

Anomaly Detection in Real Time CCTV Streams

D.M.B. Kularathne

179330X

Degree of Master of Science

Department of Computer Science Engineering

University of Moratuwa
Sri Lanka

July 9, 2020

Declaration

I declare that this is my own work and this dissertation does not incorporate without acknowledgement any material previously submitted for degree or Diploma in any other University or institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, I hereby grant to University of Moratuwa the non-exclusive right to reproduce and distribute my dissertation, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

Signature:

Date:

Name: D. M. B. Kularathne

The above candidate has carried out research for the Masters/MPhil/PhD thesis/Dissertation under my supervision.

Signature of the supervisor:

Date:

Name: Dr. Charith Chitraranjan

Abstract

Anomaly detection in video data has been a challenge always. After the introduction of many state-of-art designs, this still poses a challenge as those systems may fail to work in all types of environments. Even though many supervised methods claimed to have some good results in this domain, supervised systems may not be suitable for all the contexts such as in an open area, any type of anomaly can occur and it can be very difficult to train a system in a supervised manner to identify an unanticipated anomaly. On the other hand, it would be difficult for the user to annotate data each time when they change the context under surveillance for the device. Thus the ultimate solution should be an unsupervised solution with a appreciable accuracy. Recently deep learning techniques have emerged in many areas of computer science based solutions and so it is involved for anomaly detection tasks also. In this research, deep learning techniques are involved to solve the problem of video stream based anomaly detection of crowds.

Acknowledgements

I would like to express my gratitude for my supervisor Dr. Charith Chitranjan for the guidance and support extended since beginning. His invaluable advice greatly helped me to drive this research work on the right track. I further would like to thank my colleagues from the MSc batch who shared their knowledge and insights with me.

I am certainly in debt to my parents who guided me throughout the journey of life. I wish to give my heartiest thanks to my loving wife who stood beside me, encouraged me and supported me throughout my work and this would not have been possible without her immense support and love. Finally, I would like to extend my gratitude to all the colleagues at Synopsys who helped me in various ways to fulfill this task.

Contents

Declaration	i
Abstract	ii
Acknowledgements	iii
List of Figures	vii
List of Tables	ix
1 Introduction	1
1.1 Anomaly Detection	3
1.1.1 Definition of Anomalies	3
1.1.2 Challenges in Anomaly Detection	4
1.1.3 Aspects of Anomaly Detection Problem	4
1.1.3.1 Data As Input	4
1.1.3.2 Type of Anomaly	4
1.1.3.3 Availability of Labeled Data	5
1.2 Abnormal Human Behavior Recognition	7
1.2.1 Challenges of modeling the human behavior in videos	7
1.3 Problem	8
1.4 Objectives	9
1.4.1 Specific Objectives	10
1.5 Prior Work	11
1.5.1 Models and Algorithms for Abnormality Detection in Video Streams (Conventional Models)	11
1.5.1.1 Dynamic Bayesian Network	11
1.5.1.2 Bayesian Topic Models	12
1.5.1.3 Clustering	13

1.5.1.4	Decision Trees	13
2	Literature Review	15
2.1	Anomaly Detection using Deep Learning	16
2.1.1	Representation Learning(RL)	16
2.1.1.1	RL For Reconstruction	16
2.1.1.2	RL for Predictive Modeling	16
2.1.1.3	RL for Generative Models	17
2.1.2	Reconstruction Models	17
2.1.2.1	Principal Coponent Analysis(PCA)	17
2.1.2.2	Auto-Encoders	18
2.1.2.3	Convolutional Autoencoders	18
2.1.2.4	Sparse Autoencoders	22
2.1.2.5	Contractive Autoencoders	23
2.1.2.6	De-Noising autoencoders	23
2.1.2.7	Adversarial Autoencoder	23
2.1.2.8	Deep Belief Networks	24
2.1.3	Predictive Modeling	25
2.1.3.1	Usages of LSTM (Long-Short-Term-Memory)	25
2.1.3.2	Convolutional Long Short Term Memory (ConvLSTM)	26
2.1.3.3	Slow Feature Analysis (SFA)	26
2.1.3.4	Other Predictive Models	28
2.1.4	Deep Generative Models	28
2.1.4.1	Variational Autoencoders	28
2.1.4.2	Generative Adversarial Network(GAN)	29
2.2	State-Of-The-Art Models	32
2.2.1	AMDN	32
2.2.2	ConvLSTMAE	32
2.2.3	ConvAE	33
3	Methodology	34
3.0.1	Model Architecture	35
3.0.1.1	Motivation	35
3.0.1.2	Model Definition	37
3.0.1.3	Training	39
3.0.1.4	Anomaly Detection	39

3.0.1.5	Variations of Anomaly Score Calculation	40
4	Experiments and Results	43
4.1	Data Set	44
4.2	Experiments	44
4.2.1	Hardware Settings	45
4.2.2	Anomaly Count	45
4.2.3	Experiment Methodology	46
4.3	Results	46
4.3.1	UCSD PED1	48
4.3.1.1	True Positives	50
4.3.1.2	True Positives Outside Ground Truth	52
4.3.1.3	False Positives	56
4.3.1.4	The False Negatives	57
4.3.2	UCSD PED2	58
4.3.2.1	True Positives	60
4.3.2.2	False Positives	61
4.3.3	Improving Results - Weighted Pixel Values	62
4.3.3.1	Weighting Based Differences in Score	62
4.3.4	Further Experiments	63
4.3.4.1	Avenue Data set	63
4.3.4.2	Latent Space Patterns	65
4.4	Conclusion	66
4.5	Future Work	67
	Bibliography	69
	Appendices	76
.1	Canny Edge Detection	77
.2	MOG2 Background Subtraction	78
.3	Persistence1d Algorithm	78
.4	UCSD Anomaly Detection Data set	78

List of Figures

2.1	Spatio-Temporal Stacked frame Auto-Encoder	20
2.2	Convolutional LSTM based autoencoder	22
2.3	Adversarial Autoencoder	24
2.4	SFA	27
2.5	The basic structure of a GAN	30
2.6	ADAE	31
2.7	AMDN	32
2.8	ConvLSTMAE	33
2.9	ConvAE	33
3.1	Overall Model Architecture	38
3.2	Edges found in PED1 video clips	41
3.3	Background Subtraction on PED1 video clips	42
4.1	ROC for PED1, vanilla model with uniform weighting	48
4.2	ROC for PED1, edge based weighting	48
4.3	ROC for PED1, background subtraction based weighting	49
4.4	True positive: Test01: Biker	50
4.5	True positive: Test04: Skater	51
4.6	True positive: Test01: Van	51
4.7	True positive: Test01: Biker	52
4.8	ROC for PED1, ground truth correction	52
4.9	Ground truth correction	53
4.10	Test22: Corrected ground truth added in 4.9	53
4.11	Ground truth correction	54
4.12	Test26: Corrected ground truth added in 4.11	54
4.13	Ground truth correction	55
4.14	Test18: Corrected ground truth added in 4.13	55
4.15	False positive: Test34: change of travel pattern	56

4.16	False positive: Test12: camera shake	56
4.17	False positive: Test11: Odd direction	57
4.18	False positive: Test11: False negative	57
4.19	ROC for PED2, vanilla model with uniform weighting	58
4.20	ROC for PED2, edge based weighting	58
4.21	ROC for PED1, background subtraction based weighting	59
4.22	Test04, PED2 true positive	60
4.23	Test06, PED2 true positive	60
4.24	Test07, PED2 true positive	60
4.25	Test01, Unusual objects in hand	61
4.26	Test02, Suddenly stop moving	61
4.27	Test12, Unusual objects in hand	61
4.28	Test002 : Left to Right in order: vanilla model, edge detected, background subtraction based, background subtraction based in- creased background weights	63
4.29	Test005 : Left to Right in order: vanilla model, edge detected, background subtraction based, background subtraction based in- creased background weights	63
4.30	Test006 : Left to Right in order: vanilla model, edge detected, background subtraction based, background subtraction based in- creased background weights	63
4.31	04.avi : Running behavior	64
4.32	05.avi : Throwing objects	64
4.33	06.avi : Different types of behavior	65
4.34	Test008 : PED1 — Left:- reconstruction error, Rigt:- discriminator score	65
4.35	Test001 : PED2 — discriminator score	66
36	Maxima detection using persistence1d	78

List of Tables

3.1	Encoder Network	37
3.2	Decoder Network	37
3.3	Discriminator Network	38
4.1	The table lists the AUC(Area Under Curve) and EER(Equal Error Rate) rates of state of the art models from 2010 to 2019. Along with listed are the corresponding rates of the proposed model. Higher the AUC the better and lower the EER the better. The current state-of-the-art rates are in square brackets.	47

Abbreviations

AAE Adversarial Autoencoder. 24, 36, 55

ADAЕ Adversarial Dual Autoencoder. 29

Adam Adaptive Moment Estimation. 39

AE Autoencoder. 16, 18, 19, 21, 36

AUC Area Under Curve. ix, 32, 46, 47, 49, 52, 59, 66, 67

CAE Convolutional Autoencoder. 18–20, 31

CNN Convolutional Neural Network. 18, 19, 25, 28, 30

ConvLSTM Convolutional Long-Short-Term-Memory. v, 26

EER Equal Error Rate. ix, 32, 46, 47, 49, 52, 59, 66, 67

GAN Generative Adversarial Network. v, vii, 17, 29–31, 35

GIGO Garbage In Garbage Out. 4

GMM Gaussian Mixture Model. 13

HDP Hierarchical Dirichlet Process. 12

HMM Hidden Markov Model. 11, 12

LDA GLatent Dirichlet Allocation. 12

LSTM Long-Short-Term-Memory. v, 17, 25, 26, 28, 32, 33, 36, 37

MAE Mean Absolute Error. 24, 36

MSE Mean Squared Error. [24](#), [25](#), [31](#), [36](#), [39](#)

PCA Principal Component Analysis. [v](#), [16–20](#), [28](#)

RBM Restricted Boltzmann Machine. [24](#)

ReLU Rectified Linear Unit. [18](#)

RL Representation Learning. [v](#), [16](#), [17](#)

RNN Recurrent Neural Network. [25](#)

ROC Receiver Operating Characteristic. [46](#)

SFA Slow Feature Analysis. [v](#), [26–28](#)

STSAE Spatio-Temporal Stacked frame Auto Encoder. [20](#)

VAE Variational Autoencoder. [17](#), [36](#)

List of Appendices

Canny Edge Detection	77
MOG2 Background Subtraction	78
Persistence1d Algorithm	78
UCSD Anomaly Detection Data set	78