



Department of Computer Science
Engineering
Shri Ramdeobaba College of Engineering
Nagpur

Virtual Operation Theatre

Guided by:
Prof. Vasundhara Rathod

Members:

Nandini Jain (12)
Riya Khandelwal (15)
Saburi Hindaria (16)
Ruchit Bhardwaj (55)
Isha Bahendwar (77)

Contents

1.0 Problem Definition

1.1 Significance

1.2 Features

1.3 Planning

2.0 Methodology

2.1 Module 1 - Object Detection

2.2 Module 2 - Unity

3.0 Technologies Used

3.1 Arduino

3.2 Unity

3.3 EMG Sensors

3.4 TensorFlow

3.5 Virtual Reality

4.0 Demonstration of Completed Module

Problem Definition

To develop and implement a **VR module that can imitate and reciprocate hand movements** of a person **using Electromyographic Sensors** and **simulate the environment of an operation theatre** for assisting new doctors to identify **various segments of human anatomy** with **machine learning functionality**.

Methodology

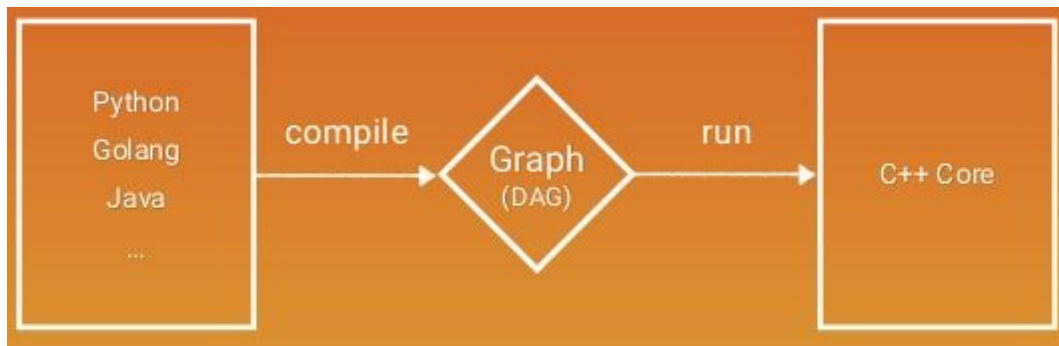
Module 1 - Object Detection

TensorFlow

- In 2011, Google created DistBelief, which used machine learning to identify what's in a photo by recognizing certain patterns.
- Created by the Google Brain team
- Open source library for numerical computation and large-scale machine learning
- TensorFlow takes the concept a step further by using deep learning, or an artificial neural network composed of many layers.

How does TensorFlow work?

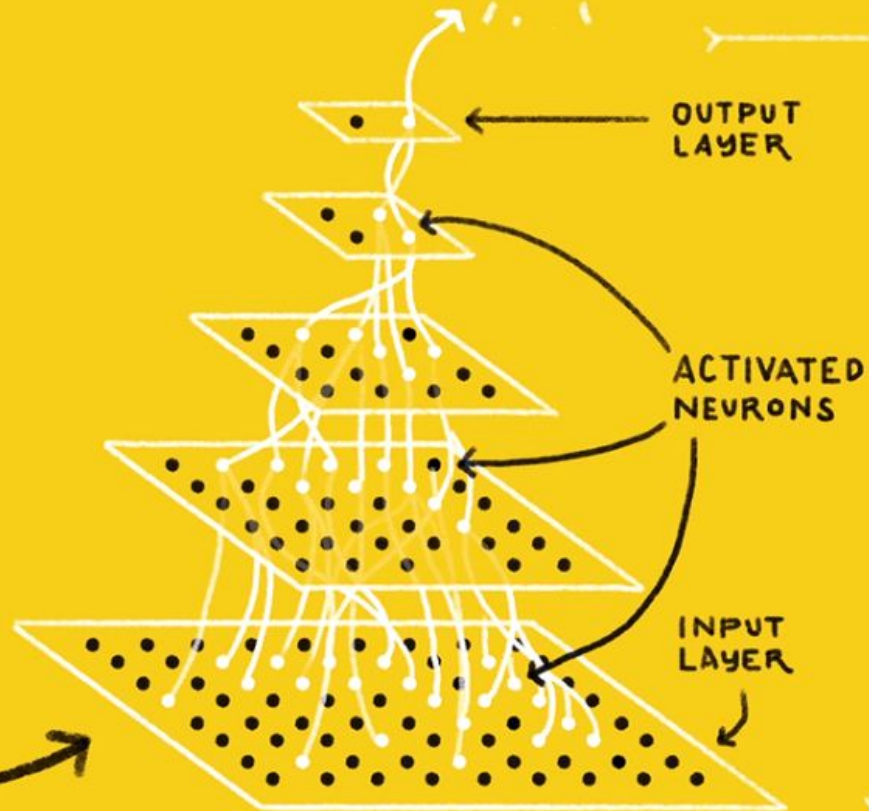
- TensorFlow sorts through layers of data, called nodes, to learn that the image it's viewing is of a cat.
- The first layer will ask the system to look for something as basic as determining the general shape in the picture. The system then moves, or flows, to the next data set - like looking for paws in the photo.



IS THIS A
CAT or **DOG**?



~~CAT~~ - DOG



DEEP
neural
NETWORK

Object Detection

Computer technology related to computer vision and image processing that deals with detecting instances of semantic objects of a certain class

- First Algorithm : Viola-Jones Algorithm (2001)
 - Face Recognition
- HOG : Histograms of Oriented Gradients (2005)
 - Face Recognition
 - Pedestrian Recognition



CNN

- Classify every single box in the image
- Sliding Window Detector
- Inefficient way : Too slow and computationally very expensive

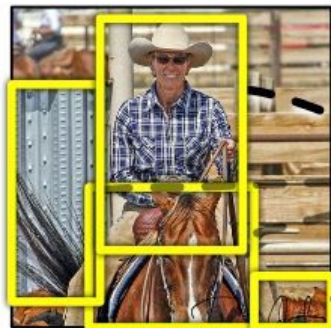
Improved version: **R-CNN**

- Uses Selective Search and Series of windows of different sizes
- Generate Region Proposals from bounding boxes (Threshold)
- Compute CNN features
- Run those images in bounding boxes in pre-trained classifier

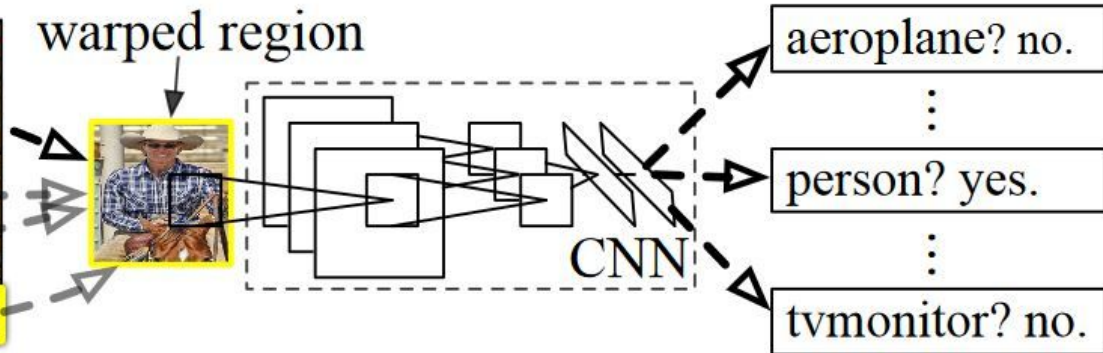
R-CNN



1. Input image



2. Extract region proposals (~2k)



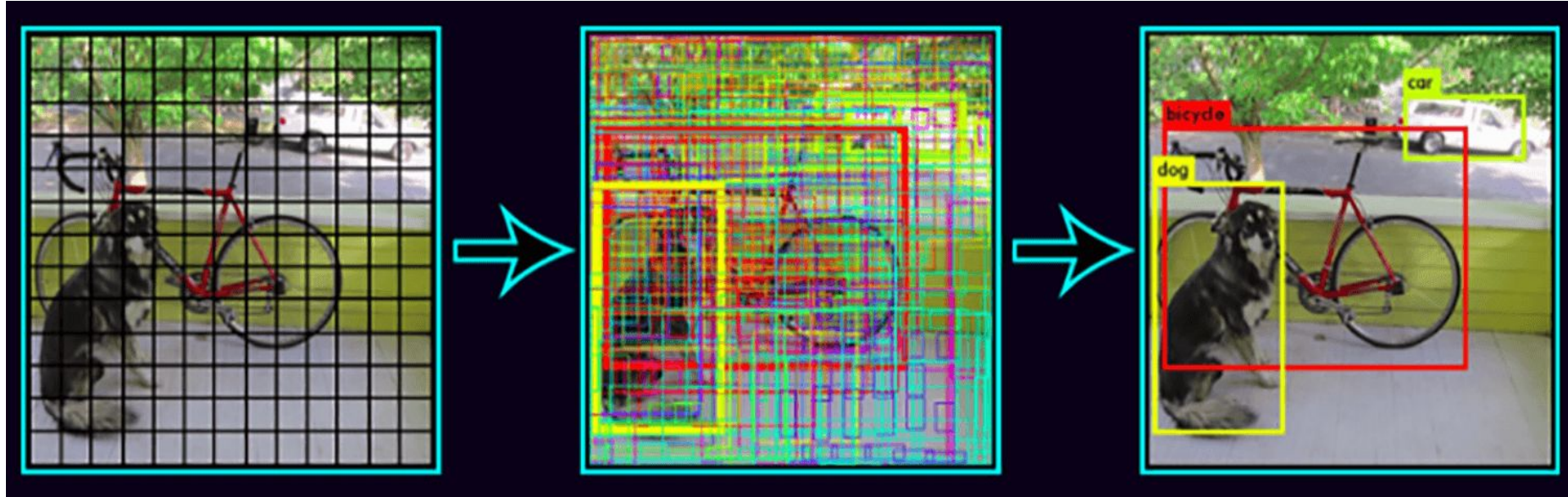
3. Compute CNN features

4. Classify regions

YOLO

- You Only Look Once
- Detects all the classes in input image / video
- 13x13 Grid
- Each cell can be a part of 5 boundary boxes
- Confidence score (Threshold)
- Higher the confidence score, Thicker the bounding box
- Cells predict : Whether or not a bounding box exists and What is its class.
- Trained on Pascal VOC Dataset

How YOLO works?



$13 \times 13 \text{ grid} = 169 \text{ cells} = 169 \times 5 \text{ bounding boxes in total} = 845$

Advantages:

- Speed (45 frames per second)
- Network understands more generalized Object Representations
- Faster Versions

Limitations:

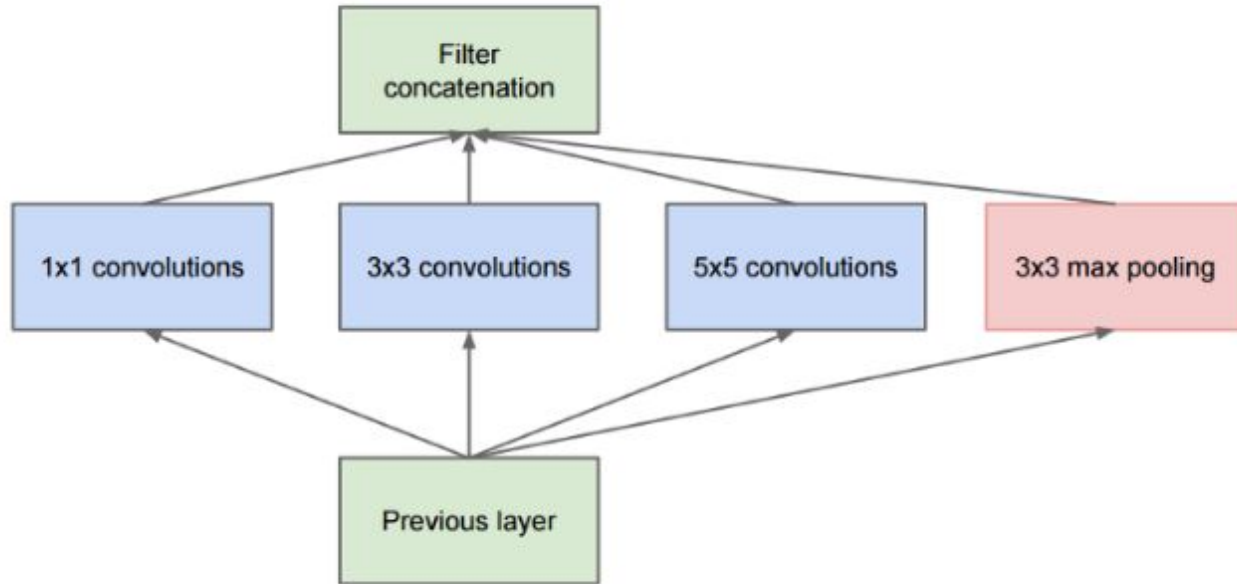
- model struggles with small objects that appear in groups
- generalize to objects in new or unusual aspect ratios or configurations

What should we do now?

Inception Model

- Google released Tensorflow framework and Inception architecture
- 'Inception' was a codename for Deep Convolutional Neural Network architecture
- Main hallmark - improved utilization of computing resources inside the computer with low computational power
- Perfect for our scenario
- D-CNN comes under the class of feed-forward ANNs (No feedback)
- D-CNNs use relatively less preprocessing compared to other Image Classification Algorithms.
- Native feed forward ANNs - 100x100 image - 10000 weights for each neuron in 2nd layer
- D-CNN - Regardless of image size, tiling regions of size 5 x 5, each with the same shared weights, requires only 25 learnable parameters

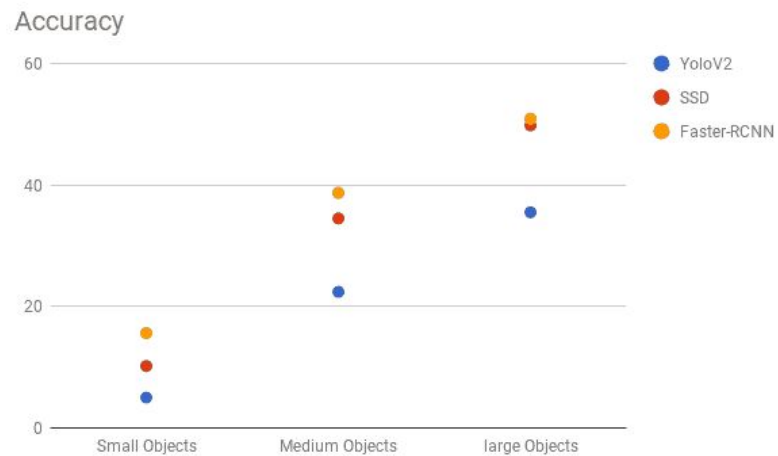
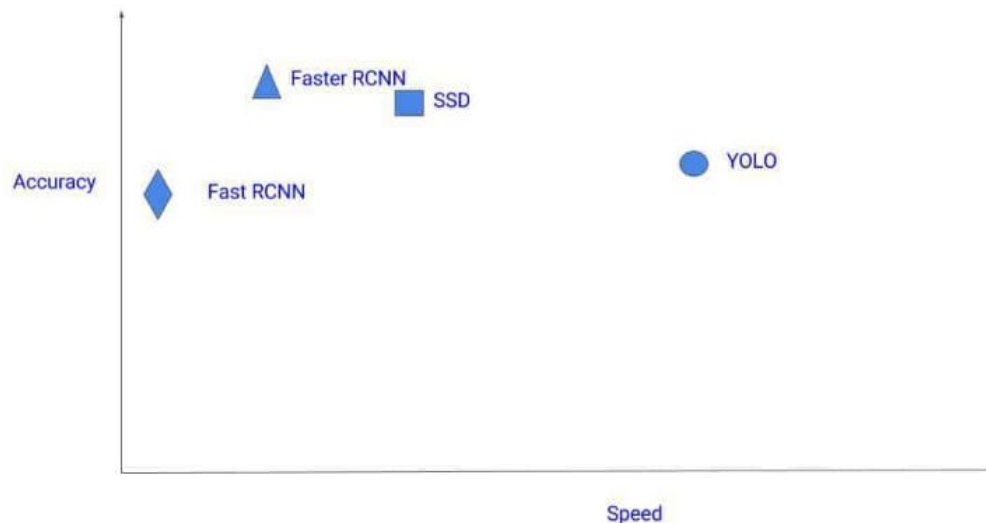
Architecture



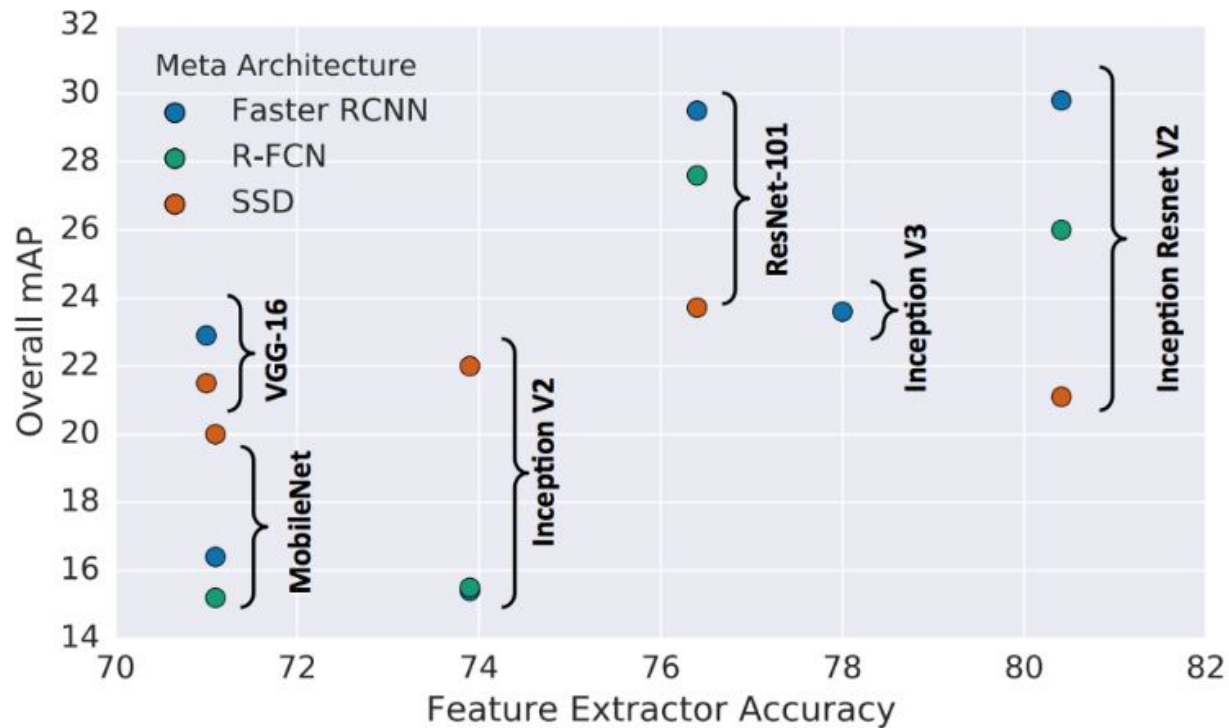
A "naive" Inception module ([source](#))

- Single Shot Detector (SSD) - Balance between speed and accuracy
- Runs a convolutional network on input image only once and calculates a feature map
- We run a small 3×3 sized convolutional kernel on this feature map to predict the bounding boxes and classification probability.
- SSD also uses anchor boxes at various aspect ratio similar to Faster-RCNN and learns the off-set rather than learning the box.
- In order to handle the scale, SSD predicts bounding boxes after multiple convolutional layers.
- Since each convolutional layer operates at a different scale, it is able to detect objects of various scales.
- SSD seems to be a good choice as we are able to run it on a video and the accuracy trade-off is very little.

Comparison (YOLO and SSD)



Versions Of SSD



Analysis

- YOLO - Fast but not so accurate
- SSD - Balance between Speed and Accuracy
- SSD seems to be a good choice as we are able to run it on a video and the accuracy trade-off is very little.
- Aim - Object Detection on a video feed, SSD proved to be a perfect choice.

Constraints

- Low computation power. CPU version of tensorflow. Training not possible. Hence pre-trained model
- Definition specific dataset not available and a generalised approach

TensorFlow Sharp

- The ML-Agents toolkit allows you to use pre-trained TensorFlow graphs inside your Unity games. This support is possible thanks to the TensorFlowSharp project. The primary purpose for this support is to use the TensorFlow models produced by the ML-Agents toolkit's own training programs, but a side benefit is that you can use any TensorFlow model.
- Still an experimental feature
- Attempt to integrate Object Detection in Unity 3D

Methodology

Module 2 - 3D Body Synthesis with Unity & Vuforia

Process Flow

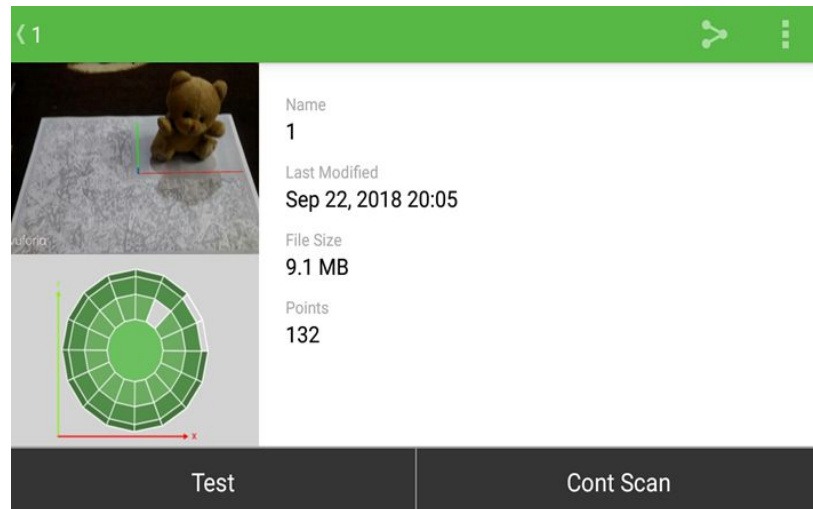
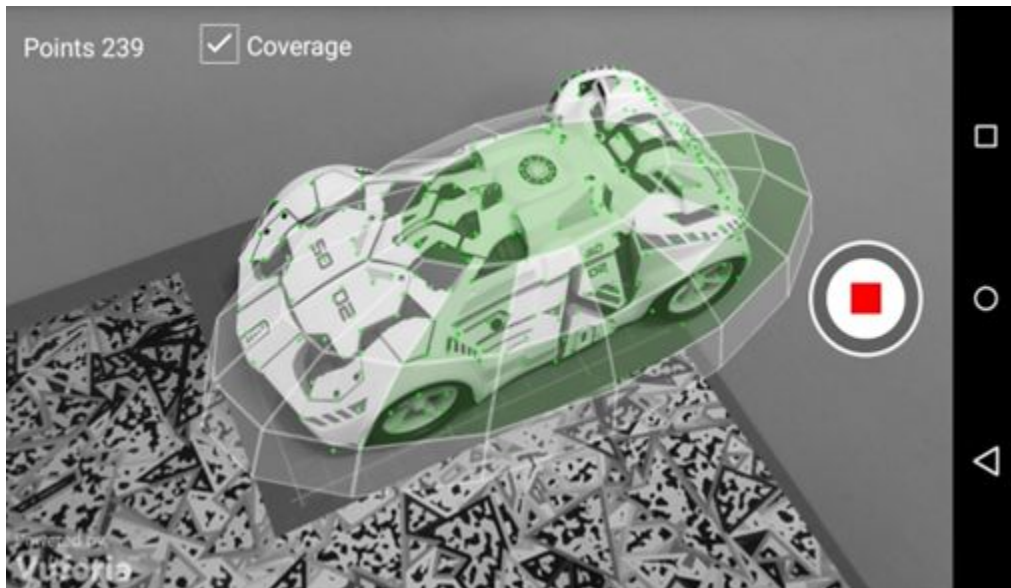
Scanning the 3D
Object

Vuforia Scanner

- The Vuforia Object Scanner allows you to create a target by scanning an object with an Android device.
- Simply install the app, place an object on the Vuforia scanning target, and start the scan.
- Gives real-time visual feedback on the scan progress and target quality
- Establishes a coordinate system so that you can build immersive experiences with precisely aligned digital content.
- The test mode allows you to evaluate the recognition and tracking quality within the app before you start any development.



Vuforia Scanner



Understanding The Object Tracking Criteria

1. Opaqueness - It should not be transparent or semi-transparent
2. Rigidity – Hard and should not be easily bent
3. Minimum moving parts



Pliable and Deformable Objects Are Not Supported

BAG



CAP

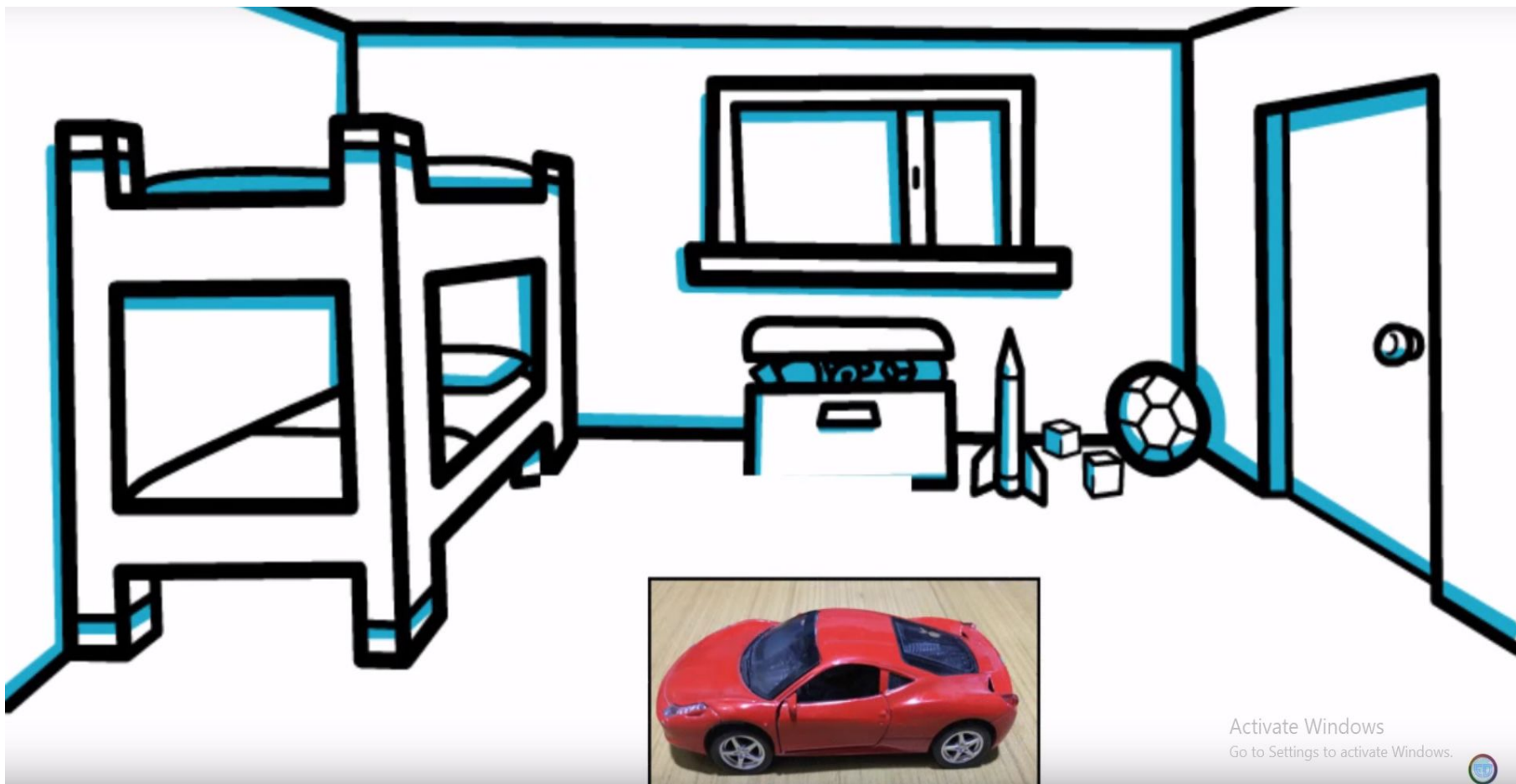


BACKCOVER



Activate Windows
Go to Settings to activate Windows.

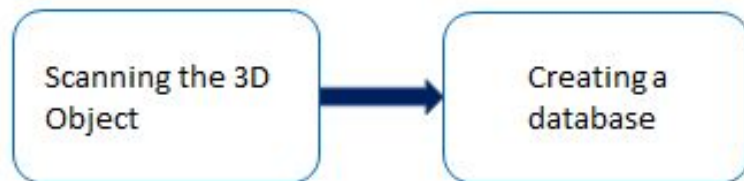
Environmental Criteria



Activate Windows
Go to Settings to activate Windows.




Process Flow



Creating The Database

 **vuforia™** Developer Portal

Hello Saburi  | [Log Out](#)

[Home](#) [Pricing](#) [Downloads](#) [Library](#) [Develop](#) [Support](#)

[License Manager](#) [Target Manager](#)

Target Manager

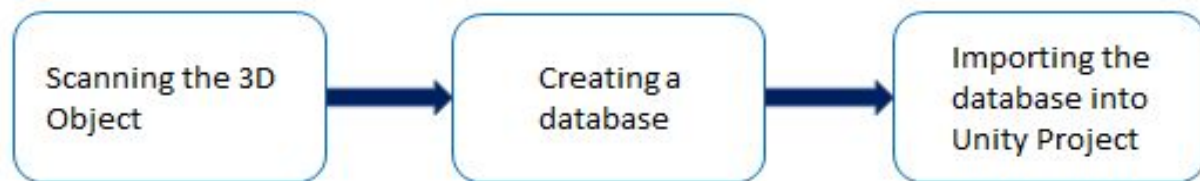
Use the Target Manager to create and manage databases and targets.

[Add Database](#)

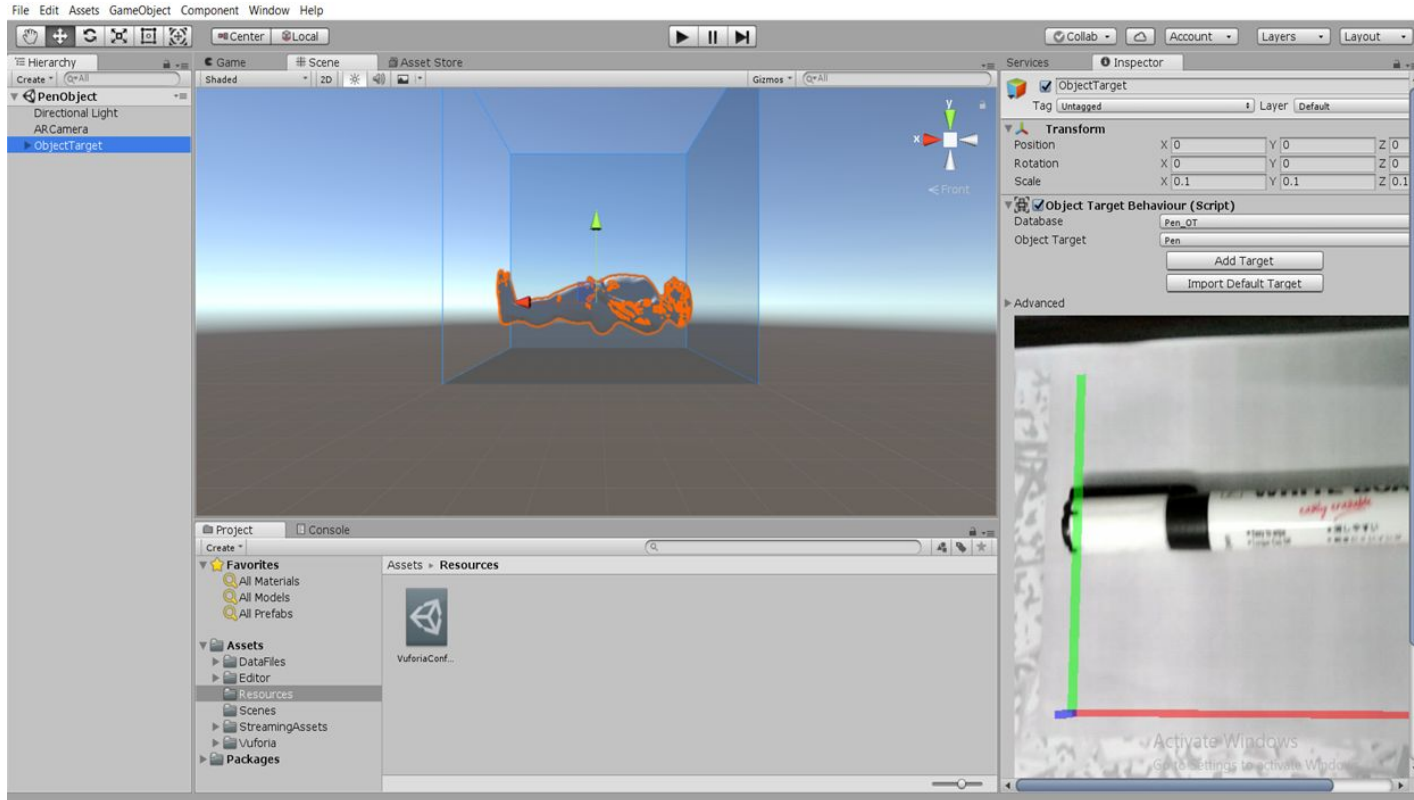
Database	Type	Targets	Date Modified
HumanBody2D	Device	1	Sep 23, 2018 19:41
HumanScan	Device	1	Sep 22, 2018 20:15
Pen	Device	1	Sep 23, 2018 11:25
PseudoHumanDB	Device	1	Sep 21, 2018 09:04
Teddy	Device	1	Sep 21, 2018 11:43

Last updated: Today 12:44 AM [Refresh](#)

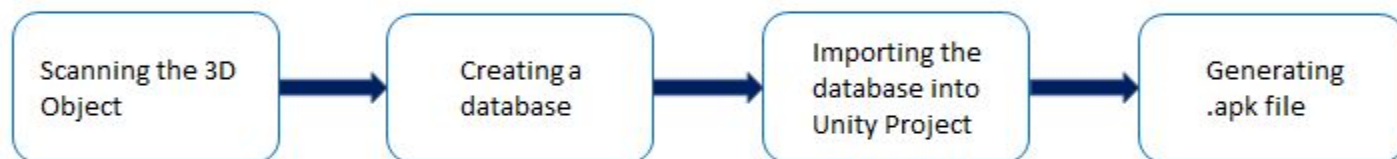
Process Flow



Importing The Database Into Unity Project



Process Flow



Build Settings



Unable to access Unity services. Please log in, or request membership to this project to use these services.

Scenes In Build

☒ DataFiles/Scenes/PenObject

0

Add Open Scenes

Platform

PC, Mac & Linux Standalone

Android

iOS

tvOS

Xbox One

PS Vita

PS4 PS4

Universal Windows Platform



Android

Texture Compression

Don't override

ETC2 fallback:

32-bit

Build System

Gradle

Export Project

☐

Run Device

Fetching connected devices..

Development Build

☐

Autoconnect Profiler

☐

Script Debugging

☐

Scripts Only Build

☐

Compression Method

Default

SDKs for App Stores



Xiaomi Mi Game Center

Add

[Learn about Unity Cloud Build](#)

Switch Platform

Player Settings...

Build

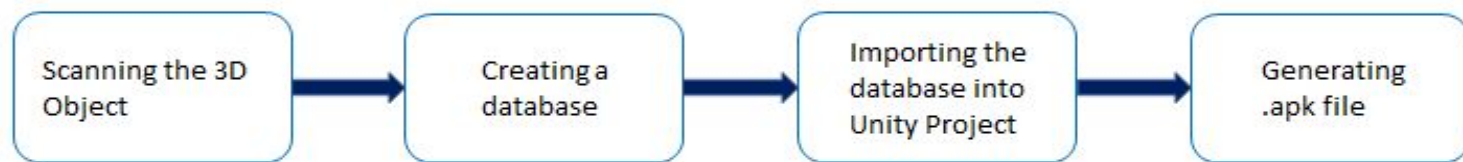
Build And Run

Select a Scene

Choose your Platform

Build a .apk file for your mobile device

Process Flow



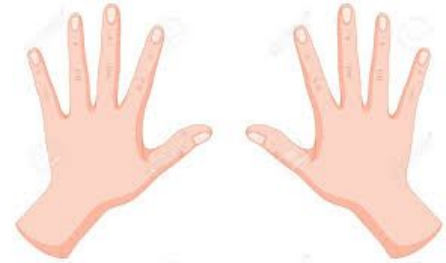
Moving a next step forward



Scans



Projected



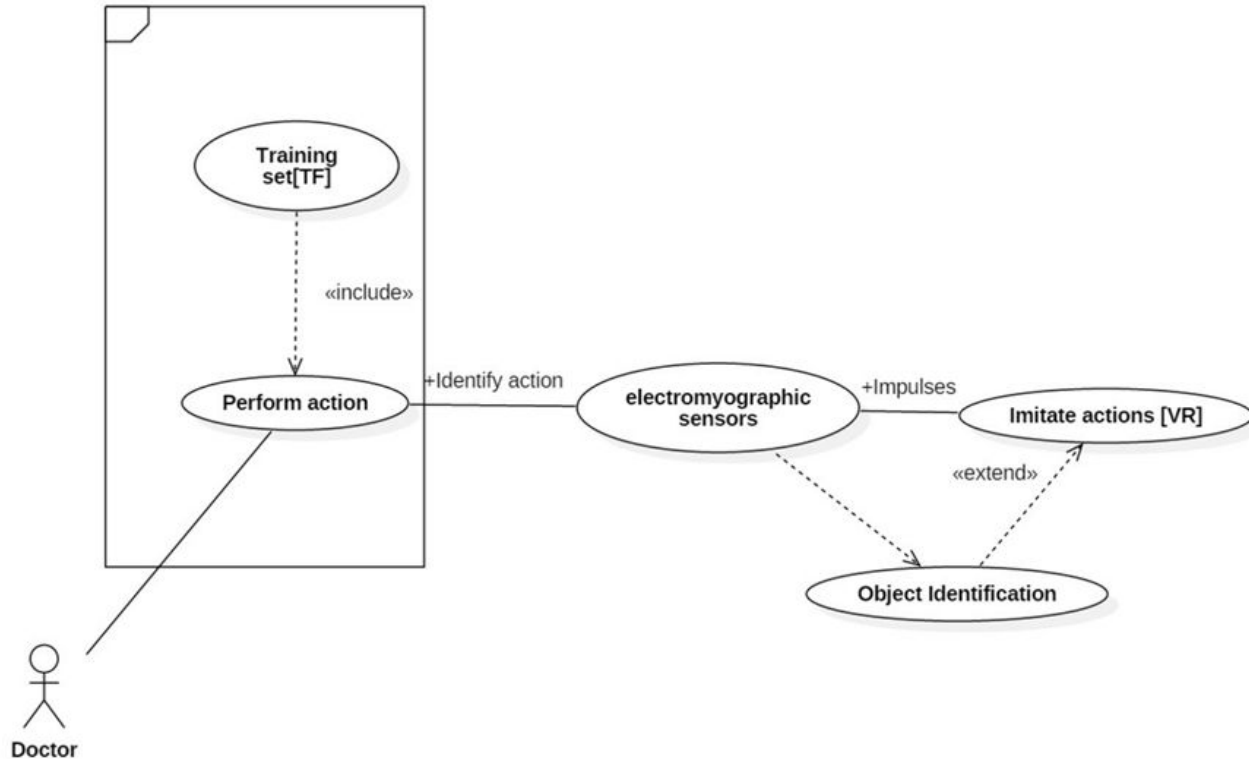
Hand Movements to
navigate the
projected 3D model

WHY?

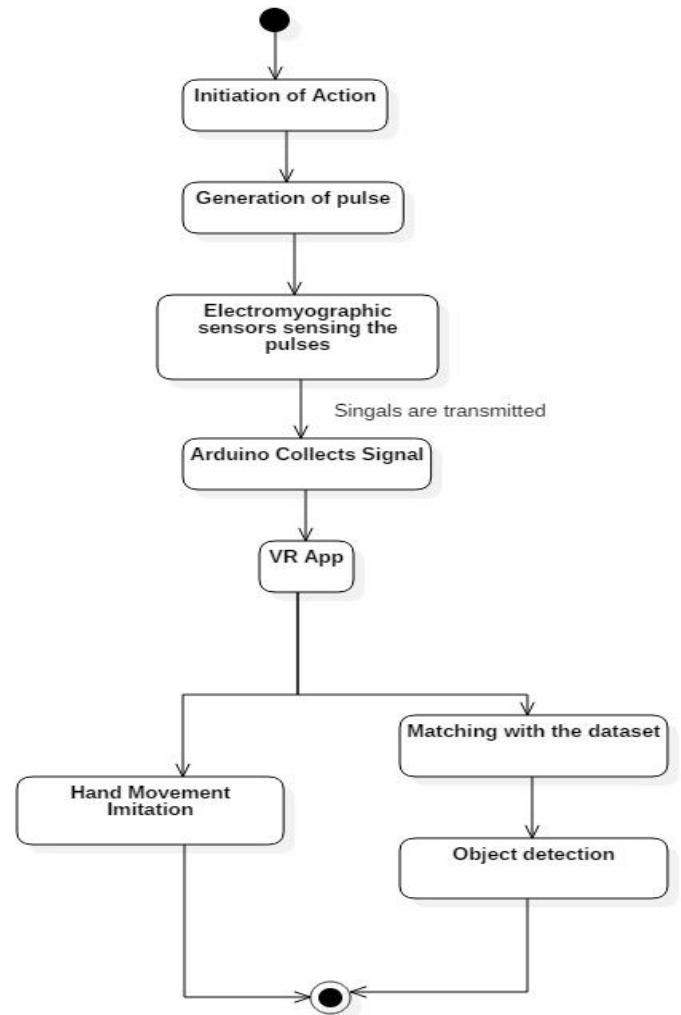
Our Project Offers

- ***Risk Free Training*** : One of the biggest issues faced by practical trainings is the large risk initiated by putting the trainees in a new and uncontrolled environment. The virtual reality simulations neutralise this risk while keeping the same training features, by creating the same environment virtually and putting the trainees inside without having to worry about the risk.
- ***Realistic Scenarios*** : The simulations are built based on real-life operations and manipulations needed in after training situations.
- ***Improves Retention and Recall*** : The main purpose of a training and what really differentiate it from theoretical learning is building a muscle memory. VR trainings solves this problem by enabling trainees to use their hands to manipulate anything inside the simulation.
- ***Suitable for different learning styles*** : When doing a training with virtual reality, the trainee will have all the freedom to test anything in the simulation in order to build an in depth understanding according to his own learning style.
- ***Simplifies complex problems/solutions*** : Virtual Reality simulations are built in such a way to simplify the most complex notions and situations that cannot be understood with traditional trainings.

Use Case Diagram



Activity Diagram



Technologies Used

Arduino Uno

- Arduino is an open-source electronics platform based on easy-to-use hardware and software.
- Arduino boards are able to read inputs, and turn it into an output.
- Arduino board designs use a variety of microprocessors and controllers.
- The microcontrollers are typically programmed using a dialect of features from the programming languages C and C++.



Unity

- Unity is a cross-platform game engine developed by Unity Technologies
- Can be used to create both 3D and 2D games as well as simulations for desktops and laptops, home consoles, smart TVs, and mobile devices.
- The engine offers a primary scripting API in C#, for both the Unity editor in the form of plugins, and games themselves, as well as drag and drop functionality.
- The Unity editor is supported on Windows and macOS, with a version of the editor available for the Linux platform(experimental)
- Written in C++ (Runtime), C# (Unity Scripting API)

EMG Sensors

- Electromyography (EMG) is an electrodiagnostic medicine technique for evaluating and recording the electrical activity produced by skeletal muscles.
- An electromyograph detects the electric potential generated by muscle cells when these cells are electrically or neurologically activated.
- EMG is performed using an instrument called an electromyograph to produce a record called an electromyogram.
- The signals can be analysed to detect medical abnormalities, activation level, or recruitment order, or to analyse the biomechanics of human or animal movement.



TensorFlow

- TensorFlow is an open-source software library for dataflow programming across a range of tasks. It is a symbolic math library, and is also used for machine learning applications such as neural networks.
- It can run on multiple CPUs and GPUs (with optional CUDA and SYCL extensions for general-purpose computing on graphics processing units).
- It is available on 64-bit Linux, macOS, Windows, and mobile computing platforms including Android and iOS.
- TensorFlow computations are expressed as stateful dataflow graphs. The name TensorFlow derives from the operations that such neural networks perform on multidimensional data arrays. These arrays are referred to as "tensors".

Virtual Reality

- Virtual reality (VR) is an interactive computer-generated experience taking place within a simulated environment, that incorporates mainly auditory and visual, but also other types of sensory feedback like haptic.
- Current VR technology most commonly uses virtual reality headsets or multi-projected environments that simulate a user's physical presence in a virtual or imaginary environment.
- VR is most commonly used in entertainment applications such as gaming and 3D cinema.
- In social sciences and psychology, virtual reality offers a cost-effective tool to study and replicate interactions in a controlled environment.
- Surgery training can be done through virtual reality. With the supervision of experts to provide feedback, simulated VR surgical environment provide effective and repeatable training at a low cost, allowing trainees to recognize and amend errors as they occur.

Future Scope

- Medical tutorials can be provided by specialists to many people
- Physiotherapy can be monitored closely
- Solving the issue of lack of cadavers.

References

- <https://medium.com/initialized-capital/we-need-to-go-deeper-a-practical-guide-to-tensorflow-and-inception-50e66281804f>
- <https://cv-tricks.com/object-detection/faster-r-cnn-yolo-ssd/>
- <https://github.com/Unity-Technologies/ml-agents/blob/master/docs/Using-TensorFlow-Sharp-in-Unity.md>
- <https://www.infoworld.com/article/3278008/tensorflow/what-is-tensorflow-the-machine-learning-library-explained.html>

Thank You