A Project Report

On

OBJECT TRACKING WITH CAMSHIFT ALGO

BY

PUSHAPM SINGH

2019B4A71272H

AT

A Practice School-I Station of

A Project Report

On

OBJECT TRACKING WITH CAMSHIFR ALGO

BY

PUSHAPM SINGH

2019B4A71272H

PREPARED WITH FULFILMENT OF THE

PRACTICE SCHOOL 1 COURSE NO.

BITS F221

AT

A PS1 CENTER OF BITS PILANI

ACKNOWLEDGEMENT

First of all, we would like to express my gratitude to the Dean, Practice School Division, and his team, for making it feasible for me to take part in the Practice School-1 even during such difficult times. We would also like to thank our Practice school faculty, Professor Chittaranjan Hota who has worked religiously in making all the preparations and arranging meets so that I do not face any problems.

Next, I would like to express my sincere vote of thanks to Chittaranjan Hota sir and Akash sir, for considering me for this internship and letting me work on the project: "OBJECT TRACKING WITH CAMSHIFT ALGORITHM". Their help and guidance have been an integral part of the working of the project.

I would also like to thank everyone else who was directly or indirectly involved to provide me with this opportunity.

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE PILANI

(RAJASTHAN)
Practice School Division

Station- **CELEBAL TECH** Centre- **JAIPUR**

Duration- 6 Weeks Start Date- 31/05/2021
Date of Submission- 25/06/2021
Title of the Project: OBJECT TARACKING WITH CAMSHIFT ALGORITHM

NAME	ID	DISCIPLINE
PUSHPAM SINGH	2019B4A71272H	MATHEMATICS + COMPUTER SCIENCE

EXPERTS:

Mr Akash Verma (ML mentor Celebal tech.)

Name of PS Faculty: Mr CHITTRANJAN HOTA

Keywords: COMPUTER VISION, OPENCV, MACHINE LEARNING, OBJECT TRACKING CAMSHIFT ALGO

Project Area: COMPUTER VISION

Abstract: Identifying where a particular object is present in the image. Object tracking allows us to apply a unique ID to each tracked object, making it possible for us to count unique objects in a video. With this technique, we can keep a log of several objects in the video. This can be used to keep surveillance and for traffic management.

Date: 25/06/2021

Table of contents:

1. INTRODUCTION

1.1 ABOUT THE COMPANY

1.1.1 Company Profile

Celebal Technologies is a premier software services company in the field of Data Science, Big Data, and Enterprise Cloud. We help you achieve a competitive advantage with intelligent data solutions, built using cutting-edge technology. Our core offerings are around Cloud Innovation, Supply Chain Analytics, Chat Bots, Power Platforms, and Data Analytics. Our solutions can help you accelerate decision-making and take giants leaps in your digital transformation journeys.

1.1.2 PARTNERS

1.1.3 Mission

Their mission is to make data simple and easy to understand for all organizations. We are committed to providing solutions powered by modern cloud and artificial intelligence that integrate with traditional enterprise software. Their tailor-made solutions help enterprises maximize productivity, improve speed and accuracy.

1.1.4 TEAM

Their Team is made up of highly motivated and talented professionals, who understand the core value of modern analytics over the traditional enterprise. We follow the concept of non-hierarchical work culture, empowering every individual in the organization to do their best. We are always on the lookout for skilled professionals who believe in integrity and respect and enjoy working in a fast-paced and autonomous environment.

1.2 ABOUT THE PROJECT

Object tracking is the process of locating a moving object (or multiple objects) over time using a camera. It has a variety of uses, some of which are: human-computer interaction, security and surveillance, video communication and compression, augmented reality, traffic control, medical imaging[1] and video editing. Video tracking can be a time-consuming process due to the amount of data that is contained in the video. Adding further to the complexity is the possible need to use object recognition techniques for tracking, a challenging problem in its own right.

In this project, I used cam-shift algorithm to detect the several objects using their pixel density in the vides,

Uses of object tracking.

1. Security
2. Surveillance
3. Video communication
4. Augmented reality
5. Traffic control
6. Medical imaging
7. Video tracking

1.3 what is computer vision

1.3.1 overview

Computer vision is a domain in which humans use computers to read and analyse images and videos. The field uses math and several coding languages and data structures to convert a 3 layer image or a frame to a binary array so that the computer can read and understand the image to process it.

Several techniques of math have a heavy influence on the field of computer vision, like linear algebra and calculus are few to mention.

Computer vision is concerned with the theory behind artificial systems that extract information from images. The image data can take many forms, such as video sequences, views from multiple cameras, multidimensional data from a 3D scanner, or medical scanning device. The technological discipline of computer vision seeks to apply its theories and models to the construction of computer vision systems.

Several open-source libraries allow a user to work on computer vision in several languages like python, c++, java etc.

For this project, I am going to use python as the primary language to code on using lib like an open cv, TensorFlow and Keras, etc.

1.3.2 OPEN CV

Open cv is an open-source free to use c language-based image processing framework that is used to work on that uses wrapper functions to make it work on python and java as well. This lib enables us to use real-time computer vision. This framework was originally developed by Intel, it was later supported by Willow Garage then Itseez.

Using this lib we can implement camshaft and mean-shift algorithms and several other computer vision features like HSV, contour detection and many more.

1.4 object tracking algorithms.

1.4.1 mean shift algorithm

This algorithm use Mean shift is a procedure for locating the maxima—the modes—of a density function given discrete data sampled from that function. This is an iterative method, and we start with an initial estimate x.

This algorithm needs to know the original position of a particular object to track it. This algorithm is not good for objects that change shape and size with time as this algo won't be able to adapt to the changes that the object is making with each frame.

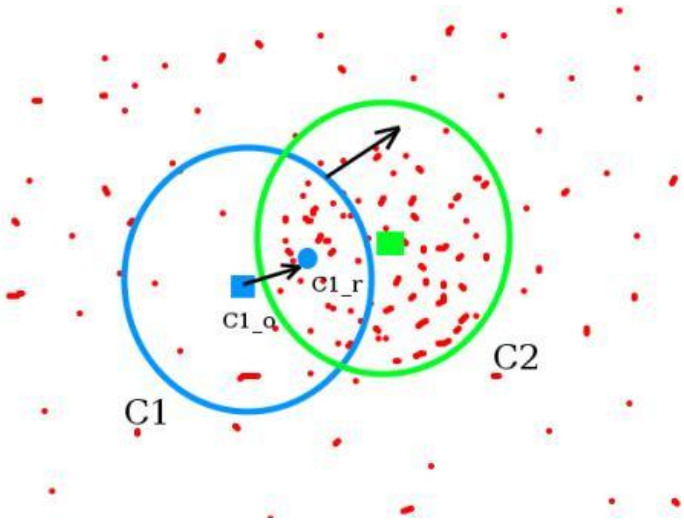
Strengths

1. Mean shift is an application-independent tool suitable for real data analysis.
2. Does not assume any predefined shape on data clusters.
3. It is capable of handling arbitrary feature spaces.
4. The procedure relies on the choice of a single parameter: bandwidth.
5. The bandwidth/window size 'h' has a physical meaning, unlike *k*-means

Weakness

1. The selection of window size is not trivial.
2. The inappropriate window size can cause modes to be merged, or generate additional “shallow” modes.
3. Often requires using adaptive window size.

1.4.2 Cam-shift algorithm



Camshift or Continuously Adaptive Mean Shift is the better version of the mean-shift algorithm that can adapt to the size and orientation change of an object with a frame where the mean shift algorithm is not. Like mean-shift, this algorithm is available for free on OpenCV.

The mean-shift is a non-parametric approach allowing (among others) to detect the modes of a probability distribution by a recursive procedure converging towards the

closest stationary points. This approach tends to ignore the effect of *outliers*, compensating for the effect of *distractors* and image noise.

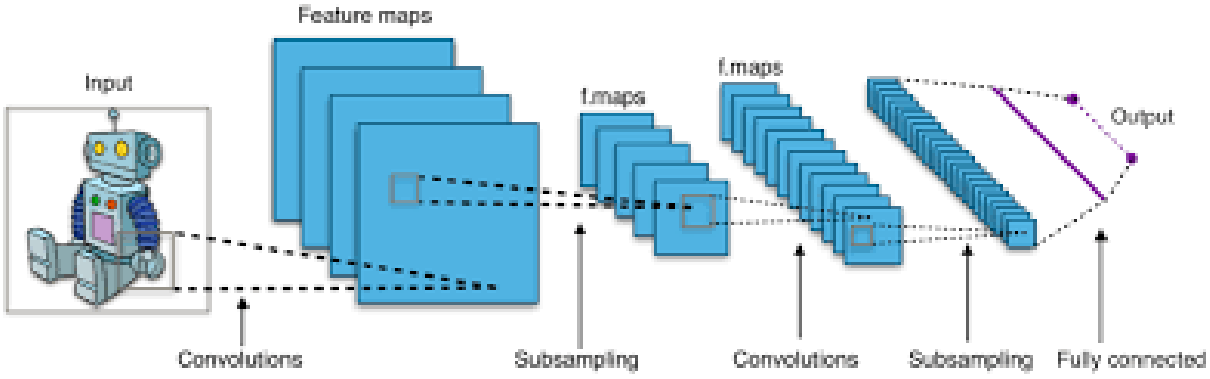
1.5 MACHINE LEARNING

1.5.1 OVERVIEW

machine learning is an algorithmic technique that tends to get better with time through experience and by the use of data. This is a part of artificial intelligence that was first developed in the 1970s. Machine learning involves computers discovering how they can perform tasks without being explicitly programmed to do so.

It involves computers learning from data provided so that they carry out certain tasks. For simple tasks assigned to computers, it is possible to program algorithms telling the machine how to execute all steps required to solve the problem at hand; on the computer's part, no learning is needed. For more advanced tasks, it can be challenging for a human to manually create the needed algorithms.

1.5.2 What is CNN



This is a part of a deep learning module called a convolutional neural network. Unlike conventional neural networks, convolutional neural networks have several convolutional, pooling, fully connected, receptive fields layers added to them.

- Architecture
    - Convolutional layers
    - Pooling layers
    - Fully connected layers
    - Receptive field
    - Weights

CNN takes tensors as data which is a data structure widely used in image processing.

A convolutional layer within a CNN generally has the following attributes:

- Convolutional filters/kernels are defined by width and height (hyper-parameters).
- The number of input channels and output channels (hyper-parameters). One layer's input channels must equal the number of output channels (also called depth) of its input.
- Additional hyperparameters of the convolution operation, such as padding, stride, and dilation.

But there is a major flaw in CNN. CNN cant work with data that has a sequential history of events which is important for the prediction of its next move.

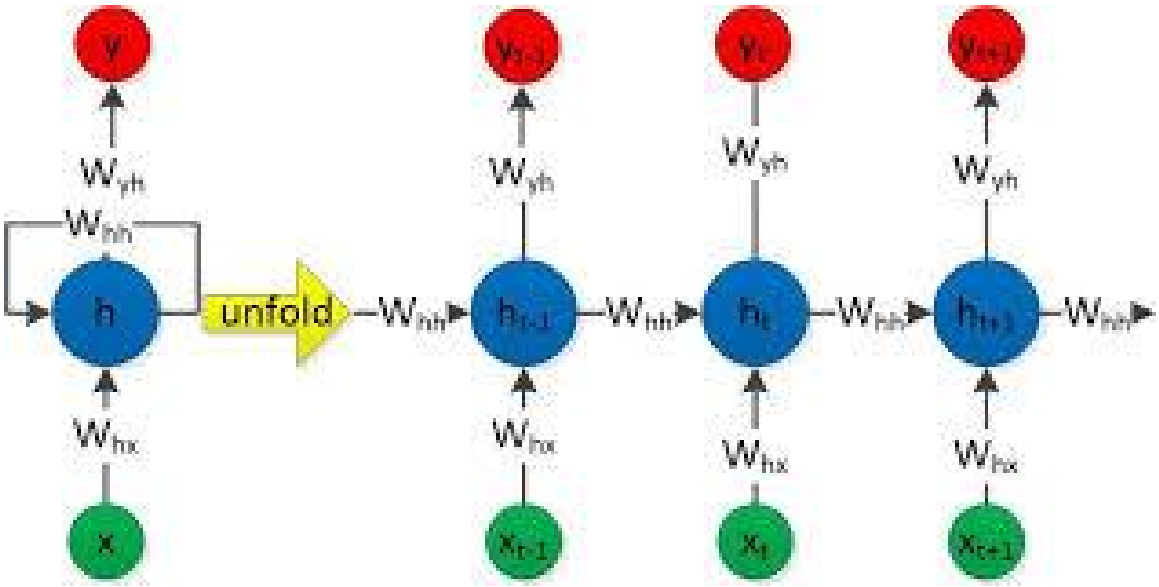
CNN has no short term memory to remember any chance of events that might be helpful for the predictions.

1.5.3 WHAT IS RNN

RECURRENT NEURAL NETWORKS or RNN is a class of artificial neural networks where connections between nodes form a directed graph along a temporal sequence. This allows it to exhibit temporal dynamic behaviour.

This allows this algorithm to work efficiently with sequential data and has a kind of memory to deal with any time event-based data. This derived from feedforward neural networks, RNNs can use their internal state (memory) to process variable-length sequences of inputs.

The following image is the basic architecture of an RNN neural network



2. PROJECT DESCRIPTION

2.1 HOW TO READ AN IMAGE.

Images are usually in 3 channel RGB mode, a computer can depict this as a 3-dimensional array with each dimension representing a colour.

In open cv using python, we can make a NumPy array that has 3 dimensions and the colour patter BGR, the first layer is of blue colour and the second layer is of green colour and lastly, the third layer is of red colour.

- Requirements

    1 numpy

    2 open cv

2.2 how to read a video

A video is a collection of frame and a code can just look at it as a collection of images. One can set a loop to keep updating the frames and that is bluehow we can read a video.

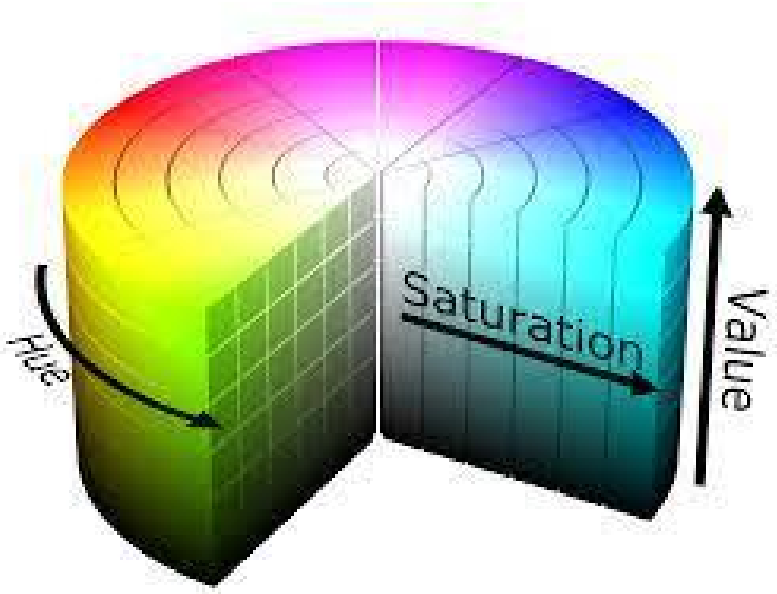
- Requirements

    1 open cv

    2 numpy

    3 the target mp 4 file

2.3 HSV COLOR COORDINATE SYSTEM



HSV or hue saturation and value is a cylindrical coordinate-based system that is used to make masks in computer vision projects.

H - hue

S - Saturation

V - Value

We can add track bars in a separate window to control the values of each of the 3 parameters in the HSV coordinate system to make suitable colour based masks.

2.4 Contour-based detection.

This technique uses the area of an object to detect its boundaries. Users can set a minimum area and maximum area limit to detect the desired object. This technique is not very useful for dynamic detection where the object changes its shape and orientation with time. In this project, we will only use this feature as the initial data provider to the camshaft algorithm as the camshaft algorithm need the initial position of the object to work on it.

- Requirements

  1 open cv

  2 NumPy

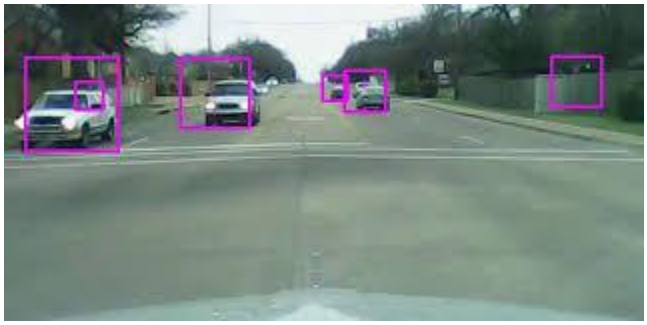
2.6 Motion-based object detection

In this method of object detection, one can use OpenCV to make a mask of the objects that are moving in the frame and hence pass into the camshaft to be tracked later. You need a steady camera video setup for this to work.

```
#parser.add_argument('image', type=str, help='path to image file')
#args = parser.parse_args()
cap = cv.VideoCapture('highway.mp4')

# take first frame of the video
ret,frame = cap.read()

# setup initial location of window
x, y, w, h = 300, 200, 100, 50 # simply hardcoded the values
track_window = (x, y, w, h)

# set up the ROI for tracking
roi = frame[y:y+h, x:x+w]

hsv_roi =  cv.cvtColor(roi, cv.COLOR_BGR2HSV)

mask = cv.inRange(hsv_roi, np.array((0., 60.,32.)), np.array((180.,255.,255.)))

roi_hist = cv.calcHist([hsv_roi],[0],mask,[180],[0,180])

cv.normalize(roi_hist,roi_hist,0,255,cv.NORM_MINMAX)

# Setup the termination criteria, either 10 iterations or move by at least 1 pt
```

// some code from the project

This is what it looks like, one can add some filters to remove irrelevant objects like moving trees



// VISUAL DEMO OF THE PROJECT

2.7 LOG FEATURE

Having a log of all the target objects in a .csv file can open a window of new options. Properly storing data is very important for an institution so that when time ask one can use that data for future analysis with the help of data science.

This project use pandas and NumPy to store data in runtime and then at the end use that data frame and arrays of data to be saved as a .csv of .excel file for future use. With this feature,, we can add all kinds of markers to a particular target so that if needed in future that data will be of proper use.

This project takes note of the position of all the corner points of an object with pinpoint date and time and will also take any specific comments that a user might want to pass in for reference. With this one can keep track of the target object im a professional way.

2.8 Whats new

This project alos include a HSV color coordinate system window so that if for some reason the object in question needs to be detected bsed solly on its color then one can use this window to set limits for

1. Hue
2. Saturation
3. Value

One can use trackbars to set the limits as this projet will show a live changing window to help the user in selecting the limit value.

This project also have a bit mask window of the region of interest so that one can see the corresponding bit mask on which the project is giving results, for this purpose the project use background subtraction available on opencv.

`One key feature of this project is that the user can save the snap of the target object with press of a button. This will record the time at which the target appear in the frame with its corresponding picture and its corresponding bit mask. A user can also save any specific comment related to the target object. This will creat a .csv file that will contain object's information with its coordinates, so that one can use the data in future.

For this task the project use pandas and numpy to make dataframe and then to save the copy of the dataframe in the same directory next to the snaps of target object.

2.9 TECH USED

This project use python as the primary language but libraries like numpy and opencv use c and c++ for the faster computation, python use a wrapper function to call these c and c++ functions. Uses get to experience faster performance as compared to python c and c++ are very fast.

The key lib that the project use is opencv, this is the primary lib used for the purpose of computer vision and video processing. This lib is written in c and c++ and this is a open source computer vision framework.

One of the most important function of opencv in this project is to detect moving objects in a video by doing frame by frame bit operations to detect the any movement in the frame. This is called background subtraction and this method make a mask of all the object that move on the frame by frame basis in a video, note that for this operation to work one need a steady stare camera setup.

Other functions of the opencv lib are video and image pre processing and this project use opencv also for saving images and video.

2.10 TARGET AUDIENCE

With the rise in traffic across the world installations like a toll gate needs optimization. Using this project one can keep track of every car, bike or any kind of vehicle that has crossed the gate with the time logs, pictures and with further improvisation even the number plate information. Some key use for this project are.

Surveillance

Traffic management

Robot vision and activity

This can help optimize the crowd management system as this can detect and notify the authorities of any kind of unusual gathering at any place. This can help with the covid situation too.

3.Conclusion

3.1 Major Milestones

1. Using open cv
2. Deploying RNN and CNN models
3. Using Keras api for image pre-processing

3.2 Technical Skills Acquired

1. Learned open cv
2. Research several ML models
3. Worked on Keras API
4. Got familiar with TensorFlow,

3.3 Soft Skills Learnt

1. Gained experience by participating in a Group Discussion.
2. Interacted with professional corporate people.

3.4 FUTURE SCOPE

Whit the help of CNN this project can also detect number plate of cars and bikes and this can be used to link directly with the the owner of a automobile. This will help our security forces to detect any kind of stolen vehicle.

Open source models are available that can detect number plates of vehicle and and with some weight manipulation this project can use those models to detect and track the number plate in question.

This model can also detect any vehicle that is moving beyond the speed limit of the road and henc this will help in the traffic situation, this can inform the police and can even fine the vehicle owner, this project will alos keep snaps and video proof of the incident.

4. REFERENCES

1 https://en.wikipedia.org/wiki/Video_tracking

2 https://en.wikipedia.org/wiki/Machine_learning

3 https://en.wikipedia.org/wiki/Computer_vision

4 https://www.tensorflow.org/

5 https://opencv.org/

6 https://keras.io/

7 https://www.youtube.com/user/Mhproductionhouse

8 https://www.geeksforgeeks.org/opencv-python-tutorial/

9 https://www.geeksforgeeks.org/opencv-python-tutorial/

10https://www.analyticsvidhya.com/blog/2020/04/vehicle-detection-opencv-python/

11https://debuggercafe.com/moving-object-detection-using-frame-differencing-with-opencv/

12https://stackoverflow.com/questions/34615331/opencv-background-subtraction