

*DEEP LEARNING PROJECT ON*

# **3D Gait Recognition on Embedded Devices**

*INTERIM PROGRESS REPORT*

*to be submitted by*

**Group-08**

**Mohit Sharma, Abhijeet Sharma, Akhil Singhal**

*for the course*

*of*

**DEEP LEARNING AND ITS APPLICATIONS**

**CS-671**



**SCHOOL OF COMPUTING AND ELECTRICAL ENGINEERING  
INDIAN INSTITUTE OF TECHNOLOGY MANDI, MANDI**

**March 2018**

# 1 Introduction

Biometrics has quickly entrenched itself as the most pertinent means for identifying and authenticating individuals. The traditional methods no longer ensure the security of our data in the hands of unauthorised persons. Biometric use unique traits of an individual for its identification. The biometrics can be broadly classified into two types namely viz, 1) Physiological and 2) Behavioral biometrics. The former includes the passive features like the fingerprint, iris, face, hand veins patterns etc and the later includes are active like gait, signature voice etc. The physiological traits of an individual are highly unique and discriminant and thus are highly reliable. The physiological biometrics require the cooperation of the individual and also a fully controlled setup for its high reliability. Compared to the conventional biometrics, the behavioural traits are non-invasive, non-contact and are perceivable from a larger distance. This kinds of traits do not require a cooperation of the individual, unlike the physiological traits. Because of these advantages, these features gait recognition can be applied in scenarios like surveillance, criminal investigation and access control. Gait recognition can be affected by many factors including pose variability, clothes, bags etc. All these factors make it a really challenging problem to solve. [1]

The majority of the research work done primarily focuses on developing the novel algorithms for the recognition tasks or improving efficiency and accuracy of existing solutions. These solutions are targeted over general purpose processors that are expensive, large with high computational power and space. Since gait recognition has a large search space and has an extensive amount of computationally intensive operations, it is reasonable to suggest an embedded implementation specifically optimized to detect humans. An embedded solution would have many many advantages including like low cost and integration within the embedded devices. Embedded devices have hardware limitations, less GPU and memory so these models are hosted on the cloud. Embedded devices utilise these models by sending the input data onto the cloud. These cloud-based implementations restrict the use cases i.e device must have internet connectivity, data privacy etc. In order to overcome these restrictions, we need to deploy these onboard. For that model, optimization is needed to make it work on the Embedded device

## 2 Background

Gait is the study of the human locomotion or in other words, it's the manner in which you walk and run. Gait features are an important bio-metric feature for the human identification. It has demonstrated an immense potential for use in the human identification systems. It has been seen that every individual has a particular style of walking which can be used as an important trait to distinguish him from other humans. This has been used for the identification since a long time. The suitability of the gait for the human identification is because it can be perceived from a distance as well as to its non-invasive nature. A huge amount of work has been done in the field of gait recognition. There are two major approaches to the problem one being the model-free approach and the other is the model-based approach. [2]

**Model-Free Approach:** In this approach, the silhouette of different individuals is collected which is then segmented into different parts which are then fitted with some shape whose parameters are calculated to generate a multidimensional vector altering with according to video. A combination of these gives us the required information. Other types of approaches include the Gait Energy Image (GEI) which is the average of silhouette over a gait cycle.

**Model-Based Approach:** In this kind of approach specific parameters are identified and matched to get the desired result. It involves identifying mathematical constructs that represent discriminative gait features with some parameters. The parameters are grouped into spatiotemporal and kinematic. Spatio-temporal parameters include the step length and height (stride) speed and gait cycle time while the kinematics include the joint rotations and the angle between different joints. This approach reliably handles the occlusion, noise, scale etc which fails in the case of model-free approaches.

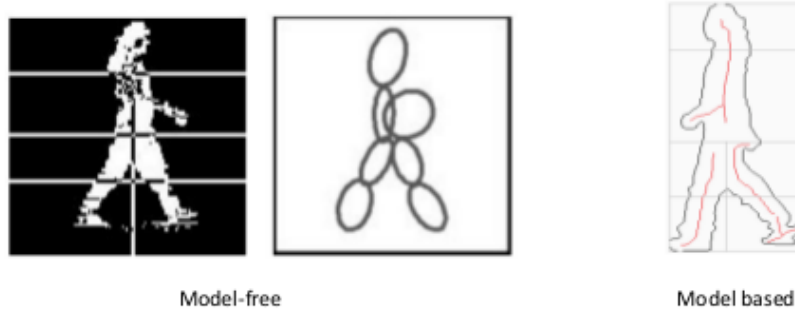


Figure 1: Model-free and Model-based approaches [3]

### 3 Problem Statement

**3D-Gait-Recognition :** Creating a deep learning pipeline for the identification of the person by the manner of its walking i.e. using his/her gait features.



Figure 2: Gait Recognition [4]

### 3.1 Dataset

The dataset that we will be using in the project will be the **Human3.6M** dataset. The dataset consists of 3.6 million different human poses collected with 4 digital cameras. The data is organized into 15 training motions containing walking with many types of asymmetries (e.g. walking with a hand in a pocket, walking with a bag on the shoulder), sitting and laying down poses, various types of waiting poses and other types of poses. The actors were given detailed tasks with examples in order to help them plan a stable set of poses between repetitions for the creation of training, validation and test sets. In the execution of these tasks the actors were however given quite a bit of freedom in moving naturally over a strict, rigid interpretation of the tasks.

### 3.2 Edge Device

**Jetson Tx2** : Jetson TX2 is the fastest, most power-efficient embedded AI computing device. The latest addition to the industry-leading Jetson embedded platform, this 7.5-watt supercomputer on a module brings true AI computing at the edge. It's built around an NVIDIA Pascal-family GPU and loaded with 8 GB of memory and 59.7 GB/s of memory bandwidth. It features a variety of standard hardware interfaces that make it easy to integrate it into a wide range of products and form factors. [5]

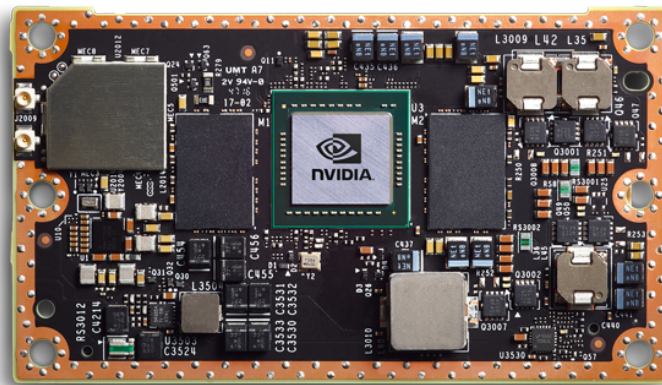


Figure 3: Jetson Tx2 Module

### 3.3 Pipeline

Below is the proposed pipeline for the project :

1. Identifying and creating the Ground Truth data.

2. Getting the individual poses for each of different concerned objects in each frame.(DensePose-RCNN) [6] [7]
3. Establishing the spatio-temporal relationships between the frames in the gait cycle using RNNs or LSTMs.
4. Optimizing the above network by creating TensorRT engine to work on the Nvidia Jetson Tx2.

## References

- [1] P. A. A. N. Daksh Thapar, Divyansh Aggarwal, “Vgr-net: A view invariant gait recognition network,” 2017.
- [2] A. M. H. J. Bogdan Kwolek<sup>1</sup>, Tomasz Krzeszowski<sup>3</sup>, “3d gait recognition using spatio-temporal motion descriptors,” 2014.
- [3] “Factors affecting identification of individuals by gait cycle.”
- [4] “Gait recognition - <https://www.youtube.com/watch?v=emjppglx14a>.”
- [5] “<https://developer.nvidia.com/embedded/buy/jetson-tx2>.”
- [6] N. N. R. A. Guler and I. Kokkinos, “Densepose: Dense human pose estimation in the wild,” 2018.
- [7] P. D. K. He, G. Gkioxari and R. G. M, “Mask rcnn,” 2017.