Sapio–Enumerator Using Machine Learning

Nikhil Singh[1], Vivek Singh[1], Manthan Thadve[1], Diwakar Jha[1], Mrs Sarita Patil[2]

[1] *UG Students,* [2] *Assistant Professors, Computer Engineering Department,*
G H Raisoni College of Engineering and Management, Pune

ABSTRACT

Sapio Enumerator based on video is an important field in Computer Vision. When Sapio Enumerator is implemented it may accompany by some problems. This paper proposed a method that may mix some existing technologies to overcome some problems. The Sapio Enumerator requires more powerful processing since it deals with real-time video, so the proposed method converts a color image into binary to minimize data of an image. Also, the image may contain a noise, the proposed method uses Erosion and Dilation processes to erase the noises. Reducing development time is an important term in Software Engineering to build an application. The method depends on existing packages for reducing development time. The proposed method is implemented in the Python programming language.

Keywords— Android application, image processing, cloud storage, motion detection, people detection, computer vision, pixel.

I. INTRODUCTION

A Sapio-Enumerator is an electronic device that is used to measure the number of people traversing a certain passage or entrance. Examples include simple manual clickers, smart-flooring technologies, infrared beams, thermal imaging systems, WiFi trackers, and video counters using advanced machine learning algorithms. They are commonly used by retail establishments to judge the effectiveness of marketing campaigns, building design and layout, and the popularity of particular brands.

Supervised Learning:

In Supervised Learning we have input variables (x) and output variables (y). Using algorithm a function is map

$$y = f(x)$$

The goal of a Supervised Learning algorithm is to map function so well that if you introduce any new (x) variable, it can predict (y) perfectly. Thus from the above example, we conclude that Supervised Learning is a method of ML in which a set of label data is used to train the model and predict the outcome. Examples of Supervised Learning: Regression, Random Forest, Support Vector Machine

Unsupervised Learning:

Unsupervised Learning only has (x) in data. There are no output variables present in the dataset. It has to build a structure based on available data to learn more about that data. These are called Unsupervised Learning because unlike Supervised there is no correct answer available, only structures are formed to learn more about data. Examples: Clustering and Association

Reinforcement Learning:

Reinforcement Learning algorithms allow computers to learn from the experience. The machine trains itself so that when it predicts the correct one reward signals are generated and punishment for the wrong one.

Algorithms (can be) used:

Many Machine Learning algorithms are present that can be used for our needs. But the dataset matters in most of the cases to determine which algorithm we choose to predict outputs. Following are algorithms that can be used for the prediction of phishing websites.

Support Vector Machine(SVM):

Support Vector Machine is a type of Supervised Learning. This method is able to perform both Classification and Regression which are types of Supervised Learning. SVM is basically used for Classification problems. In SVM Classification the data is plotted on the graph. First, we have to plot a line between the data. We have to make sure that the space between the hyperplane (the line we plot) and Support Vector (Extreme points of data on a graph) is maximum. That's how using the Support Vector Machine we classify the data into two groups where we may be able to predict the desired output.

Random Forest :

The Random Forest algorithm is also used for Regression and Classification. But it is mainly used for Classification problems. In Random Forest multiple Decision Trees are created. The number of Decision Trees decides the accuracy of the Model. More the number of decision trees more is the accuracy of the model. Decision trees have the disadvantage of overfitting. Thus, it is overcome by Random Forest. Random Forests are used in business as they can make predictions over a wide variety of data with less configuration.

Tracking:

Once the people are located in each individual image, it is necessary to track them across images. This is achieved using a particle tracker developed by Crocker et.al, available online at http://www.deas.harvard.edu/projects/weitzlab/matlab/. For each image in the blob sequence obtained above, the centroids of the blobs are located. These points and their frame number are then fed into the particle tracking algorithm. The algorithm is based upon the IDL tracking algorithm developed by Crocker et.al [6]. To assign a tracking label to a blob, it essentially finds the closest particle in the previous image and assigns that track to the current case. Constraints are set in the maximum translation a particle can move. Also, the algorithm allows a particle to disappear for several frames and still be tracked. This accounts for quick occlusions when people pass in front of each other. Furthermore, if a particle appears for less than 2 frames, it is not tracked and assumed to be noise. This accounts for quick changes in lighting and shadows. See the results of the tracking algorithm in Figure 2. The algorithm works fairly well in putting together tracks for individual particles. It does get confused when people pass over each other slowly, and on some occasions when people move very quickly such that the displacement is large. However, for the most part, so long as a track is found it is sufficient to make a decision without maintaining identity.

II. LITERATURE SURVEY

This section of the literature survey eventually reveals some facts based on thoughtful analysis of many authors work as follows.

In this paper, a brief literature survey for people counting and human detection methods are discussed elaborately and neural network based people counting and EM based individual detection methods are studied The best results for estimating the number of people has an average error of 10% over 51 test cases. These methods provide better performance and high accuracy than existing methods. This people counting and human detection process is very useful for safety control in public areas by using static camera.

1. Crowd analysis: a survey, 2008, MVA: This paper presents a survey on crowd analysis methods employed in computer vision research and discusses perspectives from other research disciplines and how they can contribute to the computer vision approach.
2. Crowd analysis using computer vision techniques, 2010, ISPM: A survey on crowd analysis by using computer vision techniques, including different aspects such as people tracking, crowd density estimation, event detection, validation and simulation.
3. A Survey of Human-Sensing: Methods for Detecting Presence, Count, Location, Track, and Identity, 2010, ACM: Computing Surveys a survey of the inherently multidisciplinary literature of human-sensing, focusing mainly on the extraction of five commonly needed Spatio-

temporal properties: namely presence, count, location, track and identity.

4. Crowd counting and profiling: Methodology and evaluation [60] 2013 MSVAC This study describes and compares the state-of-the-art methods for video imagery based crowd counting, and provides a systematic evaluation of different methods using the same protocol.

.

III. PROPOSED SYSTEM

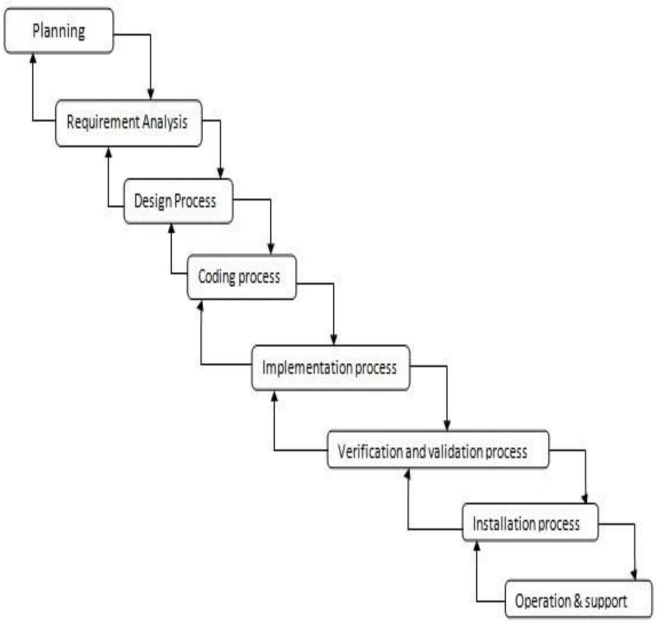

Figure 1. System Architecture Design.

The pre-processing step of importing the videos turns out to take the most time, however, this step is dependent upon platform and input data types. The step takes approximately 4-5 hours for a 45-minute video. The rest of the algorithm is implemented in Python and as quite unoptimized code. Overall it performs at approximately 2 frames/sec on a 1.6GHz Pentium M. By optimizing the code and porting it to C, a real-time implementation could easily be developed. Since this algorithm is run offline anyway, performance is not an important criterion. However, in the final implementation, it should run as fast as possible to reduce computational time.

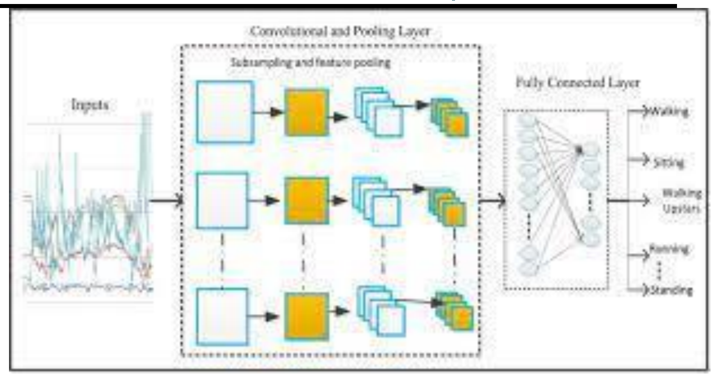

Figure 1. CNN Layer

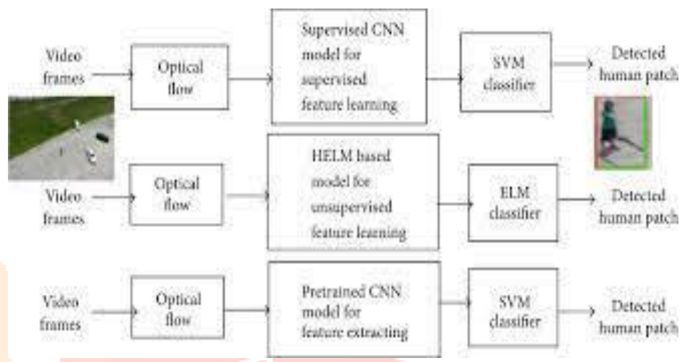

Figure 2.CNN based image processing

The scope thus far has been the completion of the basic interfaces that will be used to build the system. The main goal of this thesis is to contribute to the evaluation of the possibility to develop a technical solution that allows the use of surveillance cameras for counting people in public buildings. Our contribution is in the analyses of current real time computer vision methods to create a system capable of reliably count the number of persons in complex environments. In particular, we will focus our contribution in the development of a real time people counter in complex environments, using distributed perception systems

In this way, our objectives for this dissertation are:

• Analyze the applicability of current real-time computer vision methods as the solution to count the number of people that visit the building

- Test the implementation of different state of the art methods for real-time still camera background maintenance and people detection.

IV. Result & Discussion

Given the importance of full communication in human interaction, it can help you to have this knowledge in robotic systems where the robot can detect the motivations and emotions of a colleague. Recognizing the affected conditions from the body parts is an effective way to initiate anonymous interactions with robots. Several machine learning methods have been successfully used to affect recognition to predict the emotional state of a person given a set of physical features. However, a systematic comparison of the strengths and weaknesses of these methods has not yet been made. In this paper, we present a comparative study of four machine learning methods - the K-Nearest Neighbor, Regression Tree (RT), the Bayesian Network, and Support Vector Machine (SVM) as is used in the context of visual impact using body signals. The results showed that SVM provided excellent separation accuracy even though all methods were performed competitively. RT has provided excellent tracking accuracy and has been a great site and time-efficient. Tools make machine learning faster and faster. Machine learning tools provide a display in the language of machine learning. They provide good practice and implementation practices. Machine learning tools consist of platforms that provide the skills to use a module or project.

The relationship between the foreground pixels and the number of people becomes quite simple after the closing operation that is shown in Fig 4. Fig 4. (b) gives better performance than Fig 4. (a). This process produces high accuracy than existing methods.

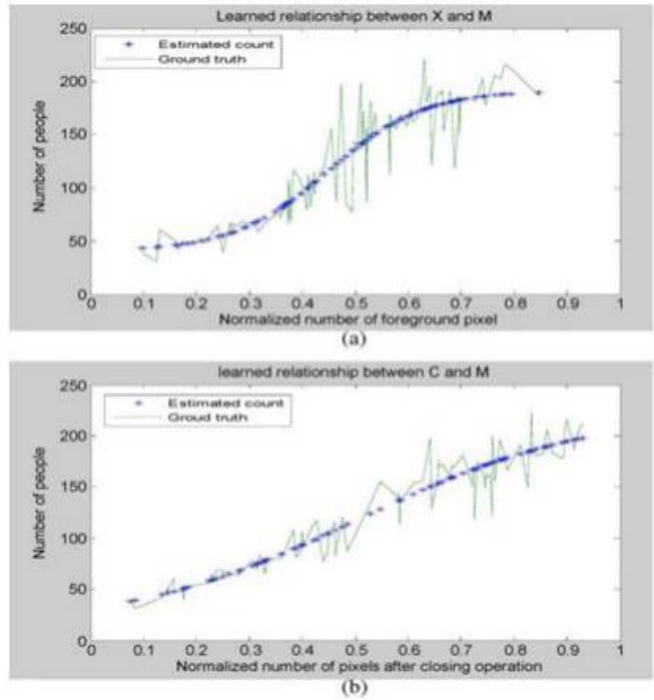

Figure 4. (a) Relationship between the number of people and the number of foreground pixels. (b) Learned relationship between the number of people and the number of pixels after the closing operation

V. CONCLUSION:

The proposed method provides faster processing for people counter application because it depends on a grayscale image which represents the image with lesser data than the colour one. Background Segmentation is a process uses to separate the background and foreground, but it isn't a sufficient process because the resulting image may contain noises. In addition to Background Segmentation, the proposed method uses the Object Detection method to detect people automatically. Background Segmentation isn't accurate when the feature of the background changes, therefore, the proposed method uses an accumulator image to be convenient with time changes.

References

[1]. T.Zhao and R.Nevatia, Tracking multiple humans in complex situations, IEEE Trans, Pattern Anal.Mach.Intell., vol.26, no.9, pp.1208-1221, Sep. 2004.

[2].T.Zhao,R.Nevatia,and B.Wu, Segmentation and tracking of multiple humans in crowded environments,IEEETrans.Pattern Anal.Mach.Intell.,vol30,no.7,pp.1198-1211,Jul.2008

[3].J.Rittscher,P.H.Tu, and N.Krahnstoever, Simultaneous estimation of segmentation and shape, in Proc,IEEE Cnf.Comput.vis .Pattern Recog.,2005,pp.486-493.

[4].A.Mohan,C.Papageorgiou,and T.Poggio, Example-based object detection in images by components, IEEE Trans.Pattern Anal.Mach.Intell.,vol23.no.4, pp.349-361,Apr.2001.

[5]. Ya-Li Hou, and Grantham K.H.Pang, People counting and human detection in a challenging situation, IEEE Trans.sys.man and cybernetics., vol.41, No.1, pp.24-33, Jan.2011.

[6]. N.Dalal and B.Triggs, Histograms of oriented gradients for human detection, in Proc.IEEE Conf.Comp.Vis.Pattern Recog, pp.886-893,2005.

[7]. B. Wu and R.Nevatia, Detection and tracking of multiple, partially occluded humans by bayesian combination of edgelet based part detectors, Int.J.Comput.Vis.,vol.75,no.2,pp.247-266, Nov.2007.

[8].G.J.Brostow and R.Cipolla, Unsupervised Bayesian detection of independent motion in crowds, in Proc.IEEE Conf.Comp.Vis.Pattern Recog, pp.594-601,2006.

[9].V.Rabaud and S.Belongie, Counting crowded moving objects, in Proc.IEEE Conf.Comput Vis.Pattern Recog.,pp.705-711,2006 Literature survey for people counting and human detection www.iosrjen.org 10 | P a g e

[10]. A.C.Davies, J.H.Yin, and S.A.Velastin, Crowd monitoring using image processing, Electron.Commun,Eng.J.,vol.7,no.1, pp. 37-47, Feb1995.

[11]. S-Y.Cho, T.W.S.Chow, and C-T.Leung, A neural-based crowd estimation by a hybrid global learning algorithm, IEEE Trans.Syst.Mcn Cybern.B.Cybern.,vol.29.no.4,pp.535-541, Aug.1999.

[12]. R.Ma,L.Li,W.Huang, and Q.Tian, On pixel count based crowd density estimation for visual surveillance, in Proc.IEEE conf.Cybern.Intell.Syst.pp.170-173,2004.

[13]. H.celik, A.Hanjalic, and E.A.Hendriks, Towards a robust solution to people counting, in Proc IEEE Int. Conf.Image Process.,pp.2401-2404,2006.

[14].P.Kilambi, O.Masoud, and N.Papanikolopoulos, Crowd analysis at mass transit site, in Proc.IEEE Intell. Transp. Syst.Conf.,pp.753-758, 2006.

[15]. A.N.Marana, S.A.Velastin, L.F.costa, and R.A.Lotufo, Estimation of crowd density using image processing, in Proc.IEEE colloq.Image Process.Security Appl.,pp.11/1-11/8,1997

[16]. A.N.Marana, L.Da Fontoura costa, R.A.Lotufo, and S.A.Velastin, Estimating crowd density with Minkowski fractal dimension, in Proc. Int. Conf. Acoust., speech, Signal Process.,pp.3521-3524,1999.

[17]. H.Rahmalan, M.S.Nixon, and J.N.Carter, On crowd density estimation for surveillance, in Proc. Inst. Eng. Technol. Conf. Crime security, pp.540-545, 2006.

[18]. X.Li, L.Shen, and H.Li, Estimation of crowd density based on wavelet and support vector machine,Trans. Inst. Meas.Control, vol.28, no.3, pp. 299-308, Aug. 2006.

[19].D.Kong, D.Gray, and T.Hat, A viewpoint invariant approach for crowd counting, in Proc. Int. Conf. Pattern Recog, pp. 1187-1190, 2006. [20]. A.B.Chan, Z.S.J.Liang, and N.Vasconcelos, Privacy preserving crowd monitoring: Counting people without people models or tracking , in Proc. IEEE Conf. Comput. Vis. Pattern Recog., pp.1-7, 2008.

[21]. W.E.L.Grimson,C.Stauffer, R.Romano, and L.Lee, Using adaptive tracking to classify and monitor activities in a site, in Proc. IEEE Conf. Comput. Vis.Pattern Recog, pp. 22-29, 1998.

[22]. C.Stauffer and W.E.L. Grimson, Learning patterns of activity using real-time tracking, IEEE Trans. Pattern Anal. Mach. Intell., vol.22, no.8, pp. 747-757, Aug.2000.