International Journal of Current Engineering and Technology
E-ISSN 2277 – 4106, P-ISSN 2347 – 5161

Available at http://inpressco.com/category/ijcet

Research Article

Real Time Person Tracking and Re-Identification using Deep Learning

Mr. Baravkar Eknath Ashok. PG Student and Prof. Rajpure Amol S. Assistant

Department of Computer Engineering, Dattakala Group of Institution Faculty of Engineering, Bhigwan, India

Received 10 Nov 2020, Accepted 10 Dec 2020, Available online 01 Feb 2021, **Special Issue-8 (Feb 2021)**

Abstract

With the unstable development of video information and the quickest improvement of PC vision technology, there are more and more relevant advancements are applied in our real life, one of which innovations are re-identification (Re-ID) technology. Object Re-ID is currently moved in the field of person Re-ID, which is basically used to realize the cross vision tracking of person and direction prediction. Video reconnaissance systems are of great incentive for open well-being. A significant trouble in such systems concerns person re-identification which is characterized as the issue of identifying persons across images that have been captured by different observation cameras without covering fields of view. With the increasing requirement for robotized video examination, this undertaking is receiving increasing consideration. What's more, it supports numerous basic applications, for example, cross camera tracking, multicamera conduct investigation and forensic pursuit. Notwithstanding, this issue is trying because of the enormous varieties of lighting, pose, perspective and foundation just as the closeness in nature like individuals with identical features, color or clothes. To handle these different challenges, right now, propose a few deep learning-based ways to deal with get a superior person reidentification execution in different manners. Person ReIdentification has become unmistakable due to different reasons significantly because of its elite techniques dependent on deeplearning. It is the procedure of person recognition from different images captured by different cameras. Images are taken from different edges and separations of a given subject so as to accomplish high exactness, so it identifies correctly. Provided two arrangement of images the purpose is to find that the given arrangement of images is identical or not. This section consolidates hypothesis and practice to clarify why the deep system can re-identify the person. To present the primary specialized course of object Re-ID, the instances of person Re-ID are given.

Keywords: *deep learning, object recognition, person location and tracking, person re-identification, feature extraction, characterizations.*

Introduction

Person tracking and re-identification is valuable for security purpose. It is useful to identifies an obscure person from swarmed area. Person tracking and re-identification systems are PC vision security systems that are ready to consequently recognized and identify a person. These systems rely upon a tracking and detection algorithms of deep learning. The initial step for a Person tracking and re- identification system is to recognize a person and concentrate it from the rest of the scene.

There are a few reasons that cause real-time person tracking and reidentification systems to come up short, for example, pose change, expression varieties, age changes, impediments, and so forth. Subsequently it is hard to get a handle on the most exceptional degree of technology. Varieties in pose of a similar client starting with one system then onto the next can't be avoided.

In a surveillance camera without covering vision, a recognized object is identified again in the wake of imaging conditions (counting observing scene, lighting conditions, object pose, object area and so on. [1]) change, which is called object re-identification (Object Re-ID). Object Re-ID technology has significant research noteworthiness in intelligent checking, multi-object tracking and other fields. In recent years, researchers have paid broad regard for it. The principle application areas of object Re-ID are person Re-ID [2], [3].

Person re-identification (Re-ID) is a technology that utilization PC vision technology to judge whether there is a particular person in the image or video succession. It is widely regarded as a sub-issue of image retrieval. Given a screen person image, retrieve the image of the line of people over the gadget. It intends to compensate for the visual impediments of the current fixed cameras, and can be joined with person recognition and walker following innovation, which can be widely utilized in intelligent video observing, intelligent security and other fields [4], [5].

Person re-identification comprises in coordinating images of a specific person captured in a system of cameras with non-covering fields of view. The task is

different from the great identification and detection tasks. The identification comprises in deciding the identity of a person in an image and the detection comprises in segregating people from the foundation without knowing the identity. Reidentification responds to the inquiry whether a given image has a place with a similar person as a question image. The identification task encourages us to realize it is and the detection task shows whether it is a person. In any case, the re-identification tells when and where this person appeared with respect to a given camera and, utilizing a few cameras, possibly takes into consideration the estimation of his/her direction over a brief time frame. In the network of present-day PC vision, person re-identification is to see if different images capture from numerous cameras are have a place with a similar subject or not.

In video surveillance, person re-identification is to judge whether questioned person show up in the perspective on another camera, set at different area or at different time [1]. The presentation of other applications, for example, examination of conduct, retrieval of an object and tracking of cross camera are firmly related to the exhibition of person re identification. In recent years headway in the representation of tests and in the assessment measurements, which assess closeness between tests, greatly increases the exhibition of person re-identification however because of debasement and variety in person images make it more testing. There exist different survey papers, reviewing about person re-identification. This survey paper represents the deep learning-based methodologies for person reidentification by characterizing them as indicated by their architecture and furthermore as indicated by the correspondence built up between the inquiry set and the info set. In the methodologies dependent on deep learning, feature learning happens naturally from information together with other segments, this greatly improves the exhibition.

Implement a real-time person tracking and re-identification system which can recognize in four sections **detect**, **track**, **identified** object and then **re-identified** that object/person at any time. To achieve that task, we have needed some following objectives such as -

- To prepare a model we need an adequate dataset with proper labeling.
- We need to detect and locate an object/person area, for example, tracking an object/person movement.
- After that utilizing Deep Learning Algorithm Identifies that object/person with the assistance of the dataset.
- In the future that object/person goes under the camera, we can ReIdentified that object/person anytime.

Literature Survey

With the improvement in multi-object identification, inquire about network has begun concentrating on following of each and every item in various situations. The total MOT (movement) issue can be considered as an affiliation issue in which the essential target is to relate the identified items [6]. Following is completed after article recognition utilizing some item locator and afterward re-recognized that object. Right now, will concentrate on the foundation of the accompanying frameworks:

A. Object detection algorithms in past years before person trackingIn the mid-1990s, object discovery was completed utilizing format coordinating based calculations [7], where a layout of the particular item is slid over the info picture to locate the most In the late 1990s, the center was moved toward the geometric appearance-based item location [8], [9]. In these techniques, the essential spotlight was on height, width, points, and other geometric properties.

During the 2000s, object recognition worldview was moved to lowlevel highlights dependent on some measurable classifiers, for example, neighborhood paired example (LBP) [10], histogram of arranged inclination [11], scale-invariant element change [12], and co-variance [13]. Highlight extraction-based item discovery and order included preparing of machine dependent on separated highlights.

For a long time in PC vision field, carefully assembled customary highlights were utilized for object location. Be that as it may, with the advancement in profound learning in the wake of achieving the amazing exhibition in 2012 picture arrangement challenge [14], convolution neural systems are being utilized for this reason. After the accomplishment of article arrangement in, scientists moved their considerations toward object recognition and characterization [14]. Profound convolution neural systems work uncommonly useful for extraction of neighborhood and worldwide highlights regarding edges, surface, and appearance.

As of late, the exploration network has moved toward locale-based systems for object recognition. This kind of item discovery is being utilized in various applications like video depiction [15]. In localebased calculations for object discovery, convolution highlights are removed over proposed districts followed by classification of the area into a particular class.

With the alluring exhibition of Alex Net [14], Girshick et al [16]. proposed object recognition utilizing convolution neural system. They utilized specific quest for proposing the zones where the potential items can be found [17]. They called their article identification organize as area convolution neural system (R-CNN).

Redmon et al [18]. proposed You Only Look Once (YOLO) for object location. They totally dropped the district proposition step; YOLO parts the total picture into networks and predicts the recognition on the bases of applicant areas. YOLO separates the total picture into S x S grids. Every matrix has a class

likelihood C, B as the bouncing box areas and a likelihood for each case. Expelling the RPN step improves the exhibition of the identification; YOLO can identify the articles while running progressively with around 45 fps.

B. Multi-object tracking-

Videos are really sequence of pictures, every one of which called an frame, showed in quick enough recurrence with the goal that human eyes can percept the continuity of its content. Clearly all image processing systems can be applied to individual frames. Also, the contents of two back to back frames are typically firmly related. Visual content can be demonstrated as a chain of importance of reflections. At the principal level are the raw pixels with color or brightness data. Further preparing yields highlights, for example, edges, corners, lines, curves, and color regions.

Object detection in videos includes checking the closeness of an object in picture groupings and potentially locating it accurately for recognition. Object tracking is to screen an objects spatial and worldly changes during a video sequence, including its presence, position, size, shape, and so forth. This is finished by taking care of the transient correspondence issue, the issue of coordinating the objective district in progressive casings of a succession of pictures taken at closely space time interval. These two procedures are firmly related in light of the fact that tracking generally begins with identifying objects, while detecting an object over and again in subsequent picture sequence is frequently important to help and check tracking [19]. Storage and retrieval of transmission has become a demand for several up to date data systems. These systems got to give browsing, querying, navigation, and, sometimes, composition capabilities involving numerous varieties of media. during this survey, we tend to review techniques and systems for image and video retrieval. we tend to 1st check out visual options for image retrieval like color, texture, shape, and abstraction relationships. The categorization techniques square measure mentioned for these options. invisible options embrace captions, annotations, relative attributes, and structural descriptions. Temporal aspects of video retrieval and video segmentation square measure mentioned next. we tend to review many systems for image and video retrieval as well as analysis, commercial, and World Wide Web-based systems. we tend to conclude with an outline of current challenges and future trends for image and video retrieval [20].

Numerous scientists have concentrated on development and spatial highlights for following the different object [21,22]. A portion of the analysts have concentrated on appearance highlights for catching the relationship between various object [23], [24].

There are some customary strategies that make forecast on outline by-outline premise. These customary methodologies include different speculation following (MHT) [25] and joint likelihood information affiliation channel (JPDAF) [26]. Both of these old procedures require a great deal of calculation for following the recognized objects. The unpredictability of these procedures increments exponentially with expanding the quantity of identifiable articles that makes them truly delayed to be utilized for online applications in complex condition. In JPDAF speculation of single state is created dependent on connection between singular estimation and affiliation probability. In MHT, a total arrangement of theories is thought about for following followed by post pruning for tractability.

Rezatofighi et al [27]. put forth an attempt to improve the JPDAF execution by giving estimation of JPDA. They abused ongoing headway in understanding m-best answer for a number program. The primary favorable position of this framework is to make JPDA not so much intricate but rather more tractable. They reclassified the technique for figuring individual JPDAF task as far as an answer for a direct program. Another gathering of scientists Kim et al [24] utilized appearance-based highlights for following the objective. They improved the MHT by pruning the diagram of MHT for accomplishing best in class execution. They utilized regularized least squares for expanding the productivity of the MHT procedure.

C. Deep learning-based identification-

Headway in convolution neural systems has accomplished noteworthy execution by expanding precision and proficiency. The exceptionally essential presumption in profound neural systems is to encourage however much information as could reasonably be expected for improving outcomes. Prerequisite of tremendous information makes profound learning-based methodologies information hungry.

Lu et al [28]. actualized a lingering system (ResNet)-based model for face acknowledgment. They partitioned their total system into three systems: one spine organizes called trunk system and two different systems called branch arranges that transmit from trunk organize. The focal system is prepared once for learning the profound highlights for face distinguishing proof. The focal system is produced utilizing leftover squares. Goals explicit coupled mapping is utilized in branch arrange for preparing. Information picture and correlation picture from display are changed to same portrayal for contrasting. In light of separation the choice is made about recognized face.

An examination bunch from Facebook, Taigman et al [29], built up a condition of the art framework for face arrangement and face acknowledgment, named as Deep Face. They utilized profound convolution neural system having nine convolution layers for removing facial highlights. Facial milestones are utilized in their framework for face arrangement. The facial milestones are assessed utilizing bolster vector regressor (SVR). Separated highlights from nine-layered system are passed to Soft-max layer for characterization. They

utilized cross-entropy to lessen the loss of right marks. They likewise proposed an enormous face acknowledgment dataset named as Social Face Dataset [29]. They utilized their dataset for preparing the framework for face ID.

We are referring a great deal of papers as appeared over, these paper shows a ton of people works on various pieces of project task related points like detection of object, tracking or capture a movement of object as well as identification of that object.

Be that as it may, in this paper we are chips away at ReIdentification of steady just as moving object/Person is our difficult works.

I. **General Process Of Person Detection**

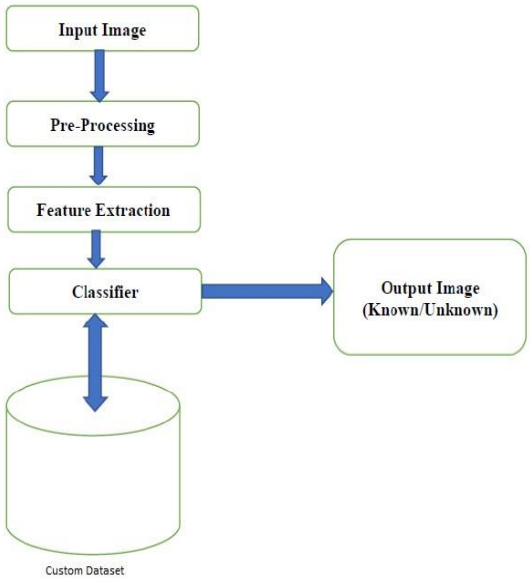


Fig 1. Block diagram of person detection/Re-Identification

Person Locating-
Person Locating is the task of finding all cases of human beings present in an image, and it has been most widely achieved via looking through all areas in the image, at all potential scales, and contrasting a little area at every area with known layouts or examples of people.

Person Tracking-

Person tracking is the process of temporally associating the person detections within a video sequence to generate persistent paths, or trajectories, of the people.
Human detection and following are commonly considered the initial two procedures in a video surveillance pipeline, and can nourish into higher-level reasoning modules, for example, activity recognition and dynamic scene investigation.

Feature Extraction and Transform-
Detection and extraction the two techniques are completed at the same time. There are different variables that makes the person detection is a difficult task. Pose presence or nonattendance of auxiliary segments, impediment, image direction. The person feature detection is the procedure to recognize the presence and area of features.

Database-

We can store annoted images for a model preparing utilizing that information we prepared a model over and over up to show signs of improvement precision of distinguishing a person. Lastly get a real time identified person.

Person recognition has been a difficult research issue for PC vision researchers for a long time. A variety of this conventional issue is that of identifying the reappearance of a similar person in different fragments. We may have video encourages from various cameras with in part covering perspectives working under widely fluctuating lighting and perceivability conditions. However, PC vision systems are tested to tracking people across cameras to scanning for them in a huge display, from gathering photographs in a photograph collection to guest examination in a retail store. In the same way as other visual recognition issues, varieties in pose, variety in areas, perspectives, enlightenment, and impediment make this issue non-minor then we center around the different person reidentification structures based on deep learning and talk about their preferences and impediments. At last, we propose the direction of further research, particularly the possibility of person re identification strategies dependent on deep learning. Aim of this paper is to detect, track movement of object/person and re-identified that object/person with better accuracy and without losses.

Methodology

For person detection and Re-identification, we used a following pipeline diagram:

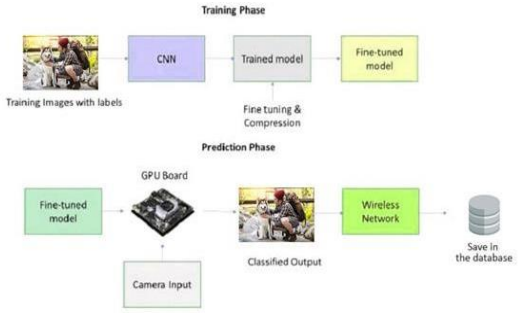


Fig. 2 Pipeline diagram for Person tracking re-identifications.

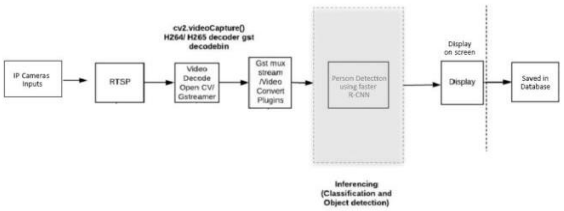


Fig. 3 Block diagram of Person Detection/Re-Identification Pipeline.

Implementation

A. Dataset Collection-

In Machine Learning projects, we need a training dataset. It is the genuine dataset used to train the model for performing different activities. ML relies intensely upon information, without information, it is incomprehensible for an "AI" to learn. It is the most urgent perspective that makes algorithm training conceivable. Regardless of how great your AI group is or the size of your informational index, if your informational collection isn't sufficient, your entire AI undertaking will fail! I have seen fabulous projects fail in light of the fact that we didn't have a decent informational collection regardless of having the ideal use case and talented information researchers.

During an AI development, we generally rely on information. From training, tuning, model determination to testing, we utilize three different informational indexes: the training set, the validation set, and the testing set. The present AI showcase part of sorts of Open Source datasets are available, for example, Object detection, Vehicle detection, Licenses number detection, vehicle number plate detection, Gender detection, Face detection, Sentiments detection, pose estimations, Person detection and the sky is the limit from there...

Here we are used only Person detection, Re-Id and tracking datasets.
There are lot of open source Datasets for it which are follows:
Market1501 [34], CUHK03 [33], VIPeR [31], PRID2011 [32], CUHK-SYSU [35], MARS [36], DukeMTMC-reID [37]. In addition to the common data sets that are already open source. There are a few more up to date data sets, for example, SYSU MM01 [38], LPW [39], MSMT17 [40], LVreID [41] the download interface isn't yet open.

There a lot of open source datasets are available, but for our Project we are used a self-made custom dataset because to prevent losses as well as increase an accuracy of model. (images of custom dataset) To create a custom dataset we are used four cameras all these are sets at different-different position and distance so we are able to capture images from different angles as well as distances.

B. Annotation-

For mode training need to utilize clean, annotated data to train machine learning models. Data annotation is a crucial phase of data prehandling in regulated learning. Machine learning models figure out how to recognize recurring examples in the annotated data. After an algorithm has prepared enough annotated data, it can begin to recognize similar examples when presented with new, unannotated data. Annotation is the craft of encoding meaning onto crude data. This is utilized in managed learning to construct a machine learning models. Raw data includes:

- Images, screenshots, documents, pdfs, Videos, Audio files, Point clouds (for example from a LiDAR system) When we are gathering a dataset after that we need to name or annot that datasets with different-different classes. For annotation there are part of annotation tolls are available in market, for example, LabelImg, CVAT and more however we are utilized LabelImg for annotation.

Encoded significance incorporates: labelImg software is utilized. Annotation xml document to show Attached a "class, for example, "Person", "car", "green marker", "duck" just as we are create names of that person as a class. With a visual annotation the area may likewise be represented with:

- Bounding boxes
- For instance, a pair of points "x min, y min" and "x max, y max"
- Polygons (any number of points)
- Pixel wise (Similar to polygons)
- Cuboids (3-dimensional box)
- Key points

Here we are utilized a LabelImg Tool for annotation.

C. Model Selection Approaches-

For a person tracking and Re-identification there are a ton of models are available however we are contemplated some deep learning algorithms and select reasonable algorithms to distinguish, attaching and re-identify the person. We can see that are as per the following:

a) Faster RCNN-

It Is a Region based Object detectors? In this way, what did Faster RCNN improve? Well, it's faster than RCNN and Fast RCNN. Furthermore, how can it accomplish that? Faster RCNN replaces specific pursuit with an extremely little convolution network called Region Proposal Network (RPN) to create regions of Interests. Faster RCNN is divided in 3 sections:

1. Convolution Neural Network (CNN)
2. Region Proposal Network (RPN) 3. Faster R-CNN Object detection pipeline.

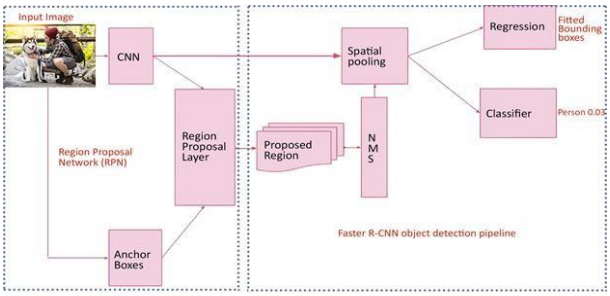


Fig. 4 Architecture of Faster R-CNN

To deal with the varieties in perspective proportion and size of objects, Faster R-CNN presents the idea of stay boxes. At every area, the first paper utilizes 3 sorts of grapple boxes for scale 128x128, 256256 and 512512. Thus, for viewpoint proportion, it utilizes three perspective proportions 1:1, 2:1 and 1:2. In this way, altogether at every area, we have 9 boxes on which RPN predicts its likelihood being foundation or foreground. We apply bounding box regression to improve the stay boxes at every area. Along these lines, RPN gives out bounding boxes of different sizes with the corresponding probabilities of each class. The differing sizes of jumping boxes can be passed further by apply Spatial Pooling simply like Fast-RCNN. The remaining system is like Fast-RCNN. Faster-RCNN is 10 times faster than Fast-RCNN with comparative exactness of datasets like VOC-2007. In any case, here we are utilized our custom dataset for high exactness. That is the reason Faster-RCNN has been one of the most precise object detection algorithms. Here is a speedy examination between different renditions of RCNN. Yet, why we see also focus on Faster RCNN rather than RCNN or Fast RCNN that is reason cleared in following table:

Table 1. Difference between R-CNN, Fast R-CNN & Faster R-CNN

Parameters	RCNN	Fast RCNN	Faster RCNN
Test Time Per Image	50 Seconds	2 Seconds	0.2 Seconds
Speed Up	1x	25x	250x

b) YOLO (You Only Look Once)-

For YOLO, detection is a basic regression issue which takes an info image and learns the class probabilities and bounding box arranges. Sounds basic? YOLO divides each image into a grid of S x S and every grid predicts N bounding boxes and confidence. The confidence reflects the exactness of the bounding box and whether the bounding box really contains an object (regardless of class). YOLO likewise predicts the classification score for each container for each class in training. You can join both the classes to figure the likelihood of each class being present in a predicted box. Thus, all out SxSxN boxes are predicted. Notwithstanding, a large portion of these boxes have low confidence scores and on the off chance that we set a threshold state 30% confidence, we can remove the greater part of them as appeared in the model underneath.

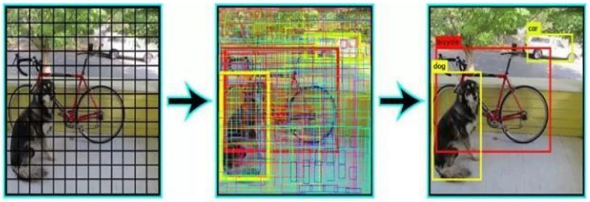


Fig. 5 Object detection using YOLO V3 algorithm.

Notice that at run-time, we have run our image on CNN only once. Henceforth, YOLO is super-quick and can be run real time. Another key difference is that YOLO sees the total image at once as opposed to looking at only a produced region proposal in the previous strategies. Thus, this logical data helps in avoiding bogus positives. Notwithstanding, one impediment for YOLO is that it only predicts 1 kind of class in one grid thus, it battles with little objects.

c) SSD v2-

Single Shot Detector accomplishes a decent harmony among speed and accuracy. SSD runs a convolutional network on input image only once and ascertains a feature map. Presently, we run a little 3X3 estimated convolutional kernel on this feature guide to predict the bounding boxes and classification likelihood. SSD likewise utilizes grapple boxes at different perspective proportion like FasterRCNN and learns the off-set rather than learning the crate. So as to deal with the scale, SSD predicts bounding boxes after different convolutional layers. Since each convolutional layer works at a different scale, it can recognize objects of different scales. The technique proposed here is divided into 2 main parts:

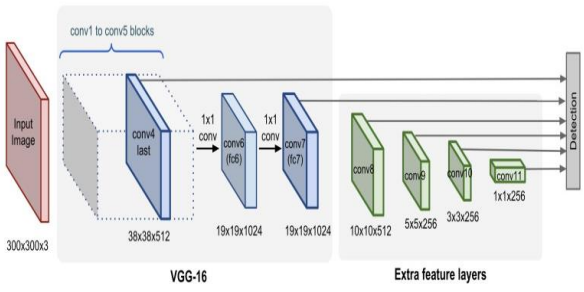


Fig. 6 Architecture of SSD V2

Person Detection-

The person detection in Real-Time is finished with the assistance of Single Shot MultiBox Detector. SSD accomplishes 75.1% mAP, outflanking a practically

identical best in class Faster R-CNN model. What's more, the SSD model is available in the TensorFlow detection zoo. The consistent combination of SSD with TensorFlow aides in further advancement and usage of the algorithm.
The SSD object detection composes of 2 sections:
(I) Extract feature maps,
(ii) Apply convolution channels to recognize objects. Despite the fact that SSD is fit for distinguishing numerous objects in the casing, right now are restricted its detection to only a person.

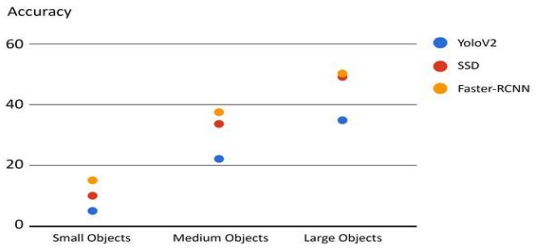
Person Tracking-
Bounding box can be accomplished around the object(person) by running the Object Detection Model in each casing, however this is computationally costly. The following algorithm utilized here is Kalman Filtering. The Kalman Filter has for quite some time been regarded as the ideal answer for some following and data prediction tasks. Its utilization in the investigation of visual movement. The purpose of separating is to extricate the required data from a sign, disregarding everything else. Right now, Kalman Filter is sustained with the speed, position and bearing of the individual which encourages it to predict the future area of the Person dependent on his previous data.

Proposed Approach

There are a ton of algorithms. Which one would it be fitting for you to use? Currently, Faster-RCNN is the decision on the off chance that you are fan about the accuracy numbers. In any case, in the event that you are lashed for calculation (presumably running it on Nvidia Jetsons), SSD is a superior recommendation. At long last, if accuracy isn't an over the top concern however you need to go super-quick, YOLO will be the best approach. As a matter of first importance, a visual comprehension of speed versus accuracy exchange off:

Graph 1. speed vs accuracy of SSD, YOLO V3 and Faster R-CNN.

In our Project we need a higher accuracy of result speed has not make any difference right now of in person following and ReIdentification case person identification accuracy is significant than speed. Along these lines, we are select a Faster R-CNN algorithm.



Graph 2. YOLO vs SSD vs Faster R-CNN for various size of objects.

Decision of a correct object detection strategy is vital and relies upon the issue you are attempting to explain and the set-up. Object Detection is the backbone of numerous viable uses of PC vision, for example, autonomous cars, security and surveillance, and numerous modern applications. Ideally, this post gave you an instinct and comprehension behind every one of the well-known algorithms for object detection. At last, we are decided to utilized Faster R-CNN algorithm based on accuracy.

Model Training

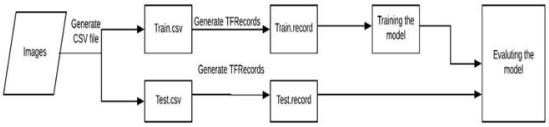
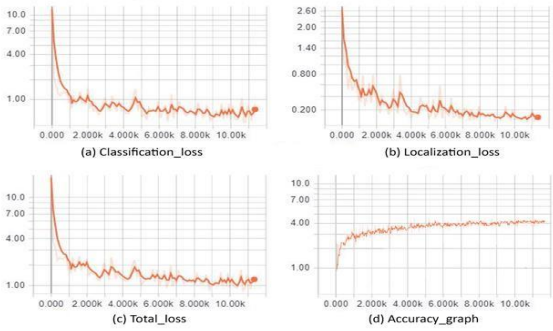


Fig. 7 Basic flow diagram of Model Training

For model training we are use train.py record, which is in the object detection/inheritance envelope. We will duplicate it into the object detection envelope and afterward we will open an order line and type: python model main.pylogtostderr–model dir=training/pipelineconfig-path=training/faster rcnn origin v2 persons.config Then on terminal model has trained and About at regular intervals the current misfortune gets logged to Tensorboard. We can open Tensorboard by opening a subsequent direction line, exploring to the object detection organizer and composing:



Graph 3. Losses and Accuracy for model train.

Above graph shows classification loss, localization loss and total loss.

Classification loss:
You will have instantly recognized it – it's a person. Take a step back and analyze how you come to this conclusion- you were shown an image and you classified the class it belonged to (a person, in this instance). What's more, that, more or less, is the thing that image classification is about. Here we have two class such as person and background compared both and assign a proper class. In that case there are some losses as shown in graph fig (a). Model not accurately classified an object that time losses occurs. At initial stage rate of losses is higher but we are trained model again then losses will be slightly decreasing so accuracy will be increased.

Localization loss:
Regression is tied in with returning a number rather than a class, for our situation we're going to return 4 numbers (x0, y0, width, stature) that are related to a bounding box. You train this system with an image and a ground truth bounding box, and use L2 distance to calculate the loss between the predicted bounding box and the ground truth. We should train the model until it reaches a satisfying loss.

Total loss:
Combination of Classification and Localization loss.

Result

For a person Re-Id, first we need to identify a steady too moving person (individual person as well as multiple person detect in single frame) and after that Re-Identified that Person.
We can Identified multiple object with person category in a single frame as shown in following figure.

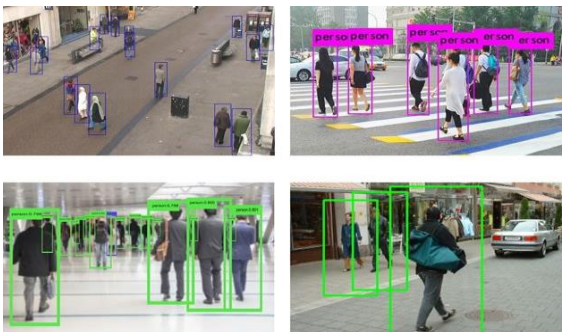


Fig. 8 Multiple person detection result.

In this paper we are done a tracking and detection of person. As well as Re-identify a person by name which are already labelled and stored in our database. By using this Dataset, we can Identify that person those are comes under our camera range. As well any unknown person detected in a frame this person detects as unknown person labelling with bounding box. Final result shows in following Fig. 9

Fig.9 Person Re-Identification Result

Conclusion

An accurate and efficient person detection and re-identification system has been designed which achieves comparable metrics with the existing state-of-the-art system. Custom dataset was created using labelImg and the evaluation was consistent. This can be used in real-time applications which require person detection for preprocessing in their pipeline.
In this project we are studied a lot of algorithms such as SSD, YOLO V3 and Faster RCNN.
We do comparative study of that algorithms and choose Faster RCNN algorithm for Person tracking and detection for the better accuracy purpose.
Finally, we are Re-Identified the persons successfully.

References

[1] Chen, Yiqiang & Duffner, Stefan & Stoian, Andrei & Dufour, Jean-Yves & Baskurt, Atilla. (2018). Deep and Low-level Feature based Attribute Learning for Person Re-identification. Image and Vision Computing. 79. 10.1016/j.imavis.2018.09.001.
[2] Sokolova, A & Savchenko, A. (2018). Data organization in video surveillance systems using deep learning. 243-250. 10.18287/16130073-2018-2210-243-250.
[3] Chellappa R, Du M, Turaga P and Zhou S K 2011 Face Tracking and Recognition in Video Handbook of Face Recognition 323-351
[4] Htun, Baby & Sein, Myint. (2017). Observation of Unattended or Removed Object in Public Area for Security Monitoring System.
536. 45-53. 10.1007/978-3-319-48490-7_6
[5] Xiying Li and Zhihao Zhou (June 23rd 2019). Object ReIdentification Based on Deep Learning, Visual Object Tracking with Deep Neural Networks, Pier Luigi Mazzeo, Srinivasan Ramakrishnan and PaoloSpagnolo, IntechOpen, DOI: 10.5772/intechopen.86564. Available from:
[6] Gulraiz Khan, Zeeshan Tariq and Muhammad Usman Ghani
Khan (May 20th 2019). Multi-Person Tracking Based on Faster RCNN and Deep Appearance Features, Visual Object Tracking with Deep Neural Networks, Pier Luigi Mazzeo, Srinivasan