International Journal of Advanced Science and Technology
Vol. 29, No. 5, (2020), pp. 2791 - 2798

Prediction of Human Motion Detection in Video Surveillance Environment Using Tensor Flow

Manikandan S [1*], Chinnadurai M [2], Thiruvenkatasuresh M.P [3], Sivakumar M [4]

1 Assistant Professor, Department of Information Technology, E.G.S. Pillay Engineering College, Nagapattinam, Tamil Nadu, India
2 Professor, Department of Computer Science and Engineering, E.G.S. Pillay Engineering College, Nagapattinam, Tamil Nadu, India
3 Professor, Department of Computer Science and Engineering, Erode Sengunthar Engineering College, Perundurai, Erode, Tamil Nadu, India
4 Assistant Professor, Department of Computer Science and Engineering, Erode Sengunthar Engineering College, Perundurai, Erode, Tamil Nadu, India
*Corresponding Author *: profmaninvp@gmail.com*

Abstract

In recent year, Deep learning concepts are used to investigate image processing and machine learning applications. As the same experience we use deep learning prediction method for finding moving object and human motion in video surveillance environment. Deep learning approach is used to detect, classify and find the moving object and classify the person in captured videos. For finding human, a set of image processing steps are applied in moving areas that contain moving objects. The selected area are applied convolutional neural network classifier which has various layers and available in Google Tensor Flow learning tool. In this approach, shaking videos, low resolution videos are taken into account and video dataset examined. The experimental studies says that the relationship between moving object and video surveillance dataset has been analyzed and prediction performance has noted as 85% accuracy and from 65 videos.

Keywords: *Deep Learning, Human Motion Detection, Convolutional Neural Network Classifier, Google Tensor Flow, Image Processing.*

1. Introduction

Nowadays most of the places video surveillance systems are used for monitoring and security purpose. Government sector also used for different areas like traffic control, security enablers, banking sectors, etc. According to the April'2019 Times of India report says that in future India most of places surveillance systems are enabled for all the monitoring and support. There are many video surveillance system are implemented for both commercial and non commercial applications such as parking, theft monitoring, retail shops, malls and cyber crimes. Laszewski et al, CCTV surveillance systems enabled for all the places and reduced human man power [1].

Image processing algorithms are developed for detecting objects from videos. Deep learning image processing methods are experienced in growth and various convolution approaches are used for specific applications. In this paper, we applied deep learning approach for detecting human from motion object in video surveillance systems [4]. More specifically the methods are used in border areas and identify very far distance objects. Kehtarnavaz et al, the human object detection or moving object monitoring system is developed and detecting human are reported each deep learning layers [2].

Adaboost classifier approach is used for detecting motion objects with the combination of histogram gradients and support vector machine features [8]. The region based convolution neural network and region based convolutional neural network are used human detection. In this paper human detection and classification techniques are applied video stream data. The object is identified using edge detection techniques and it is computationally fast calculation process for evaluation moving areas. The classification stage is a learning system which is used to find human without hold anything, human holding some objects and human with disabilities [3].

Machine learning has very become and famous approach for computational learning experiments and decision making. Statistical learning and neural networks are the approaches used for solving modelling, prediction and decision making problems [10]. Machine learning provides various methods and approaches to extract information from various data streams such as video, audio and text. It is observed and created model for predicting future and extracting information from huge data set. Lluis et al, the behaviour model are used for discovering attributes, properties and operations of each objects and a-priori methods are used for prediction and classification. Machine learning is an interesting model for infeasible and unsolved problems using exhaustive computing strategies[4].

In this paper, we used statistical machine learning approach used for detection of moving objects and deep learning approach for classification and human prediction. Section II deals with various statistical and deep learning literatures. Section III discusses system and implementation process and final section for result and discussions.

2. Learning Techniques

2.1 Statistical Learning

Machine learning techniques are applied based on statistical approach, prediction, clustering, classification and decision making. A-priori model is used for trained the data set and creating model. Supervised and Unsupervised learning are used for data streaming problems. Supervised learning is observation based approach and outputs are calculated by polynomial and linear approach. Tree, Forests, Space and Neighbour, Bayesian methods are used for supervised learning calculations. Unsupervised learning is discovering data relations and models without specifying outputs. Clustering and Dimensionality reduction techniques are applied for unsupervised learning calculations[5].

2.2 Reinforcement Learning

Reinforcement learning methods are combination of statistical learning and Q-Learning approaches. This is trial and error model with adjusting feedback based measurement procedures. The classical use of the learning is making effective decision and maximizing utilization factor [6]. Online learning is recent model with combining the entire above learning feature for Big Data and Data Analytics. This model has updating buffering and streaming data as input and predicts the future for next level [9].

2.3 Deep Learning

Neural Network is combination of all learning techniques which is used for predicting and classifying future. Artificial Neural Network are a model which collects connected neurons as input and apply regression for resulted output neurons [12]. Convolution techniques are used for image processing techniques. A neuron has perceptron with the input function f(x) and output function g(x)

where x refers neuron. The linear sigmoid function is applied for calculating error and adjusting input parameters. The quality factor can be calculated

$$Q(x) <- \text{sigmoid}(f(x),\text{sigmoid}(f(x-1))$$

Training perception is required in each stages, Feed Forward are used by arrays of storage calculations, Convolution for hidden layers, Recurrent for behaviours and Auto encoders for dimensionality reductions. The Boltzmann machine is input each layer results and new feature can be activated for expanding dimensions and new features results.

3. Video Data Set

A Video data set collected from various source like CCTV surveillance gate entrance, shopping malls, banking sectors, railway station and traffic signals from Nagapattinam areas, Tamil Nadu, India on March and April'2019. The dataset consist of 65 videos and we bundled into 5 video clips into single neuron. We applied image processing steps for image enhancement, restoration and support vector calculations. The following figure 1 shows that the moving object clips and motions.

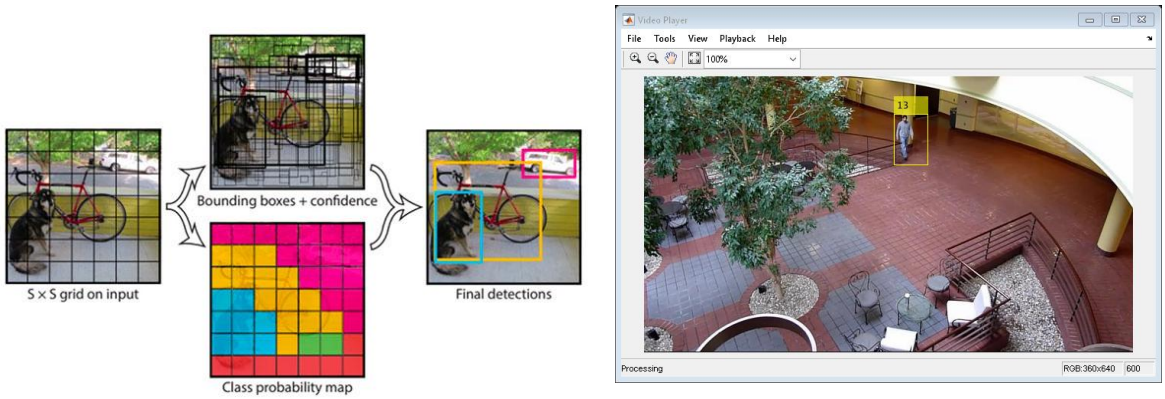

Figure 1. Moving Object Detection

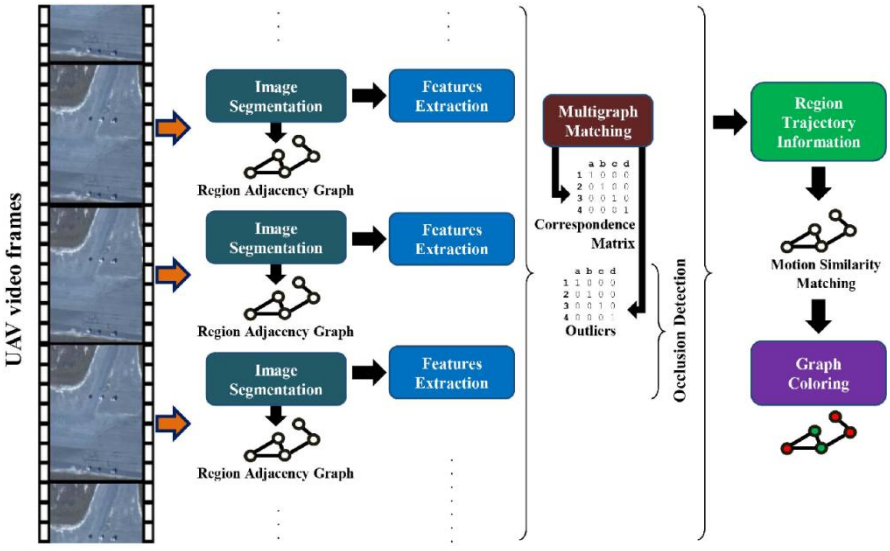

Figure 2. Processing Matrix and Deep learning classification

The video data are associated with environment so we need following challenges for identifying human. i. The video clip is long range so the human appears with various resolutions, ii. Remove fixed and unmoving objects, iii. Remove shaking, blurring and out of focus data, iv. Calculate adaboost classifier value for noise and error factor.

4. Human Detection

Detecting and Locating the human from moving objects is required at various image frames. We selected frame size as (1900 * 1020) and apply human detector algorithms applied for selected image. First the given image frame are sampled by using down factor value 2 and 4 so that we calculated efficiency. In next step the colour images are converted in gray scale image by applying binary classifier techniques. 2D convolution operation is used to obtain moving object prediction. The final step applied for Human detection algorithm with Adaboost classifier.

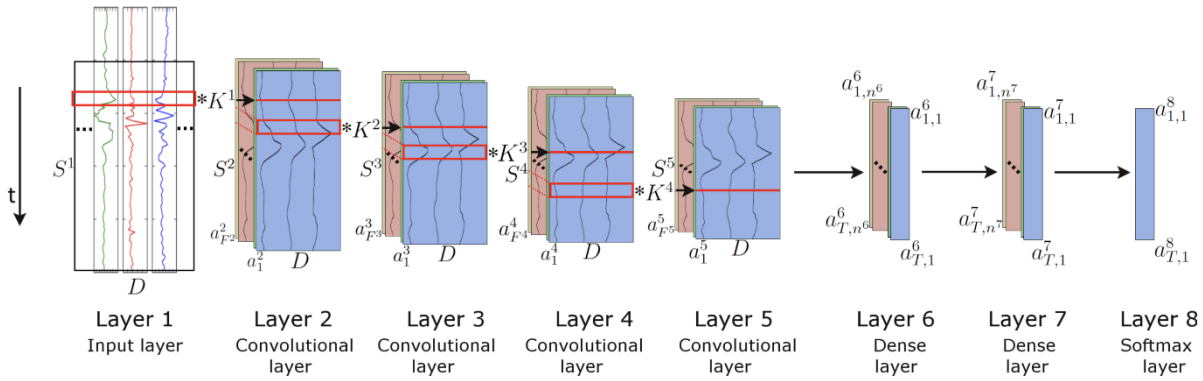

Figure 3. CNN Classifier process

After finding sub images are passed to Convolutional Neural Network classifier. In this classifier has detection stage, misdetection stage, removal stage and labeled stage. The background areas are removed and bundled. The bundled videos are trained and tested and the output are classified as Human without hold anything, Human holds with objects and Human with disabilities. The total 65 videos are trained and tested bundles. Each bundles had 5 video clips, total 13 bundles are divided into 8 training data set and 5 are testing data set.

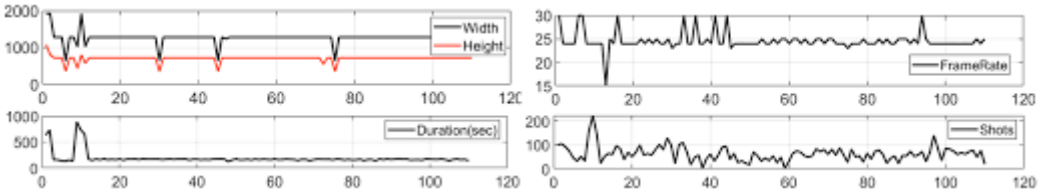

Figure 4. Tensorflow result of CNN values

5. Image based Classification

CNN models are trained from huge number of video clips and calculated by previously trained data. The fully connected layer formed from trained data set and transfer learning methods are used. The transfer learning is used by pre-trained data from AlexNet and pre-trained GoogleNet. The GPU is utilization also important for calculating accuracy factor. The three different CNN approaches, self-defined CNN for training and convolution operations, max-pooling layers for fully connected

layer calculations and AlexNet and GoogleNet for pre-trained model. This approach consists of convolution, max pooling, inception, perception and hidden layer processing.

6. Video based Classification

The video based classification is applied for each bundles, Image based classification is first step based on image result we can create behaviour and fully connected layer. Number of images are classified various cases for example (20,20,10) indicated human classification factor.

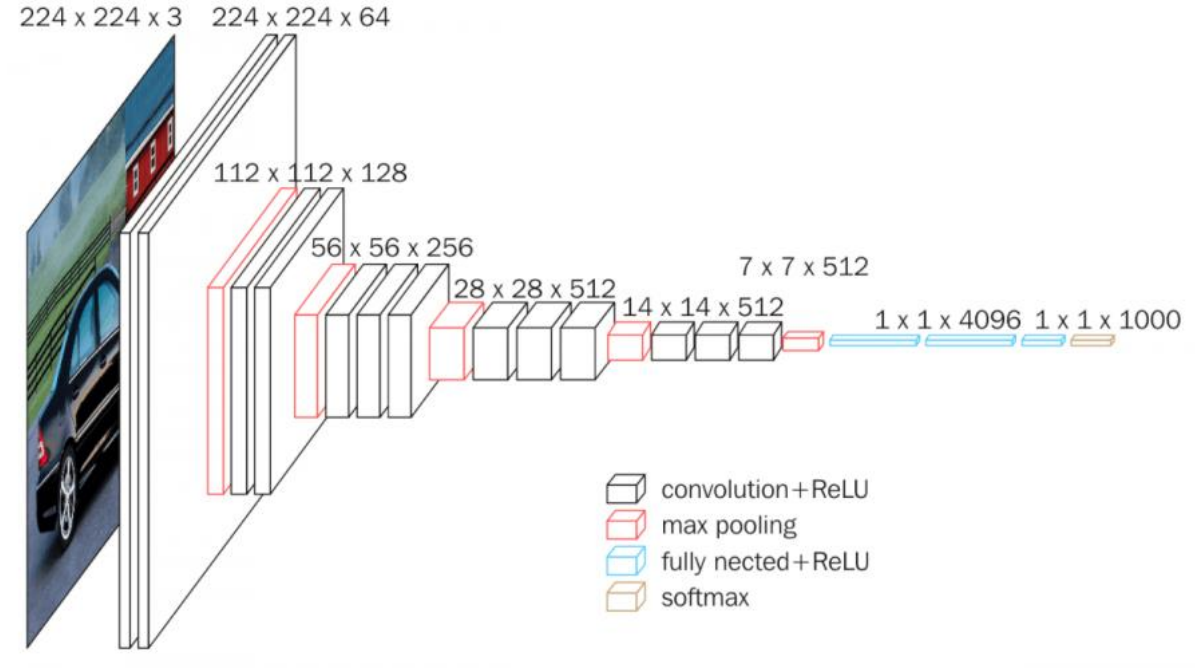

Figure 5. Convolutional and fully connected pooling results

7. Performance Evaluation

The bundled groups are classified by various factors and each layer we calculated accuracy factor. The table 1 shows the bundle group and number of image clips in each bundles.

Table 1: Bundles and Image with sigmoid values

Bundles	Images	Sigmoid(x)
1	23,400	270
2	27,890	280
3	13,457	190
4	12,688	176
5	25,789	278
6	10,678	115
7	11,874	120
8	34,657	312
9	28,098	309
10	34,878	318
11	15,081	198
12	17,898	187
13	21,048	201

ISSN: 2005-4238 IJAST

The following table 2 shows for each bundle values for test and trained result in each layer stage

Bundles	Images	I	II	III	Classifier	Test	Quality	Accuracy
1	23,400	34	32	30	31	378	78%	91%
2	27,890	21	34	32	30	457	68%	84%
3	13,457	34	45	36	40	565	72%	88%
4	12,688	25	27	25	23	476	89%	95%
5	25,789	21	23	25	21	345	78%	91%
6	10,678	56	45	47	45	409	87%	90%
7	11,874	78	78	76	76	786	89%	91%
8	34,657	56	48	67	68	791	69%	79%
9	28,098	34	45	36	37	567	89%	95%
10	34,878	45	67	76	67	456	76%	87%
11	15,081	35	27	29	31	356	89%	90%
12	17,898	45	56	65	55	567	87%	94%
13	21,048	87	87	88	86	897	88%	97%

Table 2: Bundles with image clip classifier result and accuracy rate

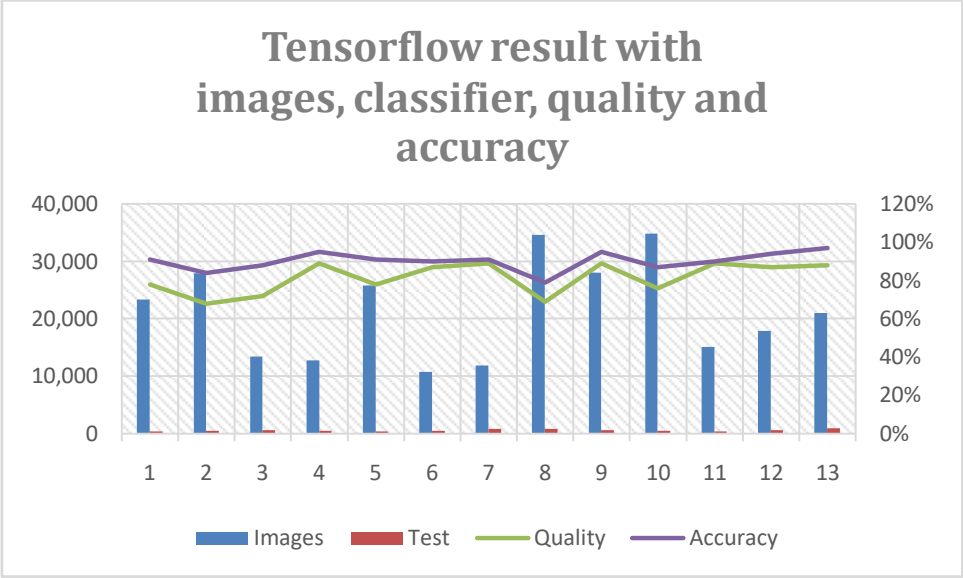

Figure 6. Tensorflow results

8. Conclusion

This paper clearly explained deep learning based approach for finding moving object and human at various classifications. Here we used statistical learning, reinforced learning, machine learning models for various prediction and decision making and deep learning for human prediction. During the detection stage, 65 video clips are bundled and human are identified. The video based classification indicated for each human actions and motions. Several convolutional neural network approaches examined by using Google Net and Google TensorFlow for prediction and decision making. The quality and accuracy rate calculated in each bundles. The average accuracy rate is calculated as 88% for given bundles. In future same approach is used for border areas and cyber crime environments.

References

[1] H.Wei, M.Laszewskai and N.Kehtarnavaz, “Deep Learning-based Person Detection and Classification for Far Field video surveillance”, IEEE, ISSN: 978-1-5386-92962-2/8,2018.
[2] JosepLluisBerral-Garcia, “When and How to apply statistics, Machine Learning and Deep Learning Techniques”, IEEE-ICTON-2018,978-1-5386-6605-0/18,pp-1-4,2018
[3] Kemal Ozkan, Osman Nuri Atak and SahinIsik, “Using Movie Posters for Prediction of Box-Office Revenue with Deep Learning”, IEEE-978-1-5386-1501-0/18,2018
[4] Deep Learning Tutorials, http://deeplearning.net/tutorial.
[5] Martın Abadi, et al., TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems, arXiv:1603.04467v2, 2015.
[6] Pandas, Python Data Analysis Library, http://pandas.pydata.org.
[7] Martın Abadi, et al., TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems, arXiv:1603.04467v2, 2015.
[8] S.Manikandan, Dr.M.Chinnadurai, "Motion Detection Algorithm for Agent Interaction Surveillance Systems", International Journal of Engineering Technology Science and Research (IJETSR), ISSN 2394 – 3386, Volume-4, Issue-11, pp:408-412, November-2017
[9] Fumiya Kudo, Tomoki Akitomi and Norihiko Moriwaki, “An Artificial Intelligence Computer System for Analysis of Social Infrastructure Data”, 2015 IEEE 17th Conference on Business Information, DOI:978-1-4673-7340-1/15, pp: 85-89
[10] S.Manikandan, M.Chinnadurai, D.Maria Manuel Vianny and D.Sivabalaselvamani, "Real Time Traffic Flow Prediction and Intelligent Traffic Control from Remote Location for Large-Scale Heterogeneous Networking using TensorFlow", International Journal of Future Generation Communication and Networking, ISSN: 2233-7857, Vol.13, No.1, (2020), pp.1006-1012.
[11] S. Manikandan and K. Manikanda Kumaran, "Identifying Semantic Identifying Semantic Relations Between Disease And Treatment Using Machine Learning Approach", International Journal of Engineering Research & Technology(IJERT), essn:2278-0181,Vol.2,June - 2013.
[12] Manikandan, S & Chinnadurai, M 2019, ‘Intelligent and Deep Learning Approach OT Measure E-Learning Content in Online Distance Education’, The Online Journal of Distance Education and e-Learning, vol.7, issue 3, July 2019, ISSN: 2147-6454.
[13] Upadhyay R. Artificial Neural Networks, Made Easy – Retail Case Study Example (Part 8) / YOU CANalytics. Explore the Power of Predictive Analytics. 29th September, 2014 Available at: http://ucanalytics.com/blogs/artificial-neural-networks-retail-case-studyexample-part-8
[14] V. Lakshmanan: Data Science on the Google Cloud Platform: Implementing End-to-end Real-time Data Pipelines: from Ingest to Machine Learning

AUTHORS PROFILE

S. Manikandan is working as Assistant Professor of IT in E.G.S Pillay Engineering College, Nagapattinam. He completed M.E-CSE in Annamalai University with First class with Distinction, 2012 and B.Tech - IT in E.G.S Pillay Engineering College with First class with Distinction, 2010. Currently he is doing Ph.D in Anna University, Chennai and his research work includes Artificial Intelligence, Network Security, Algorithms and Cloud Computing.

M. Chinnadurai is working as Professor of CSE in E.G.S Pillay Engineering College Nagapattinam. He completed his Ph.D in Anna University, Chennai in the field of VLSI at Faculty of Information and Communication Engineering, Anna University, Chennai. He is professional member in IEEE, CSI, ISTE, etc and his research work includes Artificial Intelligence, Network Security, Algorithms and Cloud Computing.

M.P.Thiruvenkatasuresh, (Mettupalayam Ponnusamy Thiruvenkatasuresh) received the B.E. Degree in Computer Science and Engineering from Madras University, Chennai, Tamilnadu, India in the year 1994 and M.E. degree inComputer Science and Engineering at Annamalai University, Chidamparam, Tamilnadu, India in the year 2005. He completed Ph.D Degree in Anna University, Chennai, Tamilnadu, India in the year 2018. He is currently working as a Professor in the department of computer Science and Engineering, Erode Sengunthar Engineering College, Perundurai,Erode, Tamilnadu, India. He has 21 years of teaching experience and 8 years in research. His research interests are Image and video processing, Data mining and cloud computing.

M. Sivakumar, received the B.E. Degree in Computer Science and Engineering from Anna University, Chennai, Tamilnadu, India in the year 2007 and M.E. degree inComputer Science and Engineering at Anna University, Coimbatore, Tamilnadu, India in the year 2010. He is Pursuing Ph.D at Anna University, Chennai, Tamilnadu, India. He is currently working as a Assistant Professor in the department of computer Science and Engineering, Erode Sengunthar Engineering College, Perundurai, Erode, Tamilnadu, India. He has 11 years of teaching experience and 3 years in research. His research interests are Machine Learning and cloud computing.