

Video Data Extraction and Processing for Investigation of Vehicles' Impact on the Asphalt Deformation Through the Prism of Computational Algorithms



Sabahudin Vrtagić^{1,2*}, Edis Softić³, Mirza Ponjavić², Željko Stević⁴, Marko Subotić⁴, Aditya Gmanjunath¹, Jasmin Kevric²

¹ Faculty of Engineering and Natural Sciences, American University of the ME, Street 250, Block 6, Egaila 54200, Kuwait

² Faculty of Engineering and Natural Sciences, International Burch University, Francuske revolucije bb, Ilidža 71210, Sarajevo, Bosnia and Herzegovina

³ Technical Faculty, University of Bihać, Irfana Ljubijankića bb, Bihać 77000, Bosnia and Herzegovina

⁴ Faculty of Transport and Traffic Engineering, University of East Sarajevo, Vojvode Mišića 52, Doboj 74000, Bosnia and Herzegovina

Corresponding Author Email: sabahudin.vrtagic@stu.ibu.edu.ba

<https://doi.org/10.18280/ts.370603>

ABSTRACT

Received: 17 August 2020

Accepted: 12 November 2020

Keywords:

Histogram of Oriented Gradients (HOG), machine learning, Support Vector Machines (SVM), video processing, asphalt deformation

There are numerous algorithms and solutions for car or object detection as humanity is aiming towards the smart city solutions. Most solutions are based on counting, speed detection, traffic accidents and vehicle classification. The mentioned solutions are mostly based on high-quality videos, wide angles camera view, vehicles in motion, and are optimized for good visibility conditions intervals. A novelty of the proposed algorithm and solution is more accurate digital data extraction from video file sources generated by security cameras in Bosnia and Herzegovina from M18 roadway, but not limited only to that particular source. From the video file sources, data regarding number of vehicles, speed, traveling direction, and time intervals for the region of interest will be collected. Since finding contours approach is effective only on objects that are mobile, and because the application of this approach on traffic junctions did not yield desired results, a more specific approach of classification using a combination of Histogram of Oriented Gradients (HOG) and Support Vector Machines (Linear SVM) has shown to be more appropriate as the original source data can be used for training where the main benefit is the preservation of local second-order interactions, providing tolerance to local geometric misalignment and ability to work with small data samples. The features of the objects within a frame are extracted first by standardizing the feature variables and then computing the first order gradients of the frame. In the next stage, an encoding that remains robust to small changes while being sensitive to local frame content is produced. Finally, the HOG descriptors are generated and normalized again. In this way the channel histogram and spatial vector becomes the feature vector for the Linear SVM classifier. With the following parameters and setup system accuracy was around 85 to 95%. In the next phase, after cleaning protocols on collected data parameters, data will be used to research asphalt deformation effects.

1. INTRODUCTION

A novel way of different engineering fields working together on the roadway pavement deterioration problem due to traffic load is an important aspect of the road safety as all the research until now are based on the simulated scenarios which are insufficient as the number of vehicles increases drastically day by day. Namely, with combination of electrical, computer and civil engineering, a new type of video data collection as a solution to asphalt deformation problem can be obtained.

Object detection and location in digital images has turned out to be one of the most significant applications for businesses to ease client, spare time and to accomplish parallelism. Various terms can be linked to object detection like are Edge matching, Gray scale matching and gradient matching. Typically, object classification approach is based on shape, motion, colour, and texture [1].

Generally, vehicle background detection is accomplished

by using a plain feature extraction like: taillight standing-out method, adaptive thresholding, centroid recognition, and taillight coupling algorithm [1]. An example is a Gaussian Mixture Based foundation model with a lot of frontal area models of item size, position, speed, and shading appropriation in which every pixel in the scene is 'clarified' as either foundation, or as clamour. A projective ground-plane change is utilized inside the closer view model to reinforce item size and speed consistency suppositions. An educated model of common street travel bearing and speed is utilized to give an earlier gauge of object speed, which is utilized to initialize the speed model for each of the frontal area objects [2].

Another used approach is to distinguish the vehicles on the interstates depending on the picture handling strategies, for example, foundation subtraction, Prewitt channel and different morphological activities. Shortly by utilizing versatile foundation subtraction strategy to extricate the moving objects of the present casing, utilizing morphological activities to expel the random items to the vehicles where test outcomes

demonstrate the high exactness of the projected strategy alongside low time unpredictability [3].

In this research, a specific algorithm design for obtaining an unambiguous set of data that will be used for further study of the vehicles' impact on the asphalt deformation is achieved by Linear SVM classifier. Data parameters of interest in our research are the following:

1. To detect vehicles in frame
2. Vehicles size
3. Track vehicles
4. Detecting the set of speeds of the identified vehicles (km/h) in a video sequence
5. Time intervals for region of interest
6. Export collected data as csv file

Given data parameters are selected based on the correlations of given parameters and their impact on asphalt deformation along its surface and other layers [4].

The main contribution achieved by the following work is a novel concept solution aiming at more precise digital data extraction for foggy/day/night intervals from video file sources by extracting HOG features from mentioned intervals for the Linear-SVM training and ability of a framework to choose a trained model with the best fit.

The paper is organized into several sections. In the introduction, the way of machine learning usage in a specific data collection to solve a civil engineering problem is introduced. The second section contains the background research on object detection and reasoning on why SVM method is selected. In methodology section a deception of raw data, unambiguous problem, and the implementation of HOG and SVM for data collection is presented. In the result and conclusion section algorithm accuracy, data set results and future work are described.

2. BACKGROUND

Haar cascades are very popular for vehicle detection demonstrating diverse vision-based vehicle location frameworks for machine learning based systems to identify vehicles from video. The Haar highlight-based arrangement is quicker and equipped for identifying vehicles at a quickly evolving condition. Trial results demonstrate that the precision is above 90% [5].

Typically vehicle background detection is accomplished by using evident features which are collected. Features like rear light detection, centroid detection, adaptive threshold and rear light coupling algorithm. Moving vehicle detection depends on mentioned feature, classifications are being made by using rear-view of vehicle as presented by Sun et al. [6], and method by which this is achieved is use of Gabor filters for vehicle feature collection and Support Vector Machines for vehicle classification. Features like area, centroid and aspect ratio constraints are figured and given per entity and describe the state vector for the SVM. The output of SVM is basically the signed distance of the test instance from the separating hyper plane where given output specifies whether the evaluated entity corresponds to a vehicle's light. The recognition algorithm extracts the luminous entities at night using thresholding and subdivision approach and appraisals diverse entities. The properties of detected luminous object like area, aspect ratio, symmetry, distance are used to train the classifier for needs of classification. Aside from negligible false positives, results that are found are 100% in favour of

detection [6].

A similar way for detection is by rear-lamp pairs in a forward-facing colour video. If system hardware was a choice, a classic low-cost camera with a corresponding metal-oxide semiconductor sensor and Bayer RGB colour filter can be used for full-colour image display. Colour threshold is directly derivative from automotive regulations and modified as per the hue-saturation-value (HSV) colour space [7]. Lights are matched by colour cross-correlation symmetry and map out using Kalman filter. A postulation that the goal vehicle lights are symmetrical generally holds for same lane or for distances greater than 15m, same logic holds if in the adjacent lane on a straight fragment of the road. In some cases, such as, during sharp turns, turns, interchanges, and the latter stages of passing manoeuvres we can say the logic is not applicable. Perspective projection calculation with a pin-hole camera model was applied to estimate distance to or from object. To enhance the system, by improving the resilience to distortions of symmetry, a tracking-based detection algorithm has been implemented [7]. Similar on-road machine learning based method with SVM is suggested for vehicle detection and acceptable results are observed demonstrating the effectiveness of proposed system, as well as, other benefits for vehicle detection and classification, collision notice and beam ray status regulation [8].

Non-maximum Suppression is a method known for getting a full scan of image to remove any undesirable pixels which are not establishing edges [9]. Hysteresis thresholding phase clears out which edges are accepted as edge and which are not. Parameters that are among the used thresholds must be categorized as edges or false-edges based on their connectivity, provided that they are linked to "sure-edge" pixels, they are considered to be part of edges. When all above are put in a single HOG function [9], Linear SVMs is being trained on HOG entities [10].

One of the facts and reasons why the SVM was selected can be seen from the following results where a model to analyse two diverse grouping strategies with K-Nearest-Neighbor (KNN) and Support-Vector-Machine (SVM) is applied. In which a main benefit of SVM classification is that SVM has superior performance on datasets having multiple attributes, even when there are insufficient accessible data samples for the training process. However, some weaknesses of SVM classifications are felt in its resource intensive computational requirements, which are seen both during training phase and testing phase. Additionally, performance is subject to kernel function parameters being used. The overall accuracy rate for the SVM classifier was 92%. Unlike KNN, SVM classification was more proficient at vehicle classification which is why it is better in overall performance [11].

By implementation of Gaussian mixture model, the foreground analyses of the vehicle are obtained. The entities vector includes size, aspect ratio, width and solidity of the vehicle foreground blob. A Gaussian mixture model helps out to segment out moving vehicle blob. Firstly, colour recognition is applied, to reduce computation complexity, and to save memory space, making sure to lessen the impact on actual hardware system implementation. Suggested structure uses an 8-bin colour histogram as SVM vector. A 100% correct detection was obtained for main colours (green, red, blue, yellow) in the COIL database. Next step is type classification where the type vector components like size, aspect ratio and solidity of the foreground (vehicle) form blob [12].

Some models use collections of vehicle images analysed so that representative Eigen spaces can be obtained. Model build with Eigen spaces for Support Vector Machine (SVM) will become a consistent classifier for vehicle detection. For night conditions however, SVM method has been used. The results are confirmed to be robustly accurate, equally in speed and accuracy for vehicle classification. Important property of Eigenvectors is invariance in direction during a transformation, which can be used as a representation set of the whole big dataset. Those modules are called Eigen faces in face detection use [11] and Eigen vehicles in vehicle detection application [8]. Research has demonstrated that a small set of the eigenvectors is sufficient to shape up the whole image characteristic. In system that has been tested, top "P" eigenvectors are unchanging. Number of important entities from the vehicle Eigen space is being represented by "P". In the field of performance, given framework provides much higher classification rate (around 94%) than the other two (55% for Line Segment Detector LSD and 70% for Artificial Neural Network ANN), this demonstrates the effectiveness of this optimized classifier for vehicle images during night-time traffic conditions [13].

Another approach for night-time traffic scenes is with the postulation that the pictures of an item under fluctuating enlightenment circumstances lie in an arched cone framed in the picture space. Likewise, varieties because of changes in light intensity can be annulled by normalizing pictures. In view of these perceptions, proposed technique joins binary classifications utilizing discriminant hyper planes in the standardized picture space. It is shown that the problem results in a combination of binary classifications using hyper planes in the (1-D) dimensional normalized image space. It is experimentally confirmed that the use of hyper planes for normalized image space is effective, and that SVMs are better suited than Fisher's linear discriminant FLDs for obtaining discriminant hyper planes [14].

To adjust to various attributes of vehicle presence at hours of daylight and of the night, four signs including underneath, vertical edge, symmetry, and taillight are melded for the former vehicle discovery [15, 16]. Particle Filter initiating sampling, propagation, observation, cue fusion and evaluation accurately generates the vehicle distribution. In this way, the proposed framework can effectively identify vehicles follow and be powerful to various lighting conditions. In the night time, taillights luminosity is the most significant entities for a vehicle [15, 16]. In addition, a high-level following approach with BSAS clustering method is selected to deal with the case when the scene is under distortion for a short period [17]. Alternative method is the use of Gabor filters which provide a contrivance for managing degrees of invariance to concentration due to global illumination that are selectivity scaled, and have selectivity in orientation. Profoundly, they are alignment and scale tuneable edge and line identifiers. Vehicles contain solid edges and outlines that may differ in the alignment and scales, thus, the statistics of these entities (e.g., mean, standard deviation and skewers) could be very dominant for vehicle detection. Associating the performance of NNs with SVMs, SVMs outclassed NNs using either principal component analysis (PCA) or Gabor entities. In one specific case, the result of SVM classifier had reached about 9% greater accuracy than the NN with PCA entities and 11% greater accuracy in concurrent use of Gabor entities. Relations of SVM sizes, the average number of support vectors by Gabor method entities was 200, which is 1441 less than using PCA

entities. These capitals that SVM founded vehicle detection using Gabor entities are fast [18].

It can be observed that the most difficult part represents the night-time detection. As almost all frameworks and proposals are dealing with vehicle in motion and some of them is even implemented from a host vehicle in motion, an important knowledge regarding performance of different filters and different type of vector generation is noted. As the data in our case includes day and night video recordings, image matching will be used to distinguish day and night. As intensity of light varies during the night, more training and a specific data setup is applied for better accuracy. Regularization Parameter (C) is deduced using heuristics to 3 and spatial size to (3, 3) with more samples collected from night videos.

3. METHODOLOGY

3.1 Data collection

An original data includes a raw video material with equal ratio of day and night samples of video files that are collected on M18 road from a surveillance camera and transferred on a hard drive. Transfer of video files was granted by city council as per the rules and regulation of law in Bosnia and Herzegovina. Important video material properties are: frame width 1280, frame height 960 and frame rate of 25 frames per second.

3.2 Briefing on the HOG and SVM classifier

The advantage of SVM is their application to Linear Separable Data, Non Linear Separable Data, and data set which has a non-linear boundary that can be linearly separable if projected to higher dimensions.

A mathematical model and formulation for SVM classification with a simple representation can be described with the following equations:

SV classification:

$$\min_{f, \xi_i} || f ||_K^2 + C \sum_{i=1}^l \xi_i y_i f(x_i) \geq 1 - \xi_i \quad (1)$$

for all i $\xi_i \geq 0$

SVM classification, Dual formulation:

$$\min_{a_i} \sum_{i=1}^l a_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l a_i a_j y_i y_j K(x_i, x_j) \quad 0 \leq a_i \leq C \quad (2)$$

for all i $\sum_{i=1}^l a_i y_i = 0$

where, variables ξ_i are called loose variables and they measure the error given at point (x_i, y_i) , with given training vectors presented as x_i and C as a regularization parameter. Mentioned inputs are formed with the help of the HOG which is used to form input features descriptors for training Linear-SVM where the following equations are used to construct the hyperplane and where K represents kernel function. More details on the HOG features descriptors formation are given in the following paragraphs. As the training of SVM becomes quite puzzling with the number of training points is increased, some useful methods for fast SVM training have been

recommended [19-21]. Given equations are important as the training procedure of the model is the key factor for accuracy.

Some other reasoning of the SVM is because the attempt with the contours and Haar classification tested on online video files with high resolution videos and on highways without any traffic jams, the following methods were not totally suited for this research case scenario. For example, one of the issues was that Canny Edge Detection (Finding contours) loses stationary objects and so re-detection of object with different IDs, so time interval and a speed calculation algorithm were giving errors for cars/objects that go to the rest. A partial solution was to check for a change in object position for small changes in Euclidian distance. Another problem was that any change on frames with respect to the background is detected, like cars signal light blink and, braking light turning on / off. By considering given issues, an approach of classification using Histogram of Oriented Gradients (HOG) and Support Vector Machines (Linear SVC) was tested and the performance has been positive and promising. The entities of the objects within a frame are extracted by first normalizing the entities variables and then computing the first-order gradients of the frame. In the next stage, constructing encoding that remains robust to small changes while being sensitive to local frame content is finalized. Finally, the frame is normalized once more and the HOG descriptors are collected. This becomes the entities vector with which will train the Linear SVM classifier.

The result of SVM and other methods of supervised learning are data input dependent, hence pre-processing of train data is an important step since these entities extracted from this data are fit onto the SVM model, this is exemplified by scaling all dimensions as $[-1,1]$. Having unstructured data, the use of RBF kernel is a good, however reweighing training instances is important only when data set is unbalanced. In case the data is not linearly separable, the RBF or other kernels are adopted and both cost and gamma using "Grid Search" algorithm for SVM is optimized. Before a more detailed explanation on HOG and SVM implementation, a brief introduction on the problem will be specified. As a specific road intersection is considered, a description of it can be seen from the bird view of intersection taken from Google (Figure 1). The region of interest is specified and circled for a better understanding and visualization of the aim that will be presented in this study. Collection of data at the intersection for the current research includes a specific region of interest, and a specific set of data needs to be collected.



Figure 1. Intersection view with marked region of interest

Additionally, the specified section of importance is shown in Figure 2, where aim is to detect and collect the speed of the car at the very entrance to the camera's view, detect if the vehicles will stop and the speed of passage below the traffic lights. In order to measure vehicle hold-ups in the specified region, the proposed solution will measure two time intervals. The time intervals are time to rest and total time for region of interest. Time and speed detection will help to calculate deceleration intensity. In addition to speed and time detection, a lane in which vehicle is detected is one of priorities too.

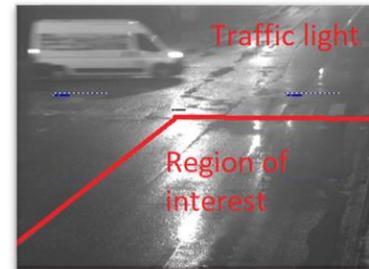


Figure 2. Region of interest on one of the frames

As explained before, the HOG and SVM will be used to obtain following parameters or data sets of interest in specified region. A HOG dataset will consist of parameters from positive and negative images. The positive images are the ROI with vehicles extracted from the source videos that can be perceived in Figure 3. On the other hand, the negative images are all that are not vehicles seen from any direction that is road, buildings, etc.



Figure 3. Positive and negative images for ROI



Figure 4. Hog descriptors of positive image

The Figure 4 represents pre-processed positive descriptors. By opening the images over 3 iterations followed by a single iteration of closing, will ensure that interesting features in the image are not lost, and will yield high quality features. Images containing single vehicles constitute the positive train data with a total of 2,954 samples, and images viewing the road and other static objects that are not vehicles constitute the negative train dataset as it is implied in Figure 5 and Figure 6 with a total of 2,860 samples make up the train dataset.



Figure 5. Hog descriptors of negative image (with zoom in, day time)



Figure 6. Hog descriptors of negative image (without zoom, night time)

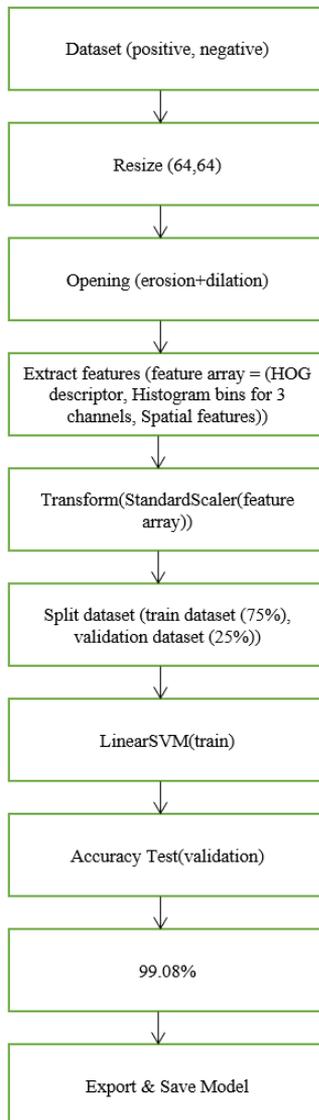


Figure 7. Training pipeline

Table 1. Configurations and functions

Positive Dataset – 2681
Negative Dataset – 2158
Original Size of the frame(w,h) – (1280, 960)
Resize dimensions – (64, 64)
Opening (erosion+dilation)–to eliminate the existing dotted boxes.
HOG descriptor – HOG (orientations=9, pixels per cell=8, cells per block=2)
Histogram Features for 3 channels (RGB) – 16 bins
Spatial features – image frame resized to (16,16)
Scaling (transformation) – remove mean and bring dataset to 0 mean and unit variance (you can write more on this) using StandardScaler()
Learning Algorithms – SVM(regularization parameter=10, loss="squared_hinge", penalty='l1')

These images are further resized to (64, 64) pixels. The feature Histogram of Oriented gradients vector of each image is obtained over 9 orientations, with 8 pixels per cell and 2 cells per block along with the histogram bins of the gradients and the spatial image is generated from resized image of (16, 16). The feature vector consists of the HOG descriptor, the histogram of the oriented gradients and the spatial image vector. Standard scaling the feature vector would remove centre the mean to 0 and scale the variance to 1.

The process explained previously for obtaining a feature vector can be summarized and visualized as shown in Table 1 and Figure 7.

4. RESULTS

A variation of SVM called a Linear SVM has been trained to distinguish between the objects of interest and those that are not. The previously generated feature vector was split into 70% and 30% as train and validation sets respectively, with the reported accuracy score of 98-100% on the test set. Given accuracy results are based on train and validation sets.

Support Vector Machines implementation steps for specific vehicle detection

- (1) Pre-processing (resizing 64,64 opening of the Figure)
- (2) Feature Extraction (HOG + Bins of HOG + spatial image (16,16))
- (3) Assigning labels (number of pos + negs)
- (4) Normalization (StandardScaler bringing the images to Unit variance and zero mean)
- (5) Training + params (LinearSVC, C=1, penalty = 'l2', loss = 'squared_hinge', results of the training precision and accuracy score)

As the video data includes parts that are zoomed in/out, a new detection and classification is done accordingly so that parameters for speed and time can be set properly. This must be done so that lines of cross for speed and time count are set properly for different view. In addition to zoom in /out a night-time, work regime is adjusted by the usage of machine learning supervised classifier so that frames lighting environment is adjusted accordingly. This needs to be adjusted so that a better accuracy is gained. Setup, functions and configurations for training are given in Table 2. In Figure 7, a training pipeline process for the feature vector is given in 10 main steps.

Once the model is trained and ready, applying sliding window concept across a region of interest with a classifier to check if the vehicle is detected in each region or not is

described in detail with the pipeline in Figure 8. Note that the size of car will vary depending on how far the car is with respect to the camera. Heat map is added for all cars detected in frame for which prediction is positive or 1 to keep only 1 window instead of multiple windows for one car. Also, some boxes will be rejected on the basis of calculated threshold values (values for pixels of 120 and below). With the following parameters and setup, system accuracy was around 85 to 95%. Errors occurring are mainly because of light reflection during the night-time from water accumulated at some region on the road. The resulting issue can be ignored as no speed can be detected for such detection. That means if no change in position for given reflections is generated, system will return 0 for given regions. Given in Table 2 is a part of data generated from original video sources. A time intervals and speeds can be observed where speed limit for specific road is given to be 50 km/h.

The time intervals of interest can be generated in two ways: by a built-in python function or with the usage of frames per second. Both ways of time generation result in very close values, so a python time function was taken for simplicity. The speed calculation is determined for a specific range of 250 pixels, by converting pixels to meters to obtain speed calculation in meters per second that will be converted to km/h as it is unit of interest.

The best reported accuracy score performance for HAAR cascade is 91.3% [22]. By comparing that to the best accuracy score of 99.08% on the test set with the proposed Linear-SVM framework, Linear SVM based on the extracted HOG features has much better detection rate for day intervals [22]. Created Linear-SVM framework overall accuracy for day, night and fog has been reported to be 85-95% due to the fog intervals which are used for the calculation of the overall accuracy rate. Additionally, the faster regions with CNN features (R-CNN) and You Only Look Once algorithm (YOLOv3) have reported accuracies above 99.5% for high video resolutions [23], but the key advantage of our framework is day, night, fog and low quality video resolutions detection with focus on detailed data extraction to be used for road erosion research.

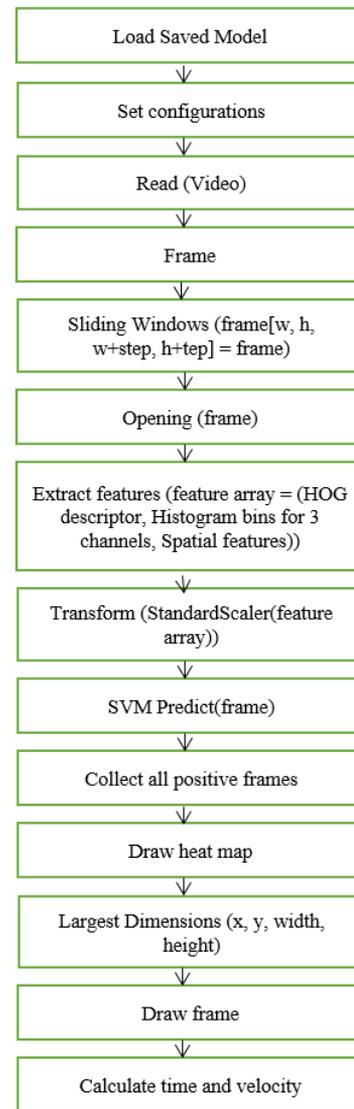


Figure 8. System detection and classification

Table 2. Video data sample

Vehicle ID	Orientation	L-lane/R-lane	Entrance Velocity (km/h)	Exit Velocity (km/h)	Time to exit the Junction (second)
10		L	46.18095	30.4022	0.779029
44		L	46.18095	10.64364	2.225199
81		L	39.36755	27.42365203	0.863641739
80		L	38.81374	37.13211928	0.637836218
79		L	40.57799	56.6620535	0.417990685
78		L	42.510283	8.682767506	2.727725983

5. CONCLUSION

With an original video from intersection and with the unique implementation of explained algorithm, all aims for data set are satisfied and ready to be implemented on a unique research that will combine multiple engineering branches to research vehicles impact on the asphalt deformation with the collected data. The following research implementation will reduce the cost of the asphalt erosion testing as the present testing needs a lot of time for implementation on data collections as they are based on a simulation. With presented research a more variety set of data in a shorter amount of time can be obtained to improve knowledge on vehicles' impact on the asphalt deformation. With the following setup, created system

accuracy was around 85 to 95%. Errors occurring are mostly because of light reflection during the night time from water accumulated at some region on the junction. The following issue is solved easily as no speed is detected because if no change in position for given reflections is generated and system returns 0 for given regions. Time intervals and speeds can be observed where speed limit for specific road is given to be 50 km/h. In addition to given data, time intervals and speeds will be used for calculation of deceleration parameter.

The data is separated into two main groups as there is a change in a video view (zoom in/zoom out), so a full data samples are obtained for the left lane while the right lane data sample is rated at 75%. Additionally the full data set is providing traffic flow intensity for a specific time of a day.

Full data analysis will be presented in future papers in which the vehicles' impact on the asphalt deformation, reasoning of traffic jam, street lights improvement for better traffic flow which will improve traffic safety. The data is already out there as just about each person today carries a smart phone and there are surveillance cameras at all places. All that is required is a software solution to ingest it and harvest actionable knowledge. Some available method is promising which proves that there are understandings to be expanded from existing data and that ensemble approaches are better suited than single classifiers when dealing with extremely mutable dynamic environments [24-26].

As an example with similar idea implementation on traffic flow improvements with machine learning and (CCTV) datasets with analysing of the road traffic flow outlines through fine-tuning pre-trained CNN prototypes on domain specific low quality images, as detained in several weather conditions and seasons [25]. This analysis can at last help the understanding of traffic interruption, during and after an outrageous climate condition occasion. It very well may be recommended that the adjusted MobileNet model with 98.2% accuracy, 58.5% review and 73.4% consonant mean might be utilized for a continuous traffic investigation with big data [27]. Given example is based on traffic flow improvements whereas purpose of this research is to collect data on traffic flow and to see the correlation with asphalt deformation.

As per the given details in paper the most suitable algorithm for specific data collection is Linear SVM. The main advantages are:

- Preparing and teaching model based on original data set extracted from video materials
- Parameter adjustments for better results
- Ability to detect object that are not in motion
- To improve results with more data

As the artificial intelligence and computing power is improving, various engineering fields will find the use and a new research fields will emerge. Examples like self-driving cars and smart road junctions will increase data collection, hence novel ways for traffic safety and traffic flow will be introduced. As a future work, a vehicle fluid consumption and environment effect could be obtained for an improvement in road construction planning from video data of road junctions and roundabouts.

ACKNOWLEDGMENT

This work is supported by the Municipality of Živinice, by providing necessary permission for data usage.

REFERENCES

[1] Gupta, B., Chaube, A., Negi, A., Goel, U. (2017). Study on object detection using open CV-Python. *International Journal of Computer Applications*, 162(8): 17-21. <https://doi.org/10.5120/ijca2017913391>

[2] Magee, D.R. (2004). Tracking multiple vehicles using foreground, background and motion models. *Image and vision Computing*, 22(2): 143-155. [https://doi.org/10.1016/S0262-8856\(03\)00145-8](https://doi.org/10.1016/S0262-8856(03)00145-8)

[3] Javadzadeh, R., Banihashemi, E., Hamidzadeh, J. (2015). Fast vehicle detection and counting using background subtraction technique and Prewitt edge detection.

International Journal of Computer Science and Telecommunications, 6(10): 8-12.

[4] Softić, E., Jusić, E., Morina, N., Dubravac, M. (2018). Selection, effectiveness and analysis of the utilization of cement stabilization. In: Avdaković S. (eds) *Advanced Technologies, Systems, and Applications III. IAT 2018. Lecture Notes in Networks and Systems*, vol 60. Springer, Cham. https://doi.org/10.1007/978-3-030-02577-9_20

[5] Choudhury, S., Chattopadhyay, S.P., Hazra, T.K. (2017). Vehicle detection and counting using Haar feature-based classifier. *2017 8th Annual Industrial Automation and Electromechanical Engineering Conference (IEMECON)*, Bangkok, pp. 106-109. <https://doi.org/10.1109/IEMECON.2017.8079571>

[6] Sun, Z., Bebis, G., Miller, R. (2006). On-road vehicle detection: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(5): 694-711. <https://doi.org/10.1109/TPAMI.2006.104>

[7] O'Malley, R., Jones, E., Glavin, M. (2010). Rear-lamp vehicle detection and tracking in low-exposure color video for night conditions. *IEEE Transactions on Intelligent Transportation Systems*, 11(2): 453-462. <https://doi.org/10.1109/TITS.2010.2045375>

[8] Sutar, V.B., Admuthé, L.S. (2012). Night time vehicle detection and classification using support vector machine. *IOSR Journal of VLSI and Signal Processing (IOSR-JVSP)*, 1(4): 1-9. <https://doi.org/10.9790/4200-0140109>

[9] Klimaszewski, J., Kondej, M., Kawecki, M., Putz, B. (2013). Registration of infrared and visible images based on edge extraction and phase correlation approaches. In: Choraś R. (eds) *Image Processing and Communications Challenges 4. Advances in Intelligent Systems and Computing*, vol 184. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-32384-3_19

[10] Bristow, H., Lucey, S. (2014). Why do linear SVMs trained on HOG features perform so well? *arXiv preprint arXiv:1406.2419*. <https://arxiv.org/abs/1406.2419v1>.

[11] Kim, J., Kim, B.S., Savarese, S. (2012). Comparing image classification methods: K-nearest-neighbor and support-vector-machines. In *Proceedings of the 6th WSEAS International Conference on Computer Engineering and Applications*, and *Proceedings of the 2012 American conference on Applied Mathematics*, pp. 48109-2122. <http://www.wseas.us/e-library/conferences/2012/CambridgeUSA/MATHCC/MATHCC-18.pdf>

[12] Chen, Z., Pears, N., Freeman, M., Austin, J. (2009). Road vehicle classification using support vector machines. In *2009 IEEE International Conference on Intelligent Computing and Intelligent Systems*, Shanghai, China, pp. 214-218. <https://doi.org/10.1109/ICICISYS.2009.5357707>

[13] Thi, T.H., Robert, K., Lu, S., Zhang, J. (2008). Vehicle classification at nighttime using eigenspaces and support vector machine. In *2008 Congress on Image and Signal Processing*, Sanya, Hainan, China, pp. 422-426. <https://doi.org/10.1109/ICDSP.2002.1028263>

[14] Okabe, T., Sato Yoichi. (2003). Object recognition based on photometric alignment using RANSAC. *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Madison, WI, USA, pp. 221-228. <https://doi.org/10.1109/CVPR.2003.1211357>

[15] Chern, M.Y., Hou, P.C. (2003). The lane recognition and vehicle detection at night for a camera-assisted car on

- highway. In 2003 IEEE International Conference on Robotics and Automation, Taipei, Taiwan, pp. 2110-2115. <https://doi.org/10.1109/ROBOT.2003.1241905>
- [16] Wang, C.C., Huang, S.S., Fu, L.C. (2005). Driver assistance system for lane detection and vehicle recognition with night vision. In 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems, Edmonton, Alta., Canada, pp. 3530-3535. <https://doi.org/10.1109/IROS.2005.1545482>
- [17] Chan, Y.M., Huang, S.S., Fu, L.C., Hsiao, P.Y., Lo, M.F. (2012). Vehicle detection and tracking under various lighting conditions using a particle filter. *IET Intelligent Transport Systems*, 6(1): 1-8. <https://doi.org/10.1049/iet-its.2011.0019>
- [18] Sun, Z., Bebis, G., Miller, R. (2002). On-road vehicle detection using Gabor filters and support vector machines. In 2002 14th International Conference on Digital Signal Processing Proceedings. DSP 2002 (Cat. No. 02TH8628), Santorini, Greece, pp. 1019-1022. <https://doi.org/10.1109/ICDSP.2002.1028263>
- [19] Burges, C.J. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2(2): 121-167. <https://doi.org/10.1023/A:1009715923555>
- [20] Cristianini, N., Shawe-Taylor, J. (2000). *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*. Cambridge University Press.
- [21] Vapnik, V. (1998). *Statistical Learning Theory*. Wiley, New York.
- [22] Atibi, M., Atouf, I., Boussaa, M., Bennis, A. (2015). Real-time detection of vehicle using the Haar-like features and artificial neuron networks. *Proc. Computer Science*, 73: 24-31. <https://doi.org/10.1016/j.procs.2015.12.044>
- [23] Benjdira, B., Khurshed, T., Koubaa, A., Ammar, A., Ouni, K. (2019). Car detection using unmanned aerial vehicles: Comparison between faster R-CNN and yolov3. In 2019 1st International Conference on Unmanned Vehicle Systems-Oman (UVS), Muscat, Oman, pp. 1-6. <https://doi.org/10.1109/UVS.2019.8658300>
- [24] Thomas, R.W., Vidal, J.M. (2017). Toward detecting accidents with already available passive traffic information. In 2017 IEEE 7th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, USA, pp. 1-4. <https://doi.org/10.1109/CCWC.2017.7868428>
- [25] Song, X.R., Gao, S., Chen, C.B. (2018). A novel vehicle feature extraction algorithm based on wavelet moment. *Traitement du Signal*, 35(3-4): 223-242. <https://doi.org/10.3166/TS.35.223-242>
- [26] Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8(6): 679-698. <https://doi.org/10.1109/TPAMI.1986.4767851>
- [27] Peppas, M.V., Bell, D., Komar, T., Xiao, W. (2018). Urban traffic flow analysis based on deep learning car detection from CCTV image series. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 42(4): XLII-4.499-506. <https://doi.org/10.5194/isprs-archives-XLII-4-499-2018>