Fractional Data Distillation Model for Anomaly Detection in Traffic Videos

Linu Shine, Vaishnav M A, Jiji C V
College of Engineering Trivandrum
Kerala, India
linushine@cet.ac.in, vaishnavma@cet.ac.in, jijicv@cet.ac.in

Abstract

Timely automatic detection of anomalies like road accidents forms the key to any intelligent traffic monitoring system. In this paper, we propose a novel Fractional Data Distillation model for segregating traffic anomaly videos from a test dataset, with a precise estimation of the start time of the anomalous event. The model follows a similar approach to that of the typical fractional distillation procedure, where the compounds are separated by varying the temperature. Our model fractionally extracts the anomalous events depending on their nature as the detection process progresses. Here, we employ two anomaly extractors namely Normal and Zoom, of which former works on the normal scale of video and the latter works on the magnified scale on the videos missed by the former, to separate the anomalies. The backbone of this segregation is scanning the background frames using the YOLOv3 detector for spotting possible anomalies. These anomaly candidates are further filtered and compared with detection on the foreground for matching detections to estimate the start time of the anomalous event. Experimental validation on track 4 of 2020 AI City Challenge shows an s4 score of 0.5438, with an F_1 score of 0.7018.

1. Introduction

Humans are amazingly fast in inferring details from the visual world and can easily spot anomalies in the world around them, even from videos. But researches show that human attention span reduces considerably after 20 minutes. Therefore, when terabytes of data generated by CCTVs installed across cities and motorways need to be monitored by human operators alone, the chances of missing abnormal events are high. Hence, an automated system that can assist humans can become life-saving in many cases. Though developing an Artificial Intelligence-based system with near-human levels of visual cognition system still seems to be a fantasy, developments in video-based anomaly detection is a sure promise in this direction.

In a broad sense, an anomaly is defined as any deviation from the normal behaviour pattern. This definition varies widely according to the situation in which anomaly detection is applied. For example in AI City challenge 2020[1], crashed or stalled vehicles are considered as anomalies. These are to be detected from video feeds available from multiple cameras at intersections and along highways. Here, all stopped vehicles may not be counted as an anomaly. Vehicles parked in parking lots or vehicles waiting for a green signal in traffic light should not be considered as anomalies, while vehicles stopped in hazard lanes or vehicles that get involved in accidents and get stalled either on roads or in grass areas should be considered as anomalies. The rarity of such contextual anomalies makes it hard for employing traditional supervised learning methods.

The goal of any anomaly detection framework is the timely detection of anomalous events. An additional constraint in this challenge is that the framework should be based on existing models which can use pre-trained weights trained in public data sets like ImageNet[7] or COCO[11]. This paper proposes a novel Fractional Data Distillation model for segregating anomalous videos. In this approach, a primary search is done in all videos using a Normal extractor to sieve the first fraction of anomalies. A Zoom extractor is used to capture the second fraction of anomalies, from whatever is left out by the Normal extractor. A rule-based decision module is developed using YOLO detector[16] to scan foreground and background images to detect anomalies. We also propose a faster method for selecting anomaly candidates by performing detection only on samples of background frames averaged at each one minute interval. This amounts to running the detections on just 14 background frames per video of 27,000 frames.

The main contributions of this paper are summarized as follows:

- A novel Fractional Data Distillation scheme in which anomalies are segregated fractionally using normal and zoom extractors.
- A fast method for selecting anomaly candidates by

scrutinizing 14 background frames per video using the YOLO detector.

The rest of this paper is organised as follows. Section 2 reviews some of the works in the field of anomaly detection. The proposed methodology for detecting anomalies is detailed in section 3. The experiments and evaluation of our method in track 4 challenge data set are explained in Section 4. Section 5 concludes the paper.

2. Related Work

The surge in the number of CCTVs installed to monitor traffic flow across the globe has prompted researchers to develop intelligent systems to analyze the recordings. Anomaly detection in traffic videos is one such budding area were a lot of research is happening. The algorithms in this area can be broadly classified as supervised, semi-supervised and unsupervised.

Anomaly detection approaches can be further classified as model-based, proximity-based and reconstruction based[9]. In a model-based approach, a set of parameters are used to represent the normal behaviour of the data. For example, in [21] an abnormality indicator is developed using hidden Markov model used to differentiate normal and abnormal frames in a video. Li *et al*. [10] used local invariant features from the video blocks to estimate the probability of normal traffic using a Gaussian distribution model. This learned model is used for detecting anomalies. Sultani *et al*. [18] used a weakly supervised learning method for differentiating normal and abnormal video segments.

Proximity-based systems use a distance-based approach to extract anomalies where it is assumed that abnormal frames have large distance vectors. Clustering of optical flow feature vectors[6] or vehicle trajectories[8] is used to identify the abnormality in videos.

A generative model-based technique is proposed in [5] where regular motion patterns are learned from normal videos using an autoencoder. This learned network produces high reconstruction error while reconstructing abnormal frames, which is used to locate the anomalies. The state-of-the-art methods using generative adversarial networks also utilize the reconstruction error in predicting abnormal frames [12]. The method used future frame prediction and utilizes motion flow vectors like optical flow for anomalous frame prediction.

Investigations of previous challenge submissions on traffic anomaly detection track reveal that most of the successful teams have utilized foreground segmentation methods to find anomaly candidates[14]. Some researchers have used traditional methods like Gaussian mixture models (GMM)[22], while [4] has employed a CNN based network for background modelling. Bai *et al*. [3] proposed a novel combination of background modeling, perspective detection module and spatio-temporal information matrix for detecting anomalies. Tracklenet tracker was used to estimate the trajectory of anomaly candidates from background detection for predicting the exact time of anomaly[20].

In this paper, we propose a novel approach which utilizes a magnification based distillation scheme for separating the anomalies in traffic videos. The proposed method also utilizes a GMM based background detection model, but instead of running the detector on the entire background video, only 14 sampled background frames are utilized. These samples are from averages of background frames in every minute and are used to select the anomaly candidates.

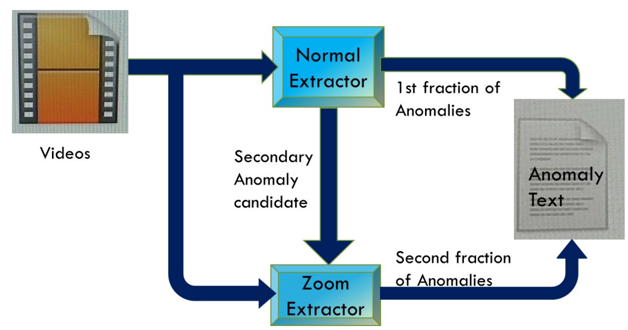

Figure 1. Overview of the proposed method.

3. Proposed Method

Track 4 of AI City challenge 2020 focuses on the timely detection of anomalies from traffic videos. Vehicles that get stalled on-road or hazard lanes are counted as anomalies. We propose a Fractional Data Distillation model for sieving out these anomalies. Different sets of anomalies get extracted at different stages of the algorithm, similar to the distillation process used in separating petroleum products from crude oil. Hence the name Fractional Data Distillation model.

Most of the vehicles that get involved in accidents are stopped immediately after the crash or within a few seconds of its occurrence. Stationary vehicles are therefore, a sure sign of possible anomalies. All static elements in a video will be grouped as *background* in background segmentation algorithms. This prompted us to select a background modelling method as the backbone for our model. An Adaptive Gaussian Mixture model serves this purpose [22].

The proposed Fractional Distillation model extracts anomalies in two stages as depicted in the block schematic in Figure 1. Traffic videos are extracted into frames and fed into the Fractional Data Distillation module. The module consists of a Normal extractor and a Zoom extractor, in which the Normal extractor is used to detect anomalies in

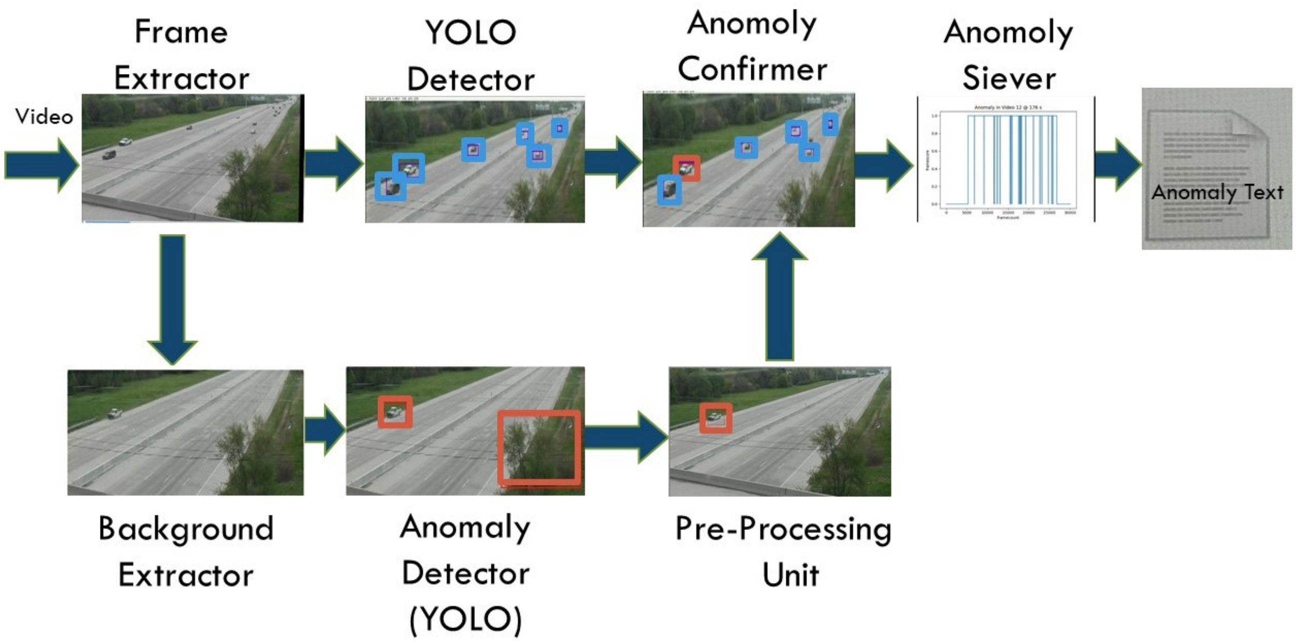

Figure 2. Block level representation of Normal extractor.

the frame level while the Zoom unit extracts anomalies at different block levels inside the frame. Each anomaly extractor unit consists of a pre-processing unit, anomaly detector and the anomaly confirmation module.

3.1. Normal extractor

An overview of the Normal extractor is shown in figure 2. Vehicles that get stalled due to accidents will be a part of the background after some time and they are possible candidates of anomalies. The background is extracted from the videos using the GMM technique. Since the method is robust against lighting variations, temporal noises and image jittering, a stable background is obtained within a few frames from the beginning of the video. Background masks are created from the background video by averaging out the generated background for one minute as shown in Figure3. The process is continued until the end of each video. So for every 15 minutes of video, 14 one-minute background masks are created. The last minute is ignored while creating the masks. Now an object detector is employed to detect vehicles in all these background masks. We have utilized the object detector API provided by[13] for detecting vehicles. The detector uses pre-trained weights of YOLOv3 trained on COCO dataset. YOLOv3 belongs to the class of single-stage object detector which provides the fastest method for object detection. The detection rate of the detector is 30 *fps*[2]. Our approach for anomaly candidate fixing is computationally efficient, since we limit our detector to scan only these 14 background frames per video instead of 27000 frames of background video. Sample detection in background is shown in Figure 4.

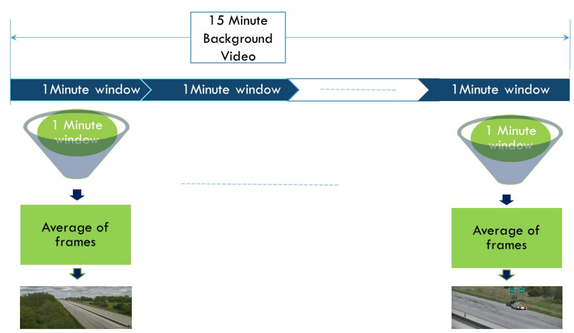

Figure 3. Background masks generated by averaging in 1 minute window

All detections in the background of a particular video are then passed to a preprocessing unit. A universal, perfect object detector with a near mean average precision of 100 is still the holy grail in the field of computer vision. Even the best performing detection model needs tuning in a particular data set to produce the best result. Manual annotation of 25 hour videos in the train data set is a near-impossible task with our limited resources. Therefore, predictions with our pre-trained detector include a small number of false predictions as well. To eliminate some of these false positives from the background detections, a preprocessing stage is employed. This method is a crude form of eliminating false positives based on the size of detection boxes and cannot eliminate all the false detections. Since real world videos are used as data set in the challenge, some videos contain frozen frames. These freeze periods are calculated and the detections within the freeze periods are eliminated in this stage. At this point, all detections in the background are considered as references for possible anomalies.

Figure 4. Detection on background masks

There is a considerable time for stalled vehicles to appear in the background, and this varies from video to video depending on its quality [17]). To find the exact time of anomaly, the anomalous position should be compared in the original video. For this purpose, the anomaly confirmation module is used. Here, the reference anomaly positions from the background are compared with detections in the original video. If a detection in the foreground matches with the background, the frame score of the corresponding frame in the foreground is incremented. The matching between two detection boxes in foreground and background is calculated using intersection over the union (IoU) between them. A sample matching detection are shown in figure 5.

Figure 5. Matching detections in Foreground and Background. Blue rectangles shows detections in original video and green rectangles indicate detections in background.

Once the frame scores are calculated for a given video, it is processed using post processing techniques to confirm the anomaly. All stationary vehicles may not constitute an anomaly. The vehicles parked in private grounds and those waiting for green signal are not anomalies. Vehicles in traffic signals also come into the background even though they are stationary for short period of time. All such detections are eliminated by applying morphological operations in the frame score. Now, the frame scores are processed to get the first frame for continuous detections. This is then handed over to the anomaly siever stage.

The Anomaly siever writes the first fraction of anomalies separated in the process into the anomaly text file. The video files in which anomalies are not detected by Normal extractor are scrutinized further to form a secondary list of anomaly candidates. In this stage, all video files in which there are no matching detections between foreground and background are added to the list of anomaly candidates. In some video files, there are no detections in background. Such videos are also added to the list. This list is passed to the Zoom extractor stage.

3.2. Zoom extractor

The challenge data set consists of videos in which anomalies arise from vehicles stalled at the farther end of roads. The size of the anomaly area is as small as 8×8 pixels in some cases. The detection of such small-sized vehicles from normal video frames is extremely challenging for the detector. To overcome this difficulty, a Zoom extractor is included. The images are zoomed and detections are run on these magnified images. Figure 7 shows the block

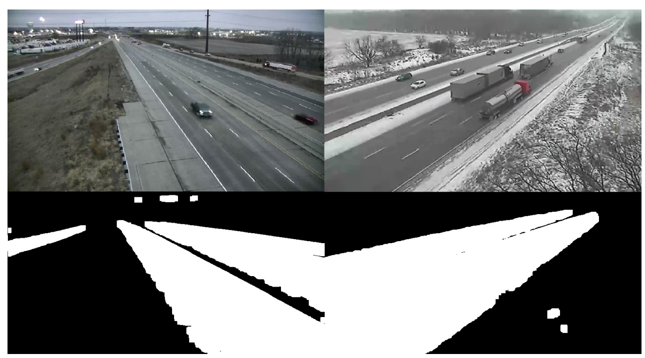

Figure 6. Road Masks generated using L)GMM R) Saliency

level architecture of the Zoom extractor. The secondary anomaly candidates from the Normal extractor is utilized in this stage. Only those videos in the candidate list alone are probed further in this module. Here the extracted frames from the videos are fed to three separate units.

The foreground extractor which is based on the GMM technique extracts the foreground objects. In traffic videos, all foreground objects correspond to moving vehicles. After employing morphological and binarization operations on the foreground image, this is seen as white blobs on a dark background. A road mask is generated using a weighted moving average of these foreground objects, obtained from each frame of the video. In some of the videos, the illumination variations and zoom variations are very large. In such cases, instead of the trailing path of vehicles the foreground

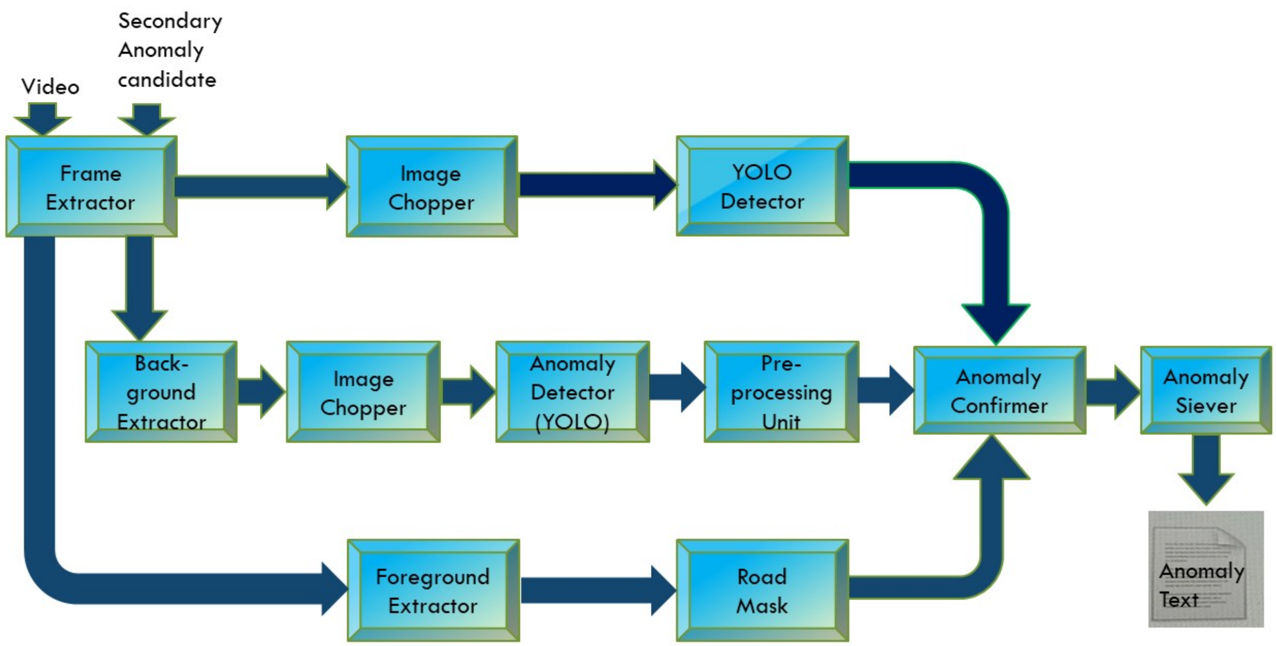

Figure 7. Block level representation of Zoom extractor

mask will be in white colour. To eliminate such discrepancies, we have adopted a saliency-based detector[19] to generate the road masks. This method shows good results compared to background subtraction, but the downside of this method is that it is computationally expensive. So we have employed this method only in the cases where GMM results in white masks. Figure6 shows the road masks generated, where continuous white regions are roads.

Background masks are generated in the same way as explained in section 2. Each mask is fed to an image chopping module. The module chops the input image into 8 non-overlapping blocks, two rows and four columns as shown in Figure8. These blocks are mutually exclusive and collectively exhaustive in nature. A single 800×410 sized image will now get changed into eight, 200×205 blocks. Each block is passed to the YOLO detector. The zooming operation is done inside YOLO. The detector internally resizes each crop into 416×416 and also preserves aspect ratio while zooming. This is achieved by zero-padding after zooming. A preprocessing step follows every detections in each background image block and eliminate possible false positives. This is followed by a confirmer module where information from road masks are available. Detections from private parking lots are eliminated using road masks. The term $Qualityfactor$ defined as

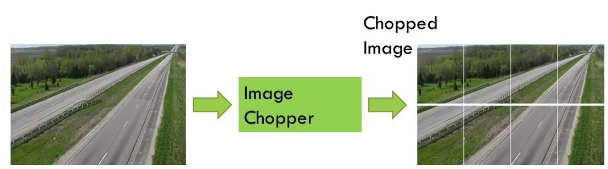

Figure 8. Image Chopper

$$Qualityfactor = \frac{BB_{Area}\ Overlap}{Total\ BB_{Area}}$$

decides whether a detection in background should be considered as a valid detection or not. Here, $BB_{Area}\ Overlap$ is the area of detection box in background that overlaps with the white region in the road mask and $Total\ BB_{Area}$ is the area of detection box in background.

Each frame from the original video is also passed through an image chopper where it gets divided into 8 pieces as explained above. Detections are run on each image block and the information is sent to anomaly confirmer. The rest of the workflow is similar to the Normal extractor. When a match between detections in original frame crop and its corresponding background tile occurs, the frame score of that frame is incremented. Post processing techniques are employed to remove false positives and first frame of continuous detections is passed to anomaly siever . The confirmed anomalies in this stage are appended to the anomaly text file as the second fraction of anomalies.

4. Experiments and Results

In this section, we present the experimental validation of the proposed Fractional Distillation method used for traffic anomaly detection. The proposed method is tested on track 4 test data of AI City Challenge 2020. The method did not use any additional datasets for training the model, but used pre-trained weights of YOLOv3 [16] on COCO Dataset[11]. The pre-trained weights used are openly available under the ImageAI Object Detection API[15].

4.1. DataSets and pre-trained weights

4.1.1 AI City Challenge Track 4 Dataset

The dataset contains 100 training and 100 test videos. Each video is approximately 15 minutes in length and has been recorded at $30fps$ and 800×410 resolution. Anomalies can be due to car crashes or stalled vehicles. Our method is unsupervised, implying that no information regarding whether a video contains an anomaly or not, is provided during training. Also there is no additional annotation made on the train dataset for this method.

The uniqueness of the data set in this challenge is that the anomalous region appears to be very small in the video. In many cases, the anomalous regions are of size less than 16×10 pixels, which is less than 0.03% of the total area of the video frame.

4.1.2 Pre-trained weights on COCO Dataset

The object detection model YOLOv3 has been used along with the API provided by ImageAI which included publicly available[13], pre-trained weights on the COCO 2014 Train Dataset.

4.2. Evaluation Metrics

The evaluation metrics of the challenge are F_1-score, root mean square error (RMSE) of detection time and S_4-score. For evaluating the F_1 score, a detection is considered true positive (TP) only if the anomaly is detected within 10 seconds from the interval between the onset of the anomaly and the end of the anomaly. F_1-score is the harmonic mean of precision and recall. A normalized RMSE (NRMSE) is obtained by min-max normalization from 0 to 300. RMSE score above 300 is normalized to 1. S_4-score is evaluated as :

$$S_4 = F_1 * (1 - NRMSE)$$

4.3. Experimental Settings

The GMM Background Extraction and the Road Mask Creation (both GMM based and Saliency) was run on a standard laptop with 2.4 GHz i5-9300H CPU and 4 GB GeForce GTX 1650 GPU and the object detection was run on a Google Colab Cloud instance with 2.3 Ghz Xeon Processor and 12 GB Tesla K80 GPU. The background extraction took about 45 hours for the 100 videos and the YOLO object detection took about 36 hours on the original 100 videos, the object detection on the cropped videos took an average of 34 hours for 27 videos. Our method has the object detection information encoded into text files. This facilitates the background subtraction and object detection to be run in parallel and finally the information is processed to find the anomalies.

4.4. Results

YOLO detections on the extracted background provide the positions of the possible anomalies in the video. A heuristic search on the original video on these positions help to find overlapping detections between foreground and background. These overlapping detections give us the anomaly frames in the video. The first frame in the continuous sequence of the anomaly frames gives us the start time of the anomaly. Our approach passes the input video to the normal anomaly extractor discussed above. The videos undetected in the Normal extractor are passed to the Zoom extractor. The experimental results from Normal and Zoom extractors are presented below.

4.4.1 Normal Extractor

Figure 9 shows the detection of the stalled vehicle in the extracted background video in the top and the bottom image shows the corresponding matching detection in the original video. We also introduce the concept of anomaly score for each frame in the video, which is the normalised sum of the IoU for the size and position overlapping detections in the extracted background and the original video.

The normalised frame score gives a clear idea about the position of the anomaly frames in the video. The sequence of continuous frames with an anomaly score greater than a threshold value is considered to be an anomaly location in the video and the first frame in this continuous set gives the time of start of the anomaly. To remove false positives and to find out the sequence of continuous sequence of anomaly frames, operations similar to erosion and dilation were performed on the normalised and thresholded frame score 1-D array. This operation ensures that false positive overlapping detections are removed and the obtained anomaly frames are continuous.

Figure 10 shows the plot of the frame-score for the anomaly video shown in Figure 9 , the top plot shows the

Figure 9. Overlapping detections on Background Video(top) for Original Video (bottom) for Video 1 in test data.

normalised frame-score of the video and the bottom plot shows the frame-score after performing the thresholding and the erosion-dilation operations explained above. From the final plot, the anomaly start time for the video can be found out directly.

4.4.2 Zoom extractor

Figure 11 shows the detection of the stalled vehicle in the cropped regions of extracted background video in the top and the bottom image in Figure 11 shows the corresponding matching detection in the original video. The original frame is cropped into eight: four horizontal and 2 vertical. These individual crops are passed separately to the object detection framework.

In the case of detections on the cropped videos, only the corresponding crop of the original video is checked for an overlapping detection. If a prospective anomaly candidate bounding box is found in crop 0 (top left) of the background video, then the search algorithm searches only on the detections on crop 0 of the original video which reduces the overall search latency.

Figure 12 shows the plot of the frame-score for the anomaly video shown in Figure 11 , the top plot shows the normalised frame-score of the video and the bottom plot shows the frame-score after performing the thresholding and the erosion-dilation operations.

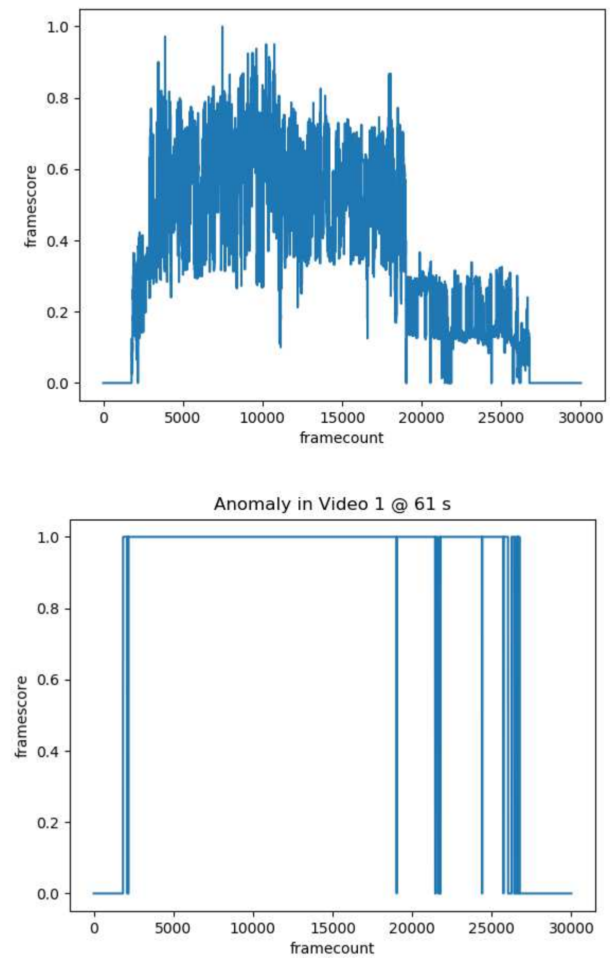

Figure 10. Frame Anomaly Score Plot for Video 1 in test data , normalized(top), thresholded(bottom)

4.4.3 Combined Results

The Normal extractor does a good job in detecting anomalies due to vehicles which are of larger size. The number of false positives is less as well. The Zoom extractor is ideal for anomalies caused by vehicles which are very small and not easily detected by traditional object detection models. The Zoom extractor produces some false positive results which reduces the F_1 score of the method. However, adding the road mask gives the best score on the combined extractor.

Table 1 gives a comparison on the results of the individual anomaly extractors discussed above on the AI City 2020 Challenge Track 4 Test Dataset, the results are obtained directly from the AI City Challenge Evaluation Server. Table 2 shows the performance of our method in the leader board of AI City challenge.

Figure 11. Overlapping detections on joined crop Background Video(top) for Original Video (bottom) for Video 20 in test data.The white lines show the lines along which the image was cropped.

Method	F_1 Score	RMSE	S_4 Score
Normal **(N)**	0.5200	51.4494	0.4308
(N)+Zoom **(Z)**	0.6316	50.7019	0.5248
(Z)+Mask	**0.7018**	**67.5044**	**0.5438**

Table 1. Comparison of Results on Different Methods

Rank	Team Name	S_4 Score
1	Firefly	0.9695
2	SIS Lab	0.5763
3	**CETCVLAB**	**0.5438**
4	UMDRC	0.2952
5	HappyLoner	0.2909

Table 2. Final Ranking and Score on Track 4 . Our team is shown in Bold

5. Conclusion

In this paper we have proposed a novel Fractional Data Distillation scheme for distilling out anomalies in traffic videos. The model follows the approach of fractional distillation and utilizes two anomaly extractors, namely Normal and Zoom. The former works on normal scale of the video, while the latter works on a magnified scale of the videos missed by the former to separate the anomalies. The Anomaly Confirmer module employed in these extractors helps in estimating the anomaly start time. The improvement in the F_1 score confirms that our distillation scheme is able to sieve a good number of anomalies with less false negatives.

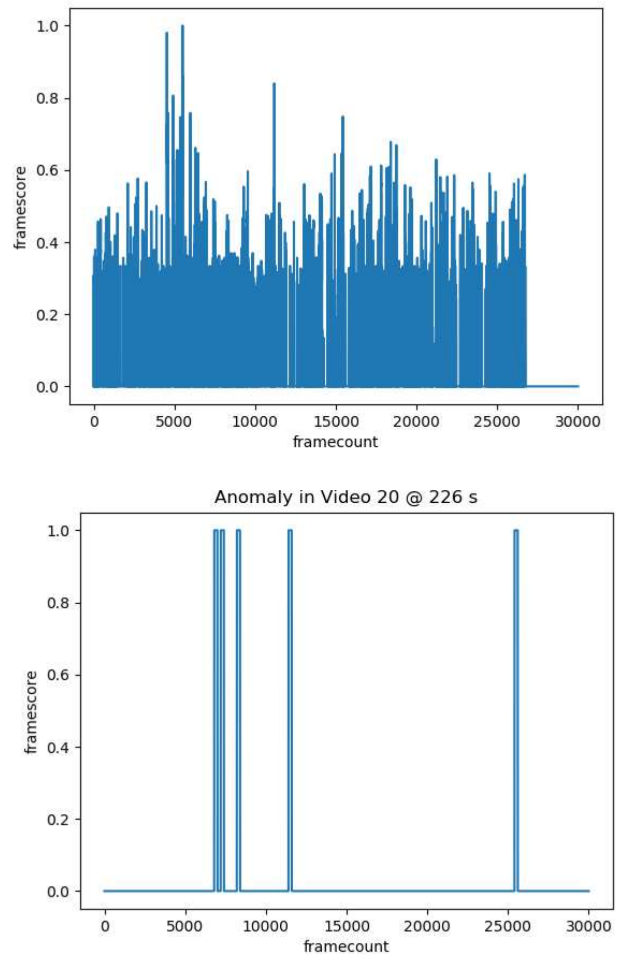

Figure 12. Frame Anomaly Score Plot for Video 20 in test data , normalized(top), thresholded(bottom)

6. Acknowledgements

The authors would like to thank all the people who supported and inspired us to continue our work in these lockdown times: Aravind Krishnan, Puthuserry Denson, Dr Ranjith S Kumar, Srichandini, Thomson M A, Nevin M T and all the members of Computer Vision Lab at CET. Special thanks to Google for providing Colab GPU instances, without which the work would not have materialized.

References

[1] The AI City Challenge. `https://www.aicitychallenge.org`.

[2] YOLOv3. `https://pjreddie.com/darknet/yolo/`.

[3] Shuai Bai, Zhiqun He, Yu Lei, Wei Wu, Chengkai Zhu, Ming Sun, and Junjie Yan. Traffic anomaly detection via perspec-

tive map based on spatial-temporal information matrix. In *Proc. CVPR Workshops*, 2019.

[4] Kuldeep Marotirao Biradar, Ayushi Gupta, Murari Mandal, and Santosh Kumar Vipparthi. Challenges in time-stamp aware anomaly detection in traffic videos. *arXiv preprint arXiv:1906.04574*, 2019.

[5] Yong Shean Chong and Yong Haur Tay. Abnormal event detection in videos using spatiotemporal autoencoder. In *International Symposium on Neural Networks*, pages 189–196. Springer, 2017.

[6] Rensso Victor Hugo Mora Colque, Carlos Caetano, Matheus Toledo Lustosa de Andrade, and William Robson Schwartz. Histograms of optical flow orientation and magnitude and entropy to detect anomalous events in videos. *IEEE Transactions on Circuits and Systems for Video Technology*, 27(3):673–682, 2016.

[7] Li Fei-Fei, Jia Deng, and Kai Li. Imagenet: Constructing a large-scale image database. *Journal of vision*, 9(8):1037–1037, 2009.

[8] Zhouyu Fu, Weiming Hu, and Tieniu Tan. Similarity based vehicle trajectory clustering and anomaly detection. In *IEEE International Conference on Image Processing 2005*, volume 2, pages II–602. IEEE, 2005.

[9] Santhosh Kelathodi Kumaran, Debi Prosad Dogra, and Partha Pratim Roy. Anomaly detection in road traffic using visual surveillance: A survey. *arXiv preprint arXiv:1901.08292*, 2019.

[10] Yanshan Li, Weiming Liu, and Qinghua Huang. Traffic anomaly detection based on image descriptor in videos. *Multimedia tools and applications*, 75(5):2487–2505, 2016.

[11] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick. Microsoft coco: Common objects in context. In *European conference on computer vision*, pages 740–755. Springer, 2014.

[12] Wen Liu, Weixin Luo, Dongze Lian, and Shenghua Gao. Future frame prediction for anomaly detection–a new baseline. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 6536–6545, 2018.

[13] Moses and John Olafenwa. Imageai, an open source python library built to empower developers to build applications and systems with self-contained computer vision capabilities, mar 2018–.

[14] Milind Naphade, Zheng Tang, Ming-Ching Chang, David C Anastasiu, Anuj Sharma, Rama Chellappa, Shuo Wang, Pranamesh Chakraborty, Tingting Huang, Jenq-Neng Hwang, et al. The 2019 ai city challenge. In *CVPR Workshops*, 2019.

[15] OlafenwaMoses. Yolo v3 pretrained weights on coco. https://github.com/OlafenwaMoses/ImageAI/releases/tag/1.0/, May 2018.

[16] Joseph Redmon and Ali Farhadi. Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767*, 2018.

[17] Linu Shine, Anitha Edison, and CV Jiji. A comparative study of faster r-cnn models for anomaly detection in 2019 ai city challenge. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 306–314, 2019.

[18] Waqas Sultani, Chen Chen, and Mubarak Shah. Real-world anomaly detection in surveillance videos. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 6479–6488, 2018.

[19] B. Wang and P. Dudek. A fast self-tuning background subtraction algorithm. In *2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 401–404, 2014.

[20] Gaoang Wang, Xinyu Yuan, Aotian Zhang, Hung-Min Hsu, and Jenq-Neng Hwang. Anomaly candidate identification and starting time estimation of vehicles from traffic videos. In *AI City Challenge Workshop, IEEE/CVF Computer Vision and Pattern Recognition (CVPR) Conference, Long Beach, California*, 2019.

[21] Tian Wang, Meina Qiao, Yingjun Deng, Yi Zhou, Huan Wang, Qi Lyu, and Hichem Snoussi. Abnormal event detection based on analysis of movement information of video sequence. *Optik*, 152:50–60, 2018.

[22] Zoran Zivkovic. Improved adaptive gaussian mixture model for background subtraction. In *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.*, volume 2, pages 28–31. IEEE, 2004.