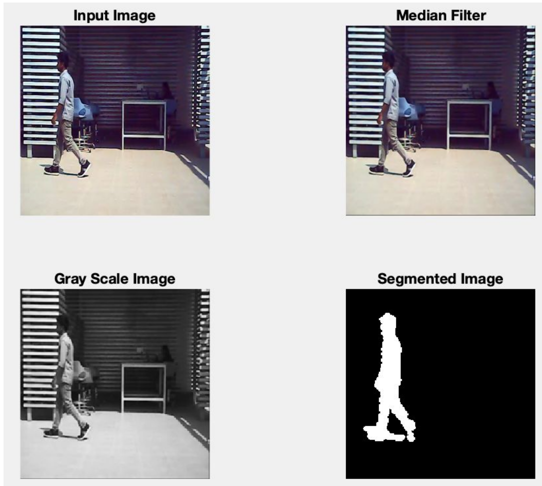


Figure-10.5 Subject_1, Background_3
Experimental Video 2 for person 1

Figure-10.6 Subject_1, Background_4
Experimental Video 4 for person 1

Table 10.2 The mathematical Results of the above Videos

Evaluation Parameters	**Standard Video**	**Experimental Video 1**	**Experimental Video 2**	**Experimental Video 3**
Mean	0.656	0.5032	0.7125	0.5482
Standard Deviation	0.1369	0.5005	0.4526	0.4977
Entropy	0.0034	1.000	0.8654	0.9933
RMS	0.6281	0.6914	0.8247	0.7118
Skewness	0.1232	0.1316	0.1569	0.1436

Graphical Representation of the features extracted from all the videos (body width) has been shown in the Figure 10.7

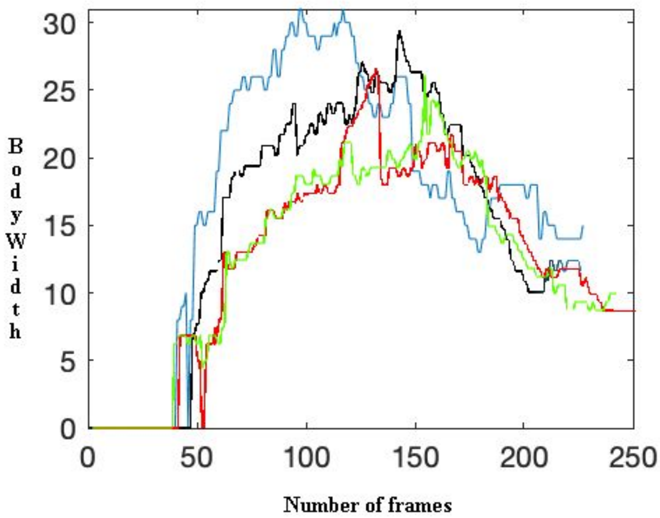


Figure 10.7 Graph of body width vs no of frames of subject_1

Blue - Standard Video
Black - Experimental Video _1
Green -Experimental Video _2
Red - Experimental Video _3

A different Subject, Subject_2 is taken in different backgrounds.

Figure 10.8 Subject_2, Background_5
Standard Video for person 2

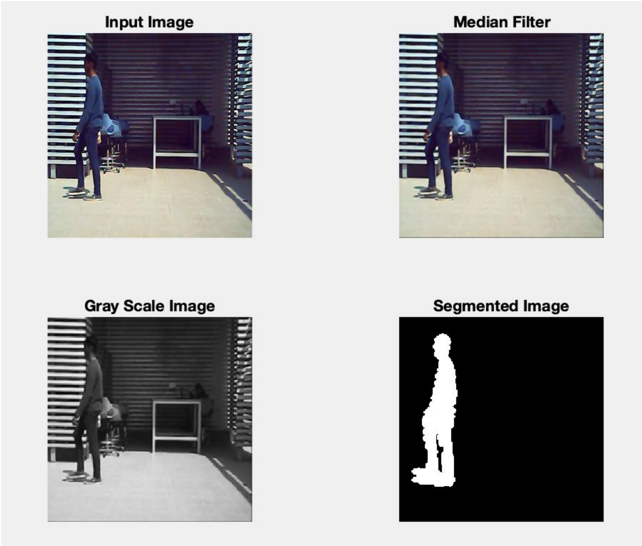


Figure 10.9 Subject_2, Background_6
Experimental Video_1 for person 2

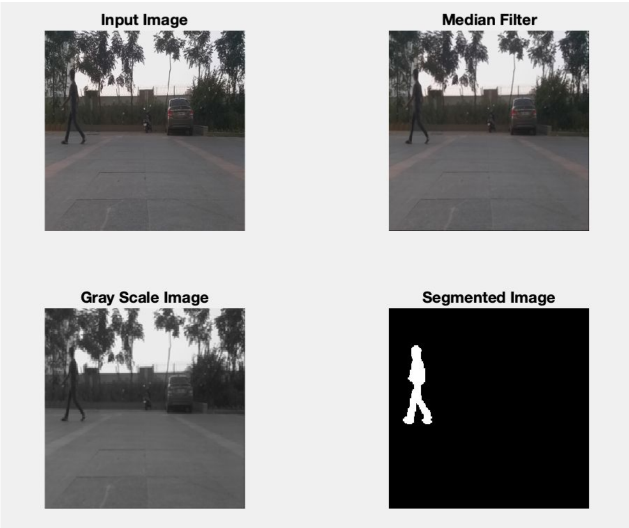


Figure 10.10 Subject_2, Background_7
Experimental Video_2 for person 2

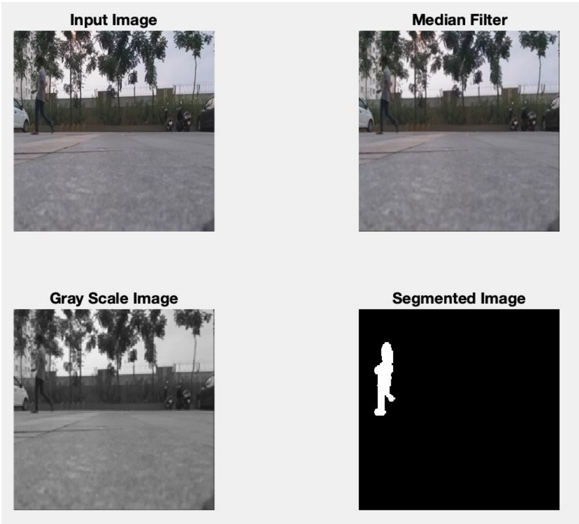
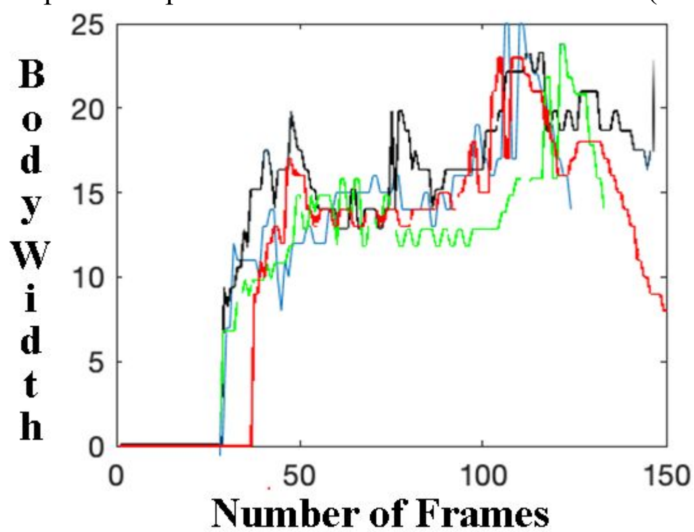


Figure 10.11 Subject_2, Background_8
Experimental Video_3 for person 2

Table 10.3   The mathematical Results of the above Videos

Evaluation Parameters	Standard Video	Experimental Video 1	Experimental Video 2	Experimental Video 3
Mean	0.7749	0.6569	0.7678	0.6630
Standard Deviation	0.4178	0.3987	0.4245	0.4216
Entropy	0.7694	0.6743	0.6534	0.5698
RMS	0.8587	0.8034	0.8234	0.7890
Skewness	0.1935	0.1859	0.1968	0.1897

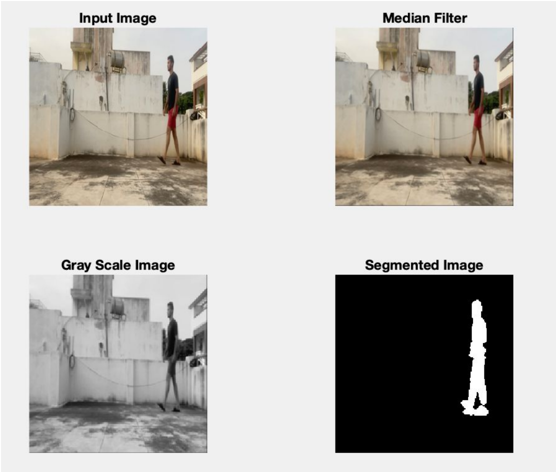
Graphical Representation of the features extracted (body width) has been done in the Figure 10.12



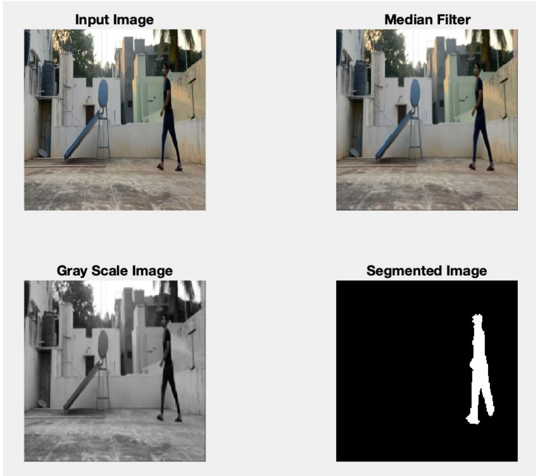
Blue - Standard Video
Black - Experimental Video _1
Green -Experimental Video _2
Red - Experimental Video _3

Figure 10.12   Graph of body width vs no of frames of subject_2

A different Subject, Subject_3 is taken in different backgrounds.



Subject_3, Background_9
Standard Video for person 3



Subject_4, Background_10
Experimental Video_1 for person 3

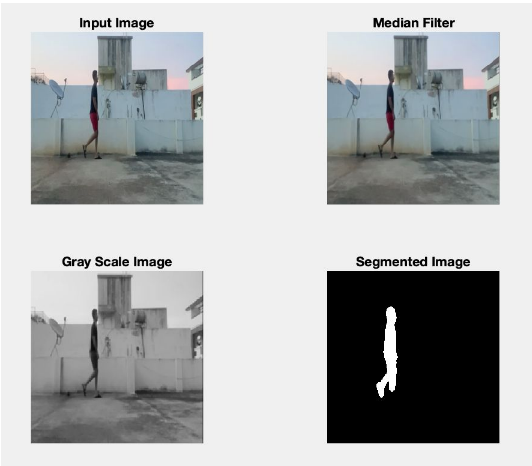


Figure 10.15 Subject_3, Background_10
Experimental Video_3 for person 3

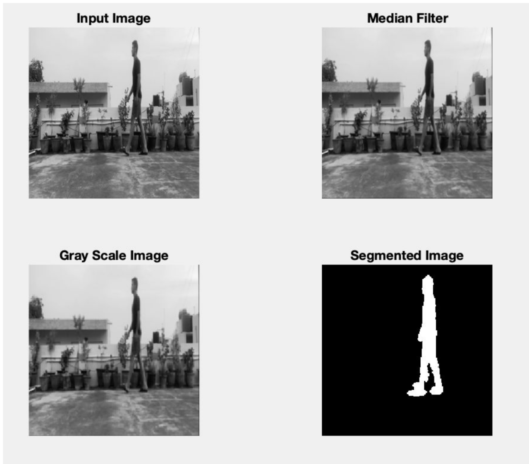
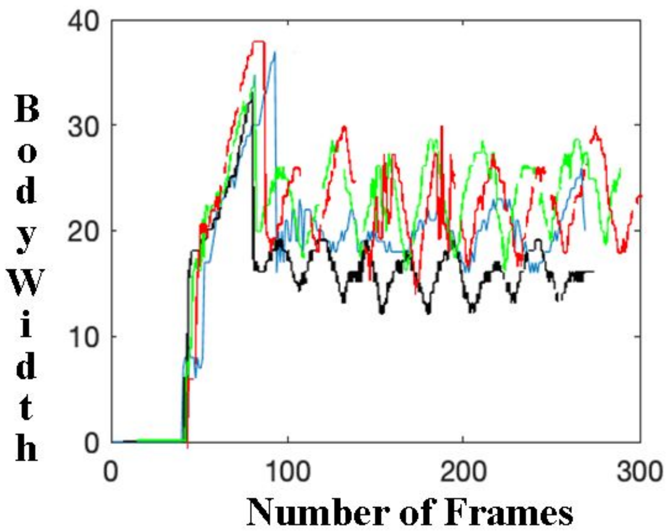


Figure 10.16 Subject_4, Background_11
Experimental Video_3 for person 3

Table 10.4 The mathematical Results of the above Videos

Evaluation Parameters	**Standard Video**	**Experimental Video 1**	**Experimental Video 2**	**Experimental Video 3**
Mean	0.4629	0.6158	0.5757	0.6006
Standard Deviation	0.4987	0.4865	0.4943	0.4898
Entropy	0.9960	0.9609	0.9834	0.9706
RMS	0.6211	0.7276	0.7083	0.7178
Skewness	0.0921	0.1245	0.1189	0.1052

Graphical Representation of the features extracted (body width) has been done in the Figure 10.17



Blue - Standard Video
Black - Experimental Video _1
Green -Experimental Video _2
Red - Experimental Video _3

Figure 10.17

Figure 10.17 Graph of body width vs no of frames of subject_3

The Graphical Representation is the Features Extracted, the individual values from the database are then fed to the Neural Networks. From these features, the weights are trained.

Once the training is completed the image 10.17 is displayed. Showing that the training is completed.

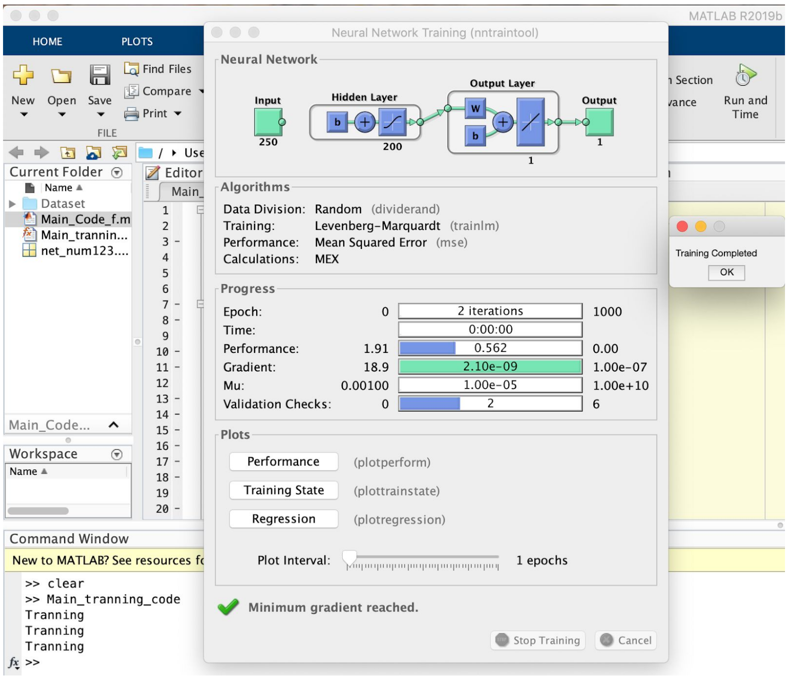


Figure-10.18 A screenshot of the integration of neural network

This imagery is the representation of the training of neural network for the database. 3 sets of people and 35 videos for each person is considered. Hence it trains according to all the person's database.

Final result obtained from the algorithm and neural network combined is

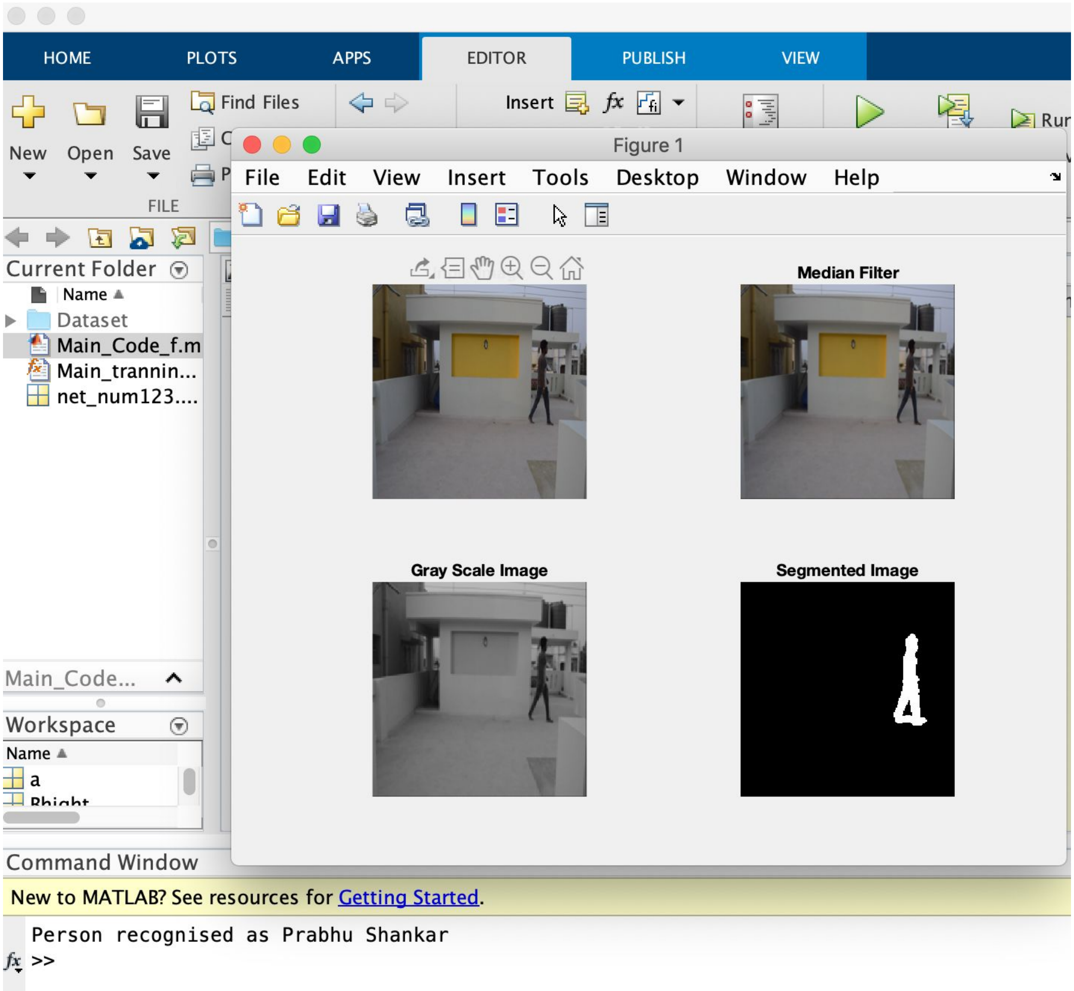


Figure-10.19 Image of One of the results of the data runs

The identification of a person using their Gait has been successfully completed.

Training of Neural Network:

Training Accuracy and Testing Accuracy of the Neural Networks have been given below.

Training of the Neural Network has been done by using 75 total videos, 25 of each person, videos each having 25 FPS.

A training Accuracy of **88.89%** was achieved and the same neural network has been used for testing.

Testing of Neural Network:

Testing of the Neural Network was done with different sets of videos.

It consists of 10 videos each of a person, a total of 30 videos in total 3750 frames.

So a total of 3750 extracted features.

A testing accuracy of **80.00%** was achieved and the results are displayed below.

Table 10.5 Accuracy in Frame difference and SSFE

Tested for a sample of 10 videos of each person. Total 30 videos	Frame Difference Method	SSFE Method
Accuracy in %	37.03 %	80.00 %

10.6 Accuracy in SSFE with and without Neural Network

Tested for a sample of 10 videos of each person. Total 30 videos	Without Neural network	With neural network
Accuracy in %	55.55 %	80.00 %

Chapter 11: Conclusions and Future developments

- A better algorithm could be developed in terms of feature extraction.
- This project is done using a static camera, in the further developments, a moving camera can be used for experimentation.
- This project is focused mainly on biometrics, hence the subject pertains to walk in a particular path. In the additional projects, the subject can be moved without any restrictions.
- Motion detection sensors can be used for a better silhouette extraction.
- Multiple cameras can be used for 360 degree view of the subject for better biometric security.

n References

1. Z Liu and S. Sarkar, "Improved gait recognition by gait dynamics normalization" i*n IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 6, pp. 863-876, June 2006.
2. S. Benbakhti and M. Benyettou, "Gait recognition based on leg motion and contour of silhouette," 2*012 International Conference on Information Technology and e-Services, Sousse, 2012,* pp. 1-5.
3. P. Nithyakalyani, A. Shanthini and G. Ponsam, "Human Gait Recognition using Deep Convolutional Neural Network," *2019 3rd International Conference on Computing and Communications Technologies (ICCCT)*, Chennai, India, 2019, pp. 208-211.
4. H. Su, Zhi-Wu Liao and G. Chen, "A gait recognition method using L1-PCA and LDA," *2009 International Conference on Machine Learning and Cybernetics*, Hebei, 2009, pp. 3198-3203.
5. K. Moustakas, D. Tzovaras and G. Stavropoulos, "Gait Recognition Using Geometric Features and Soft Biometrics," *in IEEE Signal Processing Letters,* vol. 17, no. 4, pp. 367-370, April 2010.
6. M. S. Nixon and J. N. Carter, "Advances in automatic gait recognition," *Sixth IEEE International Conference on Automatic Face and Gesture Recognition, 2004. Proceedings.*, Seoul, South Korea, 2004, pp. 139-144.
7. W. Kusakunniran, Q. Wu, H. Li and J. Zhang, "Automatic Gait Recognition Using Weighted Binary Pattern on Video," *2009 Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance*, Genova, 2009, pp. 49-54.

8. Z Liu and S. Sarkar, "Effect of silhouette quality on hard problems in gait recognition," in *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 35, no. 2, pp. 170-183, April 2005
9. L. Wang, T. Tan, H. Ning and W. Hu, "Silhouette analysis-based gait recognition for human identification," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 12, pp. 1505-1518, Dec. 2003
10. A. S. M. H. Bari and M. L. Gavrilova, "Artificial Neural Network Based Gait Recognition Using Kinect Sensor," in *IEEE Access*, vol. 7, pp. 162708-162722, 2019.

Word Count: 7026

Plagiarism Percentage 10%

Matches

1 **World Wide Web Match**
View Link

2 **World Wide Web Match**
View Link

3 **World Wide Web Match**
View Link

4 **World Wide Web Match**
View Link

5 **World Wide Web Match**
View Link

6 **World Wide Web Match**
View Link

7 **World Wide Web Match**
View Link

8 **World Wide Web Match**
View Link

9 **World Wide Web Match**
View Link

10 **World Wide Web Match**
View Link

11 **World Wide Web Match**

View Link

12 **World Wide Web Match**
View Link

13 **World Wide Web Match**
View Link

14 **World Wide Web Match**
View Link

15 **World Wide Web Match**
View Link

16 **World Wide Web Match**
View Link

17 **World Wide Web Match**
View Link

18 **World Wide Web Match**
View Link

19 **World Wide Web Match**
View Link

20 **World Wide Web Match**
View Link

21 **World Wide Web Match**
View Link

22 **World Wide Web Match**
View Link

23 **World Wide Web Match**
View Link

24 **World Wide Web Match**
View Link

Suspected Content

Chapter 1: Introduction Biometric technology analyzes human characteristics to verify a person using the person's fingerprints, eyes , face, gait, heart beat , and voice recordings. Gait is defined as locomotion of human beings, more importantly, walking style of human beings. In the process of walking, a foot from the front, reaching the back and coming back to the front is called one Gait Cycle. This profile of a person is unique and differentiable from person to person. Hence it can be used for biometrics. It can be used for various purposes such as chiropractic and osteopathic, but in this report, it is mainly focused on biometrics usage of gait. Gait in Biometrics can be carried out at farther distances or at low resolutions.The main parameters in any gait analysis is dynamic base, step length, cadence, stride length, speed. Since the gait here is adopted for biometrics, only novel features: body width, leg width(step length), leg length and body length are required. This study emphasizes the quantification of these novel features and used for biometrics. 0 1.1 Problem statement In today's world the perpetrators are trying to evade the current biometrics using various gadgets and are getting away from their criminal activities. It is very important to control such activities. So implementing a gait recognition model which would help in recognizing the person with the help of the gait patterns along with the neural network for training the patterns. The project has been done on the following parameters: ● To Intervene Biometrics using Gait Analysis ● Database Creation for Gait analysis ● Smoothening of Boundaries in each frame using Gaussian and Median Filters ● Conversion of Frames into Grayscale images ● Extracting Binary Silhouette from Grayscale images ● To Use Features which cannot be faked and lead to malicious way ● Training the Neural network to get optimised output ● To use Neural Network for minute Variations in Gait analysis ● Identification of a person 1.2 Motivation Security and Privacy go hand in hand in this current digital age, we have to make sure that an individual's privacy is not interceded. Biometrics is a way that a person's identity cannot be copied by anyone on this planet and it can be used for many security and privacy purposes. Biometrics are Fingerprint scanners, Iris Scanners, Retinal Scanners, Facial Recognition, Voice Recognition. These Biometrics can be faked and used for malicious purposes. Gait as mentioned, is a walking style of a person, this cannot be faked and used in an harmful way. Gait may change over a course of 5-6 years, but not in the span of 12-24 months (Traumatic Experiences are not taken into consideration). Since this Biometrics has lesser limitations than all the other forms of biometrics, we have considered to work on this particular project. Apart from applying gait studying for security purposes such as crime avoidance, it can be used for medical motives such as chiropractic and osteopathic. Moreover, gait biometrics can be used for public surveillance as well. In places such as airports and shopping malls. Chapter 2: Literature Survey This paper mainly proposes about the different approaches to the gait recognition. The three gait approaches mentioned in this paper are 1) Temporal alignment 2) Static parameter 3) Silhouette approach. Consideration of a Silhouette based approach has been done in our project. This approach considers the normalisation technique and linear discriminant analysis in order to maximise the differentiation between persons while reducing the variation of the same subject, thus helping in better classification. The factors like walking surface, walking speed and shadow affect the gait recognition. This method partially helps in eliminating certain factors like walking speed by normalisation. Here normalisation mainly helps in emphasising the shape of the silhouette. Thus upon emphasising the shape feature extraction becomes more easier in the further development of the project. This paper presented an algorithm which recognizes the gait of the person. This algorithm recognizes people by their style of walking. In this algorithm initially, the video is converted into a frame. In each frame the background is subtracted to segment out the person from the background. This process extracts the moving silhouettes of the subject. This is one of the methods tried in our project(frame difference method). This silhouette is categorised by computing the parameters such as contour of silhouettes and the area of the silhouettes. This paper proposes a combination of appearance based approach and model based approach to extract proper feature parameters. Feature parameters such as height, stride length, cadence and width can be derived by dividing the silhouette into different body parts. From this the features to be extracted in the project and method to extract them has been considered. Artificial Neural Network has an exceptional quality of deriving complicated, raw data into meaningful results. In this paper, consideration of DCNN (deep convolutional Neural network) has been done. In this paper, GEI (gait energy image) are recorded and fed to a neural network. According to the image matching and frame matching. The output is obtained and results are published. Since GEI extraction and feature vectors cannot be copied from another individual. Main reason to consider this paper is to add deep learning by backward propagation for feature extraction. Doing so, the minimal abnormalities in gait would be curved into the output, giving a optimal output This paper proposes gait analysis method using PCA (principal component analysis) base L1 normalisation and LDA (linear discriminant analysis) method. L1-PCA represents the features and LDA to analyze and classify these features. The importance of the width feature is stressed in this paper. L1-PCA is used for representation of data, dimension reduction and analysis. LDA helps in maximising the discrimination between features of different persons. This was taken into consideration during the early stages of our project. This paper proposes the use of different features extracted for biometrics. Particularly the features include stride length and height. The use of these features is used to increase the performance of the output obtained by feature extraction for better biometric recognition system. This paper explains about how posterior features can be used for biometrics recognition. Initially gait recognition process is done to extract silhouette image for feature extraction. Then by using a probabilistic framework, gait score obtained by gait recognition and values of features extracted are combined to generate a final score. Chapter 3: System Model 3.1 Video segmentation Video segmentation is a processing method in which video is divided into multiple frames upon which image segmentation happens. 3.2 Image Segmentation Segmentation of the image is the basic step to analyze images and extract required data from them. It is the process of division of the digital image into many segments. Segmentation is the process

> to simplify the representation of the image into something more meaningful. **20**

It consists of converting an image into a collection of pixels. That is represented by a mask or a labeled image. By converting an image into segments, only the required segments of the image can be processed instead of processing the whole image. 3.3 Testing of Different methods There are different segmentation techniques available. The two methods tried in the project are : 1. Frame difference method [1],[2] 2. Gaussian and median filter method 3.3.1 Frame difference method This method is

> the simplest form of background subtraction. This method can also be referred to as temporal difference. Uses the frame at time t-1 as the background model for the frame at time t . This technique is sensitive to noise and variations in illumination, and does not consider local consistency properties of the change mask. **6**

[1],[2] 3.3.2 Gaussian and median filter method

> A Gaussian filter being a linear filter **11**

is used to blur the image and reduce unwanted noise. The

Gaussian filter will blur edges and reduce contrast 11

in each frame. Whereas

a median filter known as a non-linear filter is used 11

in reducing salt and pepper noise. The noise caused by sudden and sharp disturbances in the image signal is the salt and pepper noise. This method has an enhanced edge detection and increased segmentation performance compared to frame difference method. Chapter 3.4 Matlab 3

.4.1 Introduction to Matlab Matlab is a high-performance language for technical computing. It integrates visualization, computation and programming in an easy-to-use environment. The problems and answers are expressed in similar mathematical 8

expressions. It consists of

a family of application-specific solutions called toolboxes. They allow to learn and apply specialised technology. Toolboxes are comprehensive collections of MATLAB functions (M-files) that extend the MATLAB environment to solve particular classes of problems. Studies for which toolboxes are available include wavelets ,signal processing, neural networks ,control systems, fuzzy logic, simulation and many others. 5

3.4.2 Tools used in Matlab 1. Computer Vision Toolbox It involves algorithms, functions and applications for testing and designing computer vision. 3D vision, and video processing systems. Object detection and tracking, as well as feature detection, extraction, and matching can be performed. 2.

Database Toolbox It provides functions and an app for exchanging data with relational and non-relational databases. It enables this exchange by automatically converting between database and Matlab data types. 9

3.

Image Processing Toolbox It provides a comprehensive set of reference-standard algorithms and workflow applications for image processing, analysis, visualization, and algorithm development. Image segmentation, image enhancement, noise reduction, geometric transformations, image registration, and 3D image processing can be performed. 7

It allows to

automate common image processing workflows. It helps to interactively segment image data, compare image registration techniques, and batch-process large data sets. 3

4. Deep Learning Toolbox It presents a framework for designing and enforcing deep neural networks with algorithms, pre-skilled fashions. With the help of this toolbox class and regression on photograph, time-collection, and textual content facts may be accomplished.With this, the community can be designed, analyzed, and skilled graphically. It helps to control multiple deep learning experiments, maintain the required training parameters, examine results, and examine code from exceptional experiments. Layer output can be visualized. 5.

Statistics and Machine Learning Toolbox It provides features to explain, analyze 3

the facts. The toolbox gives supervised and unsupervised devices gaining knowledge of algorithms, which includes aid vector machines,

k-nearest neighbour, k- manner, hierarchical clustering, Gaussian aggregate models, and hidden Markov Models. Many of the facts and system learning algorithms can be used for computations on records units that are too huge to be saved in memory. 3

6.

Vision HDL Toolbox It provides pixel-streaming algorithms for the layout and implementation of 2

imaginative and prescient structures

on FPGAs and ASICs. It provides a layout framework that supports a diverse set of interface sorts, body sizes, and frame rates. 2

The photo

processing, video, and computer vision algorithms inside the toolbox use an architecture appropriate for HDL implementations. The 2 generated HDL code is FPGA-confirmed for frame sizes as much as 8k decision and for high frame video .

3.5 Neural Network A neural network is **23**

a circuit consisting of artificial neurons which can also be called as nodes (referred to in a neural network). An artificial neuron is a mathematical function used as a version of neurons. The neurons are modelled as weights. All inputs are modified for the use of weights and summed. This network can examine from the facts, so it is able to study to realise patterns and classify facts and expect destiny occasions. Neural network breaks the input into layers of abstraction. Then it's miles educated on the usage of exclusive examples to recognize photos, speech and lots of different matters. The weights are adjusted at the same time as schooling in step with a specific rule with the intention to acquire the required outcomes.[3] 3.5.1

Feed Forward Back Propagation In the neural network a feed forward back propagation **18**

algorithm is used to train. Here the

information moves in a single direction, from the input nodes, through the hidden nodes after which to the output nodes. **15**

This does not have loops or cycles in the network. In this feed forward neural network , back propagation algorithm helps in efficiently adjusting the weights in order to get the required result.This performance

makes it possible to use gradient techniques for using multilayer networks, updating weights to minimise loss. **4**

This is how feed forward back propagation works. 3.6 Block diagram of gist of working 3.6.1 Structural silhouette and feature extraction (SSFE) Figure-3.1 Block Diagram of the working of the project An analog video is recorded using a 25 megapixels camera and is converted into digital format. The standardized video format is about 25 frames per second. Each frame is captured by the camera and played continuously to give video. All the recorded videos are saved in a database, which will be used to examine the video. Dynamic input from the camera is taken and recorded. This video is examined with the pre-recorded database created earlier. Code written, converts the video into desirable output by smoothing the frames, converting it into grayscale image, cropping the required binary silhouette for feature extraction. Silhouette is a black and white image of a person with the boundary either filled with black or white.From the above process different features are extracted which will be used to train the neural network. All the features from the silhouette frames are compared with the trained database to provide a result. This result is then checked for maximum similarity. If there is maximum similarity the person will be recognised. 3.6.2 Neural networks Gait Cycle Detection Dataset Pre-Processing Artificial Neural Network Feature Extraction Training of ANN Classification Trained ANN Recognition Figure-3.2 Block Diagram of the working of the Neural Network After Gait Detection has been done, smoothening of the image is done, the extracted features are then used for training the neural network. From the dataset created, the neural network alters the weights and trains the nodes in order to obtain the output. According to the trained weights and nodes, the neural network takes the features from the ongoing video and gives the classification of a person. Concluding with the recognition, leading to biometrics. Chapter 4: Database creation 4.1 Introduction, what is a database? Database is the basic information required in a project to get a more efficient result and also a base for training any neural network. The word itself has the meaning in it, 'Data'base, meaning it has some data which can be used for calculations. [1] Database is an organised collection of data which is accessible by users for calculations and proper training of the neural network. Database in the digital world is a crucial part of networking, without databases; it would fail to keep a check on abnormal activities, program computers in such a way they are being used now or create Artificial Intelligence. 4.2 Database created Databases that have been used in the project are our own videos recorded from a 25-megapixel camera having the recording speed of 25 frames per second. The main reason for considering this database is because of the easy access. Initial stage where the algorithm had to be developed and a neural network for the proper training of the weights, this was carried out on the basis of trial and error method until properly segmented video was obtained for better feature extraction. The accurately segmented video is considered as the standard video for all the future references. The videos considered in the database have variations in the gait patterns with respect to the subject i.e.speed, arm-swing, step width. These variations were considered because of the changes in surroundings of the subject at a particular time. These variations are not supposed to intervene in the calculations of the results. Hence a wide range of variations are considered and almost 25 videos of each person has been used to train the Neural network. These videos cover almost all types of gait, including Slow steps, Fast steps, Broad steps, Tapered steps and so on. For the better weighted average in the Neural Network, more the variations should be considered in the database. It is always known, "The more the merrier"; the same way more videos are pre recorded and saved in the database, the more accurate will be the output. Each video should be a minimum of 5 seconds and max of 15 seconds i.e minimum of 125 - 375 frames. This is required for the least gait analysis to be performed by the algorithm. Recording lesser than the threshold would be insufficient for the algorithm to perform the segmentation for each frame and extract accurate features. Anything above 375 frames (15 seconds) is just wasting time and data by recording. Hence the ideal database video for the project is between 125 - 375 frames. Figure-4.1 A screenshot of the database created Chapter 5: Human detection 5.1 Introduction Human detection is a type of object detection. Since the object considered in the project is a person, we are referring it to as human detection. Human detection(object detection) is a computer vision technique used for identification of the person's movement from the image or video. 5.2 Vision.Foregrounddetector In the project vision.foregrounddetector function is used to detect the person. This function enables in computing and returning the foreground masks with the use of the gaussian mixture model. This function compares a color or grayscale video frame to a background to determine whether the subject's

pixels are part of the background or the foreground. It

then finds the value of the foreground mask. These can be in real time from a video camera or from photographs. The detected images are enhanced using gaussian filter and median filter for the removal of noise from them. Later the RGB image is transformed into a grayscale image which is the best way to extract the edges and other features from it. As a result of this silhouette image is obtained. Figure-5.1 An example of how the median filter works Figure-5.2 An example of how the Gaussian Filter works Chapter 6: Silhouette Extraction 6.1 Silhouette Meaning Explained A silhouette is a type of image of a person, represented as a shape of a single colour, with its edge of the shape matching the outline of the subject.

Usually the solid color will be black or white, with white or black white background simultaneously. 6.2 Usage of Silhouette When the extraction of the frames from the video was done, each frame will have an image of a person with a different stance of walking. These stances are extracted from the frames and converted into silhouette images. Figure-6.1 Gait silhouette From these Silhouettes the features are then extracted and used for person detection. 6.3 Silhouette storage. Each silhouette is stored as a binary image in a matrix form. This is done by converting the RGB to grayscale, and then to binary image. A sample of a matrix is given below. Figure-6.2 An example of the matrix obtained after the extraction of silhouette The binary image is then used for further calculation of feature vectors to be given for the neural network. 6.4 Advantages of Silhouette imaging Silhouette Extraction of a subject leads to a better edging and proper boundary detection. When the silhouette is extracted from a grayscale image, the boundaries extracted from the grayscale would be accurate to 20x30 pixel level. With such crisp boundaries, the feature will be seamless, more accurate features will be extracted. The Gaussian mixture model using silhouette extraction gives a better output as shown in the table below. Hence Silhouette extraction is significantly suitable for feature extraction. Chapter 7: Feature Extraction 7.1 Information on Feature Extraction Feature Extraction is one of the main stages of the project. The features extracted are used to identify the person in the finale stage of the project. The initial dataset will have many information which is not required for our recognition of a person using their walking style. Hence the

feature extraction is a process of dimensionality reduction in the dataset to 21

exclude all the non essential information, and extract the necessary information using various algorithms. Figure-7.1 Image of neural network usage 7.2 Features considered for calculations The Following Features are considered for the identification of a person.[2],[4] 1. Height of a person 2. Arm swing (body width) 3. Leg movement (leg width) 4. Upper body length (waist and above) 5. Lower body length (waist and below) Figure-7.2 Feature consideration while standing Figure-7.3 Feature consideration while walking 7.3 Method of extracting features In the binary matrix, the top most pixel having the value 1 is taking the tip of the head, likewise the bottom most part of the matrix is taken as the leg tip. The leftmost and rightmost pixel having a value 1 will be taken for body edge measurements. The below image gives a better understanding of the feature extraction measurement in the binary image. Figure-7.4 How the features are extracted from the matrix The top most and bottom most pixel values having 1's indices are then subtracted to get the length and height. The upper body is taken to be 60% of the total height of a person. From the research conducted, the majority of people have their upper body 60% of their total height. 7.4 Advantages of using Feature extraction The main advantage is that features extracted are not the same for different persons. For example, two persons even having the same height, will not have the same arm swing and leg movement at a time. The features being used in this project cannot be easily replaced by someone with the same physique. Rest are as follows 1. Accuracy improves, as unwanted data is reduced. 2. Overfitting risk reduction. 3. As data is less, the speed of the algorithm to function increases. 4. Improved Data Visualization. 5. Increase in explainability of our model. Chapter 8: Neural network 8.1 Introduction Artificial neural networks (ANN) are computing systems which are modelled after the biological neurons that constitute the human brains. This system learns to perform specific tasks from examples (known as a learning set). It generally does this without programming with a set of rules. They do this without any prior knowledge. The neural network automatically generates identifying characteristics from the learning datasets that it processes. Every neuron can transmit a signal to another neuron via connections to each other. A neuron that receives a signal can process it and then signal additional neurons connected to it for further processing. 8.2 Neural network being used In this project we use Feed forward backpropagation neural networks. Feedforward neural network is a type of artificial neural network 17

in which the connected

nodes do not form a cycle 2

and data flow happens in a forward direction. It is the simplest type of artificial network present. There are three layers present that are

input nodes, hidden nodes and output nodes. 24

Information travels

from the input node through hidden nodes to the output node. There are no loops 16

present. Backpropagation is a type of algorithm used in feed forward neural networks used for supervised learning(it maps input to output).

Backpropagation computes the gradient of the loss function 4

by using weights of an input output mapped network. It is a very efficient network which makes it feasible for training a multi layered neural network. It uses chain rule to compute the

gradient of loss function with respect to each weight. 4

It computes

one layer at a time iterating backward from the last layer to avoid redundant calculations intermediate terms in the chain rule. 4

Figure-8.1 Layered Neural Network 8.3 Importance of neural network in our project After all the features are extracted from the videos the mathematical output extracted after feature extraction is fed

to the neural network for training the network. We train the network to 14

get the best possible values so that when we do the biometric calculations we get the best possible result. There are 200 hidden layers present in the neural network (this can be changed according to the performance of a graphics processing unit(GPU), more the hidden layer, better is the result).

Hidden layer is present between input layer and output layer. 14

It consists of artificial neurons. These neurons take input and produce outputs using an activation function. When we use a backpropagation method, all the hidden layers can be calibrated and fine tuned according to our needs. We are using 200 hidden layers because when we process our data set each video has a different subject and we have to calibrate the hidden layer according to the subject in our dataset, therefore we use 200 hidden layers. After all the databases are trained we give a dynamic input from the camera and the subject is recognized however he walks in the given condition. 8.4

Advantages of neural networks ● Neural networks have the ability to learn by themselves and produce output, inputs need not be provided all the 10

time. ● No need of creating a separate database to store inputs. It stores all the data in its network, so loss of data is minimized. ● Ability to learn by themselves

and apply them when a similar event arises therefore making 10

it useful for real-time events. ● Even if there is a piece of information missing they can produce the desired outputs. 8.5 Advantages of feedforward backpropagation ● It is the simplest type of neural network which is easy to use. ● Many numbers of hidden layers which can be fine tuned or calibrated according to our needs can be used. ● Output is efficient, therefore is feasible for multi layered training. Chapter 9: Algorithms Developed 9

.1 Frame difference Method In Frame difference Method, each frame is subtracted with the 22

previous frame. It is one of the easiest methods which can be used for silhouette extraction. In this method detection of object is done by sequence of frames from a static camera. Its objective

is to detect the moving object from the difference of existing frame and the reference frame. It is the most commonly used method 12

of object detection. It uses

pixel based difference to detect the moving 12

objects. 9.1.1 Implementation of frame difference method In the beginning , the function “uigetfile('*.mp4;*.avi','Select Video')” is used to select the video from the dataset which was created. This selected video is then assigned to a variable to be used for calculations using the function VideoReader([path,file]). For the background to be subtracted, the algorithm should be fed with an image of the background. Initially when the video has begun, the average value of the first 10 frames is taken as a background. This frame of averaged value, is then used as (n-1)th for frame subtraction. Then later the number of frames in the recorded video is found by using the function “obj.reader.NumberOfFrames”. A loop is created to read all the frames (20:n-20 frames , where n - number of frames in the video taken). Each frame is extracted from the recorded video which has a 20:n-20 number of frames using the function “read(obj.reader,i” (where i varies from 20:n-20). n value is considered from 20 to n-20, to avoid the initial errors in the gaussian filter application on the frames. A morphological operation called “strel” is used as a structural element which is used for dilation or erosion. This function is used to create a disk shaped structural element as the video is in RGB format(which is a model in 3-dimension). Gaussian filters are used to smoothen the object in the frames to extract an almost perfect silhouette. This filter is used for all the extracted frames. The function for smoothening frames is “imgaussfilt(frame1,1)”, this filter smoothens each frame with a Standard Deviation of 1. After the image is smoothed the frames extracted are subtracted to the previous frame, for example, if Frames extracted are frame 1 and frame 2, subtract frame 2 from frame 1 to obtain a subtracted frame. If the pixel value of the object in a subtracted frame is less than 30 then it is neglected. The function used is “diffrenceimge = (frame2-frame1)>30”. Later resizing of frames to pixels is done with a 256x256 matrix. Resizing is done, as the applied Gaussian filter causes delay if used for higher magnitude image size, disturbing the silhouette extraction. Hence for the gaussian filter to be applied smoothly and without much delay it is resized to the standard value of 256x256. After the subtraction of frames is done we resize the subtracted frame and create a mask with the function “mask = imresize(diffrenceimge,[300 300)”. To make the output frame more enhanced and to remove small and unwanted objects “bwareaopen” function is used. It removes all the unwanted small objects with pixel value of p from the binary image bw to give bw2. Therefore this function is used to remove unwanted objects with pixel value less than 100 in subtracted and resized frames. Then dilation of the frame is done to get a perfect silhouette image of the desired object using the function “imdilate(mask,se)”. It dilates the surrounding pixels using strel operation and returned by the strel function. Final stage the object then obtained from all the frame subtraction is then converted into a silhouette for feature extraction. The mean, entropy, Standard Deviation and RMS values are calculated to verify our accuracy of the outputs with the Structural Silhouette and Feature Extraction (SSFE) method. 9.1.2 Advantages of SSFE method As the frame difference subtraction was completed, the below parameters are used in checking the accuracy of the algorithm. Mean of the Silhouette extracted was 0.0684. Standard Deviation extracted was 0.2524. Entropy Extracted was 0.3599. RMS was 0.1617. The standard Deviation and entropy Extracted was disruptive and discontinuous. Figure-9.1 Silhouette extraction from frame Subtraction method with ideal background Here is a snapshot of the extracted Silhouette from a completely non disturbed background. This Silhouette would be difficult to calculate the feature we consider for the recognition of a person. Hence a new algorithm called Structural Silhouette and Feature Extraction (SSFE) is developed in this project.. 9.2 SSFE Method In Structural Silhouette and Feature Extraction (SSFE), extraction of the foreground (human) by using a matlab function of vision.ForegroundDetector. Then Gaussian filters for boundary clearance. The function needs to be specified as how many frames it should consider for training the function in differentiating the foreground and background. The minimum background ratio is a threshold to determine the background model, stated as a numerical scalar. The extracted foreground in RGB is then converted into grayscale. The human silhouette is extracted in terms of a binary image. Hence by this the feature extraction is done in the similar way as has been discussed in frame subtraction method. The feature extraction parameters remain the same. Figure-9.2 Silhouette extraction from SSFE method with busy background The above image is one of the examples of the silhouette extraction process. Due to better silhouette extraction in SSFE method, feature extraction is done by using the same algorithm. 9.3 Algorithm used in neural network In the neural network the backpropagation algorithm is used in order to train the network. This algorithm trains the network based on the chain rule method. In this

chain rule method, after each forward pass through a network, backpropagation performs a backward pass while adjusting the **13**

model's parameters. The above algorithm shows a mathematical explanation of how our algorithm works. Since we compute the error vector backwards , it is called backward propagation. And 200 hidden layers are considered . 9.3.1 Advantages of backpropagation algorithm: ● Relatively simple implementation. ● Mathematical formulas in the algorithm can be used in any network for training. ● Computing time is less comparatively ● It does not require any prior knowledge about the inputs and is flexible Chapter 10: Results and Discussions In this report, presentation and discussion have been done of two methods of Gait Analysis which can be used for biometrics; one is Frame Difference Method and the other method is SSFE (Structural silhouette and Feature Extraction). Neural networks also have been adopted to avoid the minor changes in Gait. Both methods are based on silhouette extraction of a walking human. The first method that is Frame subtraction, silhouette extracted by subtracting each frame was discussed and the results are displayed. Since the method of frame subtraction is a traditional method used in all the algorithms of video encoding, silhouette extraction was improper in this method. In SSFE method, the silhouette extraction is better and more stable. Hence feature extraction like body width, leg width, leg height and body height (explained in a detailed way in feature extraction chapter) is more feasible in this method. The results of both the methods have been displayed in the below diagrams and mathematical values. Figure-10.1 Frame Subtraction Method Figure-10.2 SSFE Method The Fig 10.1 and 10.1 are of the same video, 10.1 is from frame subtraction and 10.2 is from SSFE. It can be clearly seen from the image, segmentation of the silhouette is not enough to consider feature extraction to be done on the silhouette. Hence SSFE would be a better method to extract features of a human walking. The following table shows the comparison of frame difference method and SSFE method. The experimentation is done with a data set of 75 videos, each video of 5 seconds,with a frame rate of 25 frames per seconds. Table 10.1 Difference Between Frame difference and SSFE Methods Mean of the matrix Standard deviation of the matrix Entropy of the matrix Frame difference method 0.221 0.431 0.236 SSFE method 0.753 0.346 0.236 The values in the above table are obtained by performing operations such as Mean, Standard Deviation and Entropy on the binary matrix of the silhouette extracted. In a binary image, the pixel has a value of either 0/1. Hence taking a mean for the matrix, it is known that the unwanted information is not being extracted, as the value is closer to 1 in SSFE method. Standard Deviation and Entropy is lesser than the Frame Difference Method, Hence the disturbance in the matrix and deviation is less. Better silhouette extraction is being done. A perfect Silhouette extracted video is taken as a standard video from our database and compares all our values to this standardised video. The same subject is taken in a different background and the silhouette is extracted. The values and results are recorded. The imagery results of the many have been displayed below. Figure-10.3 Subject_1, Background_1, Standard Video for person 1 Figure-10.4 Subject_1, Background_2 Experimental Video 1 for person 1 Figure-10.5 Subject_1, Background_3 Experimental Video 2 for person 1 Figure-10.6 Subject_1, Background_4 Experimental Video 4 for person 1 Table 10.2 The mathematical Results of the above Videos Evaluation Parameters Standard Video Experimental Video 1 Experimental Video 2 Experimental Video 3 Mean 0.656 0.5032 0.7125 0.5482 Standard Deviation 0.1369 0.5005 0.4526 0.4977 Entropy 0.0034 1.000 0.8654 0.9933 RMS 0.6281 0.6914 0.8247 0.7118 Skewness 0.1232 0.1316

0.1569 0.1436 Graphical Representation of the features extracted from all the videos (body width) has been shown in the Figure 10.7 Blue - Standard Video Figure 10.7 Graph of body width vs no of frames of subject_1 Black - Experimental Video _1 Green -Experimental Video _2 Red - Experimental Video _3 A different Subject, Subject_2 is taken in different backgrounds. Figure 10.8 Subject_2, Background_5 Standard Video for person 2 Figure 10.9 Subject_2, Background_6 Experimental Video_1 for person 2 Figure 10.10 Subject_2, Background_7 Experimental Video_2 for person 2 Figure 10.11 Subject_2, Background_8 Experimental Video_3 for person 2 Table 10.3 The mathematical Results of the above Videos Evaluation Parameters Standard Video Experimental Video 1 Experimental Video 2 Experimental Video 3 Mean 0.7749 0.6569 0.7678 0.6630 Standard Deviation 0.4178 0.3987 0.4245 0.4216 Entropy 0.7694 0.6743 0.6534 0.5698 RMS 0.8587 0.8034 0.8234 0.7890 Skewness 0.1935 0.1859 0.1968 0.1897 Graphical Representation of the features extracted (body width) has been done in the Figure 10.12 Blue - Standard Video Black - Experimental Video _1 Figure 10.12 Graph of body width vs no of frames of subject_2 Green -Experimental Video _2 Red - Experimental Video _3 A different Subject, Subject_3 is taken in different backgrounds. Subject_3, Background_9 Standard Video for person 3 Subject_4, Background_10 Experimental Video_1 for person 3 Figure 10.15 Subject_3, Background_10 Experimental Video_3 for person 3 Figure 10.16 Subject_4, Background_11 Experimental Video_3 for person 3 Table 10.4 The mathematical Results of the above Videos Evaluation Parameters Standard Video Experimental Video 1 Experimental Video 2 Experimental Video 3 Mean 0.4629 0.6158 0.5757 0.6006 Standard Deviation 0.4987 0.4865 0.4943 0.4898 Entropy 0.9960 0.9609 0.9834 0.9706 RMS 0.6211 0.7276 0.7083 0.7178 Skewness 0.0921 0.1245 0.1189 0.1052 Graphical Representation of the features extracted (body width) has been done in the Figure 10.17 Blue - Standard Video Figure 10.17 Black - Experimental Video _1 Figure 10.17 Graph of body width vs no of frames of subject_3 Green -Experimental Video _2 Red - Experimental Video _3 The Graphical Representation is the Features Extracted, the individual values from the database are then fed to the Neural Networks. From these features, the weights are trained. Once the training is completed the image 10.17 is displayed. Showing that the training is completed. Figure-10.18 A screenshot of the integration of

neural network This imagery is the representation of the training of neural [2]

network for the database. 3 sets of people and 25 videos for each person is considered. Hence it trains according to all the person's database. Final result obtained from the algorithm and neural network combined is Figure-10.19 Image of One of the results of the data runs The identification of a person using their Gait has been successfully completed. Training of Neural Network: Training Accuracy and Testing Accuracy of the Neural Networks have been given below. Training of the Neural Network has been done by using 45, 15 of each person, videos each having 25 FPS. So a total of 5625 frames in total i.e having 5625 features extracted. A training Accuracy of 88.89% was achieved and the same neural network has been used for testing. Testing of Neural Network: Testing of the Neural Network was done with a completely different database of videos. It consists of 5 videos each of a person, a total of 15 videos in total 1875 frames. So a total of 1875 extracted features. A testing accuracy of 80.00% was achieved and the results are displayed below. Table 10.5 Accuracy in Frame matching and SSFE Tested for 3 subjects with 5 videos of each Frame Difference Method SSFE Method Accuracy in % 37.03 80.00 Table 10.6 Accuracy in SSFE with and without Neural Network Tested for 3 Subjects with 5 videos of each Without Neural network With neural network Accuracy in % 55.55 80.00 Chapter 11: Conclusions and Future developments ● A better algorithm could be developed in terms of feature extraction. ● This project is done using a static camera, in the further developments, a moving camera can be used for experimentation. ● This project is focused mainly on biometrics, hence the subject pertains to walk in a particular path. In the additional projects, the subject can be moved without any restrictions. ● Motion detection sensors can be used for a better silhouette extraction. ● Multiple cameras can be used for 360 degree view of the subject for better biometric security. Biometrics using

Gait Recognition Biometrics using Gait Recognition Biometrics using Gait Recognition Biometrics using 1

Gait Recognition Biometrics using Gait Recognition Biometrics using Gait Recognition Biometrics using 1

Gait Recognition Biometrics using Gait Recognition Biometrics using Gait Recognition Biometrics using 1

Gait Recognition Biometrics using Gait Recognition Biometrics using Gait Recognition Biometrics using 1

Gait Recognition Biometrics using Gait Recognition Biometrics using Gait Recognition Biometrics using 1

Gait Recognition Biometrics using Gait Recognition Biometrics using Gait Recognition Biometrics using 1

Gait Recognition Biometrics using Gait Recognition Biometrics using Gait Recognition Biometrics using 1

Gait Recognition Biometrics using Gait Recognition Biometrics using Gait Recognition Biometrics using 1

Gait Recognition Biometrics using Gait Recognition Biometrics using Gait Recognition Biometrics using 1

Gait Recognition Biometrics using Gait Recognition Biometrics using Gait Recognition Biometrics using

1

Gait Recognition Biometrics using Gait Recognition Biometrics using Gait Recognition Biometrics using

1

Gait Recognition Biometrics using Gait Recognition Biometrics using Gait Recognition Biometrics using

1

Gait Recognition Biometrics using Gait Recognition Biometrics using Gait Recognition Biometrics using

1

Gait Recognition Biometrics using Gait Recognition Biometrics using Gait Recognition Biometrics using

1

Gait Recognition Biometrics using Gait Recognition Biometrics using Gait Recognition Biometrics using

1

Gait Recognition Jan-May 2020 1 Jan-May 2020 2 Jan-May 2020 3 Jan-May 2020 4 Jan-May 2020 5 Jan-May 2020 6 Jan-May 2020 7 Jan-May 2020 8 Jan-May 2020 9 Jan-May 2020 10 Jan-May 2020 11 Jan-May 2020 12 Jan-May 2020 13 Jan-May 2020 14 Jan-May 2020 15 Jan-May 2020 16 Jan-May 2020 17 Jan-May 2020 18 Jan-May 2020 19 Jan-May 2020 20 Jan-May 2020 21 Jan-May 2020 22 Jan-May 2020 23 Jan-May 2020 24 Jan-May 2020 25 Jan-May 2020 26 Jan-May 2020 27 Jan-May 2020 28 Jan-May 2020 29 Jan-May 2020 30 Jan-May 2020 31 Jan-May 2020 32 Jan-May 2020 33 Jan-May 2020 34 Jan-May 2020 35 Jan-May 2020 36 Jan-May 2020 37 Jan-May 2020 38 Jan-May 2020 39 Jan-May 2020 40 Jan-May 2020 41 Jan-May 2020 42 Jan-May 2020 43 Jan-May 2020 44 Jan-May 2020 45 Jan-May 2020 46