International Journal of Advanced Science and Technology
Vol. 29, No. 9s, (2020), pp. 6860-6864

Human Action Recognition In Streaming Video Using Deep Learning

G.Divya[1], SP. Chokkalingam[2]

[1]Research Scholar, Department of Computer Science and Engineering, Saveetha School Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai
[2]Professor, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai.
Email:mailtodivya16@gmail.com, cho_mas@yahoo.com

Abstract

Human action recognition in streaming video is an open issue in surveillance monitoring system in the field of Artificial Intelligence. To protect and prevent from robbery, chain snatching issue surveillance monitoring system is necessary to the society. In this paper we discussed about a novel architecture of Deep learning based CNN is used to address the above issue. Human Action frame prediction model is used to predict the various action of human in streaming video. UCF101 data set is used to predict the various actions with learning rate of 0.01. The proposed approach yields 72.34% accuracy of identifying the human actions compared to various traditional models.

***Keywords**: Deep Learning, Convolutional Neural Network, Video Processing*

1. INTRODUCTION

In today's world, CCTV (Closed Circuit Television Camera) play integral part of our day to day life . CCTV is installed and equipped in public places like malls, railways, schools, Hospitals. Apart from public places, CCTV are installed in our home to monitor the activities of the people for safety purpose like to avoid from theft, wandering, etc. Video Surveillance system plays important research area in the emerging field of Artificial Intelligence. Managing and monitoring the videos in sequence is an open issue in the research field of AI. Abnormal activities should be monitored in streaming video to ensure the security of the place and to identify abnormal person in the system from robbery, chain snatching etc. The main objective of the paper is to provide security of the system and to identify the Human Actions in various Perspectives. One of the best solutions to provide and enhance the existing security is surveillance monitoring.

Deep learning can be explained with the help of industries that invest on artificial, intelligence, big data and Analytics. Deep learning has become increasingly popular now-a-days and is considered as an important requirement of vast number of industries. Social networking sites like Amazon and Netflix uses deep learning to understand the requirement and behaviour of their customers and Google uses it for voice and image recognition. To our astonishment, MIT's researchers are now trying to use deep learning for predicting the future. This is the impact of deep learning on various industries in the world.

1.1 Draw Backs of Machine Learning

- When a high dimensional data with large number of inputs and outputs is to be analyzed and worked upon, the traditional ML algorithm might not be useful. For instance, we have a large amount of input and also various types of inputs associated with various type of handwritings in "hand writing recognition".

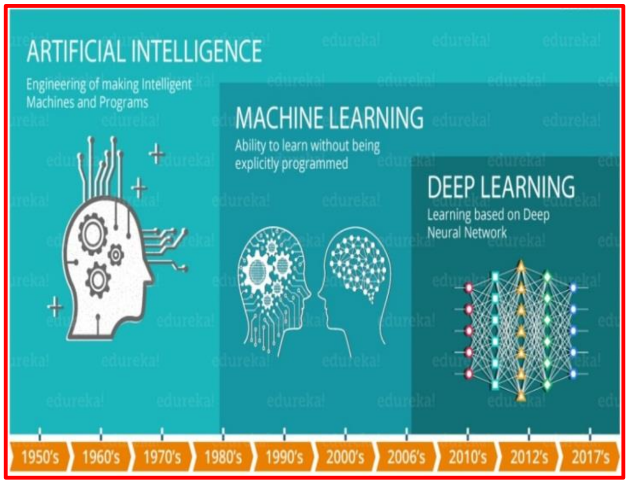

Figure-1. Time Line of Deep Learning

- The feature extraction is the next important challenge which plays a predominant role in predicting the future outcome and also in achieving better accuracy. For this purpose, input has to be given to the computer about the features it should look for.

Applying machine learning algorithms is more perilous in solving complex problems such as object detection, recognition and NLP process. This is because in traditional machine learning, when an algorithm is fed with new data, it hardly works and this makes feature extraction a unfavorable part of machine learning workflow. The efficiency of the Programmer is therefore affected by these drawbacks and this in turn reduces the effectiveness of the algorithm.

2. PROPOSED APPROACH

In the proposed approach, Convolutional Neural Network has been widely used in video processing. Generally CNN is applicable to AI techniques like Hand character recognition, detection of classification of objects in streaming videos ,speech recognition etc. In this approach human actions are classified based on time series using Region of Interest (ROI) and irrelevant sequences of images. LSTM is very good approach to model long sequences than RNN (Recurrent Neural network).LSTM predict the next frame automatically in the given figure 2 which has 2048 dimensional from each frame RESNET-50.

2.1 Algorithm

- 2 layers of LSTM, First layer consider as encoder and second Layer consider as decoder.
- Squared error loss function is applied to train the model
- Let R_i be input to another network which act as classifier
- The output of the encoder is trained which will be given input to other layer.

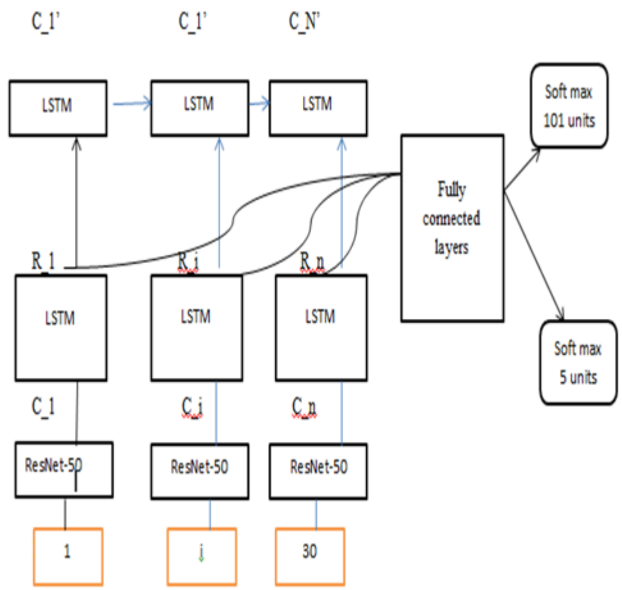

Figure 2 Network for prediction model –Human Action

In this approach, this network is fully connected with 1024 units classified as linear and non linear units. Output layer is used to predict the action and classify with 101 and 5 units.

3. EXPERIMENTAL RESULTS:

UCF-101 Data set is used to analyses the human action .Adam optimizer is used to train the networks with learning rate of 0.01. To avoid over fitting both LSTM and Fully connected layer is used . To predict the next frame for 30 epochs is used .Figure 3 shows loss function with the number of epochs . and shows the accuracy of the model over epochs. Table 1 gives the results of our model comparing with various models.

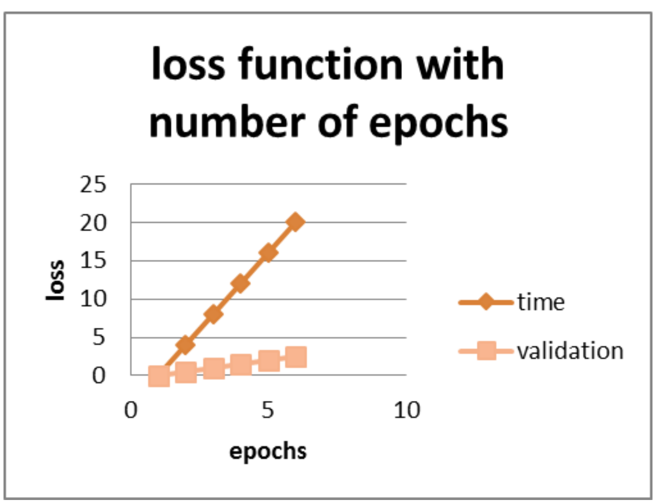

Figure 3- Loss function

MODEL	ACCURACY
CNN only(ResNet-50)	56.8%
Super class model	62.333%
Next Frame Prediction	72.34%

Table 1 – Summary of accuracy with various models.

4. CONCLUSION

Image representation from ResNet-50 significantly out per-forms Alex Net than previous methods .This suggests that an excellent representation of static images is critical for good video category. From associated works we can finish that optical drift is a good motion function and captures temporal statistics that enables motion recognition. Further we also examine that pre-schooling a model on larger relevant dataset likeSports-1M or ImageNet enables the network learn a better representation. This is expected because of large type of information contained in those datasets. Finally, predicting the subsequent frame of a video appears to analyze as a minimum some relevant features. We look at a small improvement in accuracy, however given the number of classes and trouble of the undertaking, this small improvement is big. The encoder-decoder based model for video illustration may be used for an expansion of video.

REFERENCES

1. Roberto Arroyo, J. Javier Yebes, Luis M. Bergasa, Iv´an G. Daza and Javier Almazan, “Expert video-surveillance system for realtime detection of suspicious behaviors in shopping malls”, *Expert Systems with Applications*, Vol. 42, No. 21, 2015, pp. 7991-8005.
2. Kun Wang, Stanley Langevin, Corey O’Hern, Mark Shattuck, Serenity Ogle, Adriana Forero, Juliet Morrison, Richard Slayden, Michael Katze, Michael Kirby, “Anomaly detection in host signaling pathways for the early prognosis of acute infection”, *PloS one*, Vol. 11, No. 8, 2016.
3. Yudong Zhang, Genlin Ji, Jiquan Yang, Shuihua Wang, Zhengchao Dong, Preetha Phillips and Ping Sune, “Preliminary research on abnormal brain detection by waveletenergy and quantum-behaved PSO”, *Technology and Health Care*, 2016, pp. 1-9.
4. Soumi Ray and Adam Wright, “Detecting Anomalies in Alert Firing within Clinical Decision Support Systems using Anomaly/Outlier Detection Techniques”, *Proceedings of the 7th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics, USA*, 2016, pp. 185-190.
5. Wei Wang, Lin Chen, Kang Shin, and Lingjie Duan, “Thwarting intelligent malicious behaviors in cooperative spectrum sensing”, *IEEE Transactions on Mobile Computing*, Vol. 14, No. 11, 2015, pp. 2392-2405.
6. Oscar Rojas and Clesio Tozzi, “Abnormal Crowd Behavior Detection Based on Gaussian Mixture Model”, *Proceedings of European Conference on Computer Vision, Springer International Publishing, Cham*, 2016, pp. 668-675.
7. Andrea Pennisi, Domenico Bloisi, and Luca Iocchi, “Online real-time crowd behavior detection in video sequences”, *Computer Vision and Image Understanding*, Vol. 144, 2016, pp. 166-176.
8. Supriya Mangale and Madhuri Khambete, " Camouflaged Target Detection and tracking using thermal infrared and visible spectrum imaging”, *Proceedings of International Symposium on*

Intelligent Systems Technologies and Applications, Springer International Publishing, Cham, 2016, pp. 193-207.

9. Yang Xian, Xuejian Rong, Xiaodong Yang, and Yingli Tian, "Evaluation of Low-Level Features for Real-World Surveillance Event Detection", *IEEE Transactions on Circuits and Systems for Video Technology,* Vol. 27, No. 3, 2016, pp. 624-634.
10. Serhan Coşar, Giuseppe Donatiello, Vania Bogorny, Carolina Garate, Luis Otavio Alvares, and François Brémond, "Toward Abnormal Trajectory and Event Detection in Video Surveillance", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 27, No. 3, 2017, pp. 683-695.
11. Sarita Chaudharya, Mohd Aamir Khana, Charul Bhatnagar, (2018), "Multiple Anomalous Activity Detection in Videos", *Procedia Computer Science, Vol. 125, pp.* 336–345.
12. "CMU Graphics Lab Motion Capture Database.", http://mocap.cs.cmu.edu/, last accessed 2018/1/2.
13. Ryoo, M.S. and Aggarwal, J.K.: "UT-Interaction Dataset, ICPR contest on Semantic Description of Human Activities (SDHA)", http:// cvrc.ece.utexas.edu /SDHA2010 /Human _Interaction.html.
14. "Peliculas Movies Fight Detection Dataset", http://academictorrents.com/ details /70e0794e2292fc051a13f05ea6f5b6c16f3d3635 /tech&h it=1&filelist=1, last accessed 2018/1/5.
15. E. Bermejo, O. Deniz, G. Bueno, R. Sukthankar: "Violence Detection in Video using Computer Vision Techniques", Proceedings of Computer Analysis of Images and Patterns. (2011).
16. "CRF Web Dataset", http://crcv.ucf.edu/projects/Abnormal_Crowd/#WebDataset, last accessed 2018/1/5.
17. http : // www . svcl . ucsd . edu / projects / anomaly / dataset.htm.
18. Balasundaram and Chellappan, (2018), " An intelligent video analytics model for abnormal event detection in online surveillance video", Journal of Real-Time Image Processing, Doi: 10.1007/s11554-018-0840-6.
19. M. Bertini, A. Del Bimbo, and L. Seidenari, "Multi-scale and realtime non-parametric approach for anomaly detection and localization", Compt. Vis. Image Und., 116(3):320–329, 2012.
20. V. Reddy, C. Sanderson, and B. C. Lovell, "Improved anomaly detection in crowded scenes via cell-based analysis of foreground speed, size and texture", In CVPR Workshops, pages 55–61, 2011.
21. V. Saligrama and C. Zhu, "Video anomaly detection based on local statistical aggregates", In CVPR, pages 2112–2119, 2012.