



Data Mining in A Smart Traffic Light Control System Based on Image Processing and KNN Classification Algorithm

Abdullah Yusefi^{1*}, Adem Alpaslan Altun² and Cemil Sungur³

¹ Computer Engineering Department, Konya Technical University, Turkey

² Faculty of Technology, Selçuk University, Turkey

³ Electrical-Electronics Engineering Department, Konya Technical University, Turkey

(1st International Conference on Computer, Electrical and Electronic Sciences ICCEES 2020 – 8-10 October 2020)

(DOI: 10.31590/ejosat.819762)

ATIF/REFERENCE: Yusefi, A., Altun, A. A. & Sungur, C. (2020). Data Mining in A Smart Traffic Light Control System Based on Image Processing and KNN Classification Algorithm. *European Journal of Science and Technology*, (Special Issue), 461-465.

Abstract

In today's modern world, communication, transportation and the movement of people and merchandises are important, and doing so in the shortest possible time is also essential and vital. In the past decade, due to the significant increase in the number of passengers and vehicles along with the capacity limitations of communication arrays, it is absolutely necessary to apply new technologies to intelligent traffic control and management. The intelligent transportation system (ITS) utilizes advanced technologies in the fields of information processing, telecommunications and electronic control to meet transportation needs. The purpose of these systems is to streamline traffic in important and sensitive routes, and in addition to providing traffic safety, information, timely traffic control and the use of optimal capacity of transport arteries. This paper presents new method for extracting traffic parameters associated with a signalized highway using image processing and data mining KNN classification algorithm. These parameters include the length of red light LED, the volume of passing vehicles and the volume of pedestrians passing the highways in the green phase. In what follows, a Data Mining Traffic Light Control System is introduced, which by receiving the three traffic parameters mentioned above, proceeds to optimize the traffic signal timing. At the end, a two-phase common highway is simulated in the MATLAB software environment, and the results of the image processing algorithms and the Data Mining Traffic Light Control System designed for it are evaluated.

Keywords: Two-phase Thresholding, Blocking, Data Mining, Traffic Simulation, Classification, Face Detection, Vehicle Detection, KNN classification

1. Introduction

Today, many highways around the world are equipped with intelligent traffic light systems. One of the most common methods for measuring the highway traffic parameters is the use of inductive loop traffic detectors. Inductive loop traffic detectors are only used in the front of intersections or highways. For this reason, in addition to the proper functioning of the SCATS system in allocating timelines that are consistent with the instantaneous traffic conditions, it has a drawback that it lacks the direct and accurate calculation of the red line's queue length.

In this paper, a new method is proposed for extracting traffic parameters by applying image processing's face detection and vehicle detection capabilities. The extracted parameters are then processed by data mining algorithms to define the timelines of traffic lights. In image-based systems, the red line's instantaneous queue length can be obtained and taken into account in decision making. Even critical modes can be defined for the length of the queue, for example, if the queue length of an intersection reaches its adjacent intersection, the traffic light changes immediately. In the field of extraction of traffic parameters such as: the volume, the number of passing vehicles, the length of the queue and the traffic flow through the image processing, there have been many articles and much work in recent years. The parameters extracted from these studies are then used in traffic systems to control the traffic signals. There

* Corresponding Author: Konya Teknik Üniversitesi, Bilgisayar Mühendisliği Bölümü, Konya, Türkiye, ORCID: 0000-0001-7557-8526, e168129001005@ktun.edu.tr

have been different approaches proposed for tackling this problem. For example, in [3] using virtual loops, there is a way to count the vehicles passing the highway. In [5], a method for measuring the traffic flow using the concept of entropy is discussed. Here, flow is the number of vehicles passing an intersection during one hour. Also, in [4], by developing a new concept and simultaneous use of both the optimal area processing and perspective transformation, images are made as spatio-temporal images, and from these images various parameters such as speed, distance between vehicles and flow are obtained. Moreover, in [1], a method for detecting the cycle failure in a signalized intersection is presented. In this method, the end of the queue is followed by the red light after the light is green, and if a vehicle fails to illuminate the intersection during the green phase, this will be considered as a cycle failure. But it does not directly focus on measuring the red light queue length and changing the timing according to that. It only suggests that it is possible to extend the green time so that the queue of the red light is discharged and, in the case of this extension, the length of the red phase queue, which is currently red, is not considered. There have been different approaches proposed for tackling this problem. A hierarchical clustering data mining method is used for designing a TOD signal control system [10]. System signal detectors are used to collect the data which is the number of vehicles. In [2] digital image processing has been used to detect and count the vehicles. The proposed traffic control system in [8] is by installing inductive loops and controlling the traffic light and also detecting the vehicles which stop beyond the zebra crossing. However, none of these studies use the number of pedestrians and data mining classification technique for controlling the traffic lights.

However, none of the papers and works mentioned above has taken the number of pedestrians in to account as a parameter. Moreover, in studies done in literature old methods of inductive loops and sensors were used to control the traffic light system and they have not used the modern methods such as data mining techniques. In this paper, a new method for counting the number of vehicles in the red light queue, which is a special parameter of the intersections, and less work has been done on it, is given. Moreover, this paper takes the number of pedestrians waiting for the green light in account for calculating the timeline of green lights. In the case of counting vehicles passing through the zebra cross in the green phase, despite the similarity in general concepts, such as the use of the optimal area processing method used in most articles [3-5], an innovation has been given in the method used in this paper. Also, a new partitioning method based on the data mining techniques is developed in this paper which achieves a high degree of segmentation. In the following of the article, based on traffic parameters, a data mining control system is designed for a common two-phase highway cross, which takes into account the synchronous information of the green and red phases on the timing of the traffic light. The final section relates to the simulation carried out in the MATLAB software environment in which the control system designed for a common two-phase highway cross is evaluated. Of course, it's necessary to point out that the video is about the highway cross is recorded on a sunny day, and because image processing methods directly depend on the environmental conditions of the environment such as: rainfall conditions and nighttime condition, the algorithms have to be modified accordingly. We hope to address these issues in the next articles as well.

This paper is organized as follow: section 2 describes the use of image processing, face detection and vehicle detection methods used in this paper. In section 3 classification method is described. Section 4 reports the traffic control system's experimental results. The paper is concluded with section 5.

2. Image Processing

Traffic monitoring is one of the most important tools for developing intelligent traffic systems. Application of image processing techniques and machine vision in analyzing traffic video frames and converting them into statistical sequences for collecting traffic information is much more accurate and less costly than other method of collecting traffic information. Methods such as microwave detectors, light sensors and the use of inductive sensors and circuits will usually have problems like high installation and maintenance costs and blocking paths for vehicles during installation. Furthermore, when it comes to detecting slow-moving or stopped vehicles, these methods face some difficulties. Image processing has had dramatic advances in both theoretical and practical aspects, and many sciences depend on it. Various applications can be used for image processing algorithms, including industrial automation, medical image analysis, video and image compression, transport, meteorology, urbanization, industry, agriculture, military and security science, astronomy, medicine, antiquities Science, cinema, geology, food industry and even psychology. This paper uses two image processing techniques for extracting traffic parameters. First, face detection has been applied to detect the number of pedestrians awaiting the green light. Second, the system uses the vehicle detection to detect the number of vehicles in the street. These two parameter are later used for controlling the traffic signal statuses.

2.1. Face Detection

Face detection is a practical and important area of study that has made significant advances in recent decades. These Systems have numerous applications, including security and monitoring. In the last decade, major changes have taken place in the field of face detection that allows the use of these systems in every aspect of sectors. However, applying it in controlling traffic has not been done yet. In this section, the main parts of the system for detecting and verifying the faces and their technical details are described.

As it is known, the images recorded by traffic light cameras provide small and low quality images of pedestrians. These pedestrians stand side by side or behind each other next to zebra crosses in the highway streets. Since the people standing in the back are occluded by the people in front of the camera, it makes the counting pedestrians a difficult process. Due to this issue, human detection method is not a suitable choice for counting the waiting pedestrians. Therefore, we chose the face detection method for solving this issue.

In order to do this we used a pre-trained model of faces and then performed the classification of detected object in the image using this model. These objects are defined by clustering the input image which is based on the Jaccard distance d .

$$d(f_i, f_j) = 1 - J(f_i, f_j) \quad (1)$$

Here, d is the Jaccard distance, f_i and f_j are the shapes of objects and J is the standard Jaccard similarity. Figure 1 depicts the test image of pedestrians before applying the face detection and figure 2 shows the same image after detection of pedestrians. The results shows that this method can detect small faces in the images with high accuracy and is an appropriate method for the purpose of this paper.



Fig. 1. Pedestrians before applying face detection



Fig. 2. Pedestrians after applying face detection

2.2. Vehicle Detection

One of the most important applications of traffic surveillance images are to detect and extract traffic information from traffic images. In this paper, first, the range of the vehicles are determined using the background image, the image containing the vehicles and image filtering in RGB and grayscale models. So the difference between the background image and the image containing the vehicles are calculated and then the median filter is used to remove the noise from the image. After eliminating noise, the number of cars calculated and prepared to be used in this paper. The results show that the method is sufficiently accurate and acceptable. The performance of the vehicle detection technique is shown in figures 3 and 4.

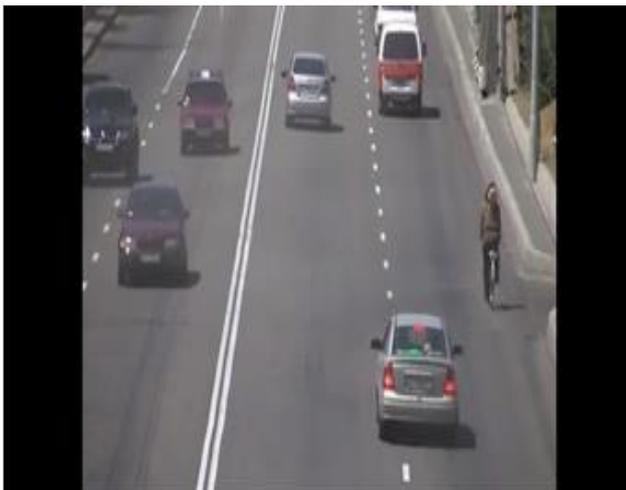


Fig. 3. Vehicles before applying vehicle detection

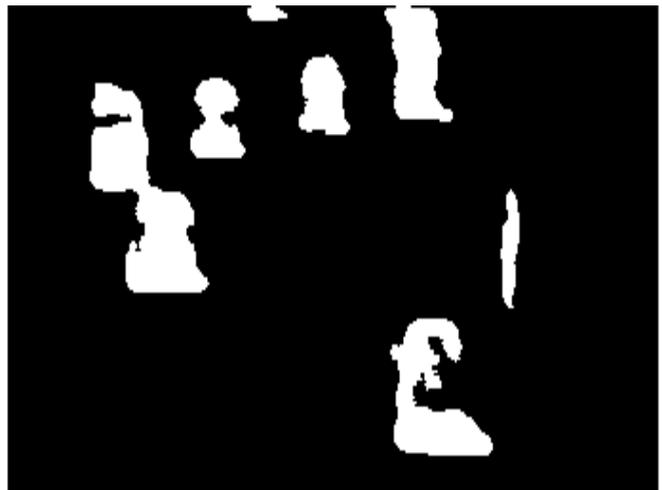


Fig. 4. Vehicles after applying vehicle detection

3. Data Mining Traffic Light Control System

With the rise of computer systems and the spread of information technology, the main topic in computer science has changed from how information is gathered to how information is used. Data mining systems allow the user to interpret the mass of data collected and extract the knowledge they contain. Today, it has attracted many researchers as one of the most important issues in the

field of artificial intelligence and databases. In this paper, one of the popular and robust data mining classification algorithms called k-nearest neighbor (KNN) algorithm is used for classifying the data collected from traffic surveillance cameras. The KNN classification algorithm is one of the best and most widely used classification algorithms, which is widely used in various sectors. This algorithm measures the distance between input data and trained data set and classifies the data based on the nearest neighbor to input data. The distance between two data is obtained formulas like Euclidean distance:

$$d_E(x, y) = \sum_{i=1}^N \sqrt{x_i^2 - y_i^2} \quad (2)$$

The algorithm used in this paper creates a model from trained data and classifies the number of detected faces and detected vehicles into classes based on that pre-trained model. The model has 4 classes PLVM, PMVL, PLVL, PMVM. These classes represent four possible statuses of traffic and pedestrians. The status are when the pedestrians are less but vehicles are many, when pedestrians are many but vehicles are less, when both pedestrians and vehicles are less, when both pedestrians and vehicles are many, respectively.

After classifying the input data, the traffic light mode changes according to the class selected for that status. In the PLVM status the traffic light will increase the green phase for vehicles and likewise decrease the red phase. Using this procedure vehicles will have more time to pass and the traffic congestion is prevented. In PMVL status, the traffic light will decrease the green phase and increase the red phase for vehicles. Since the number of vehicles are less in this status, there is no traffic congestion issue and thus many pedestrians waiting for the traffic light will have more time to pass. Furthermore, in PLVL status and PMVM status, the traffic light will slightly increase the green phase for pedestrians and give the pedestrians higher priority of passing, so that the number of pedestrians will decrease in the next iteration of classification process. In figure 5, the general flowchart of the proposed traffic light control system is depicted.

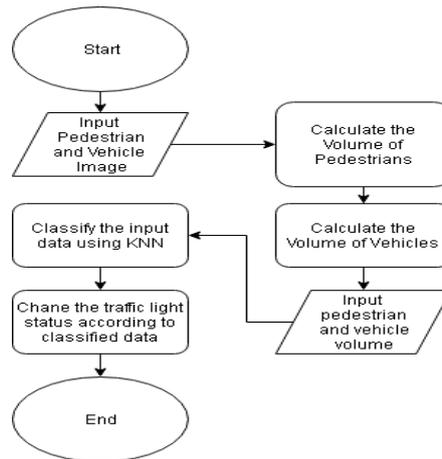


Fig. 5. Proposed Traffic Light Control System

4. Experimental Results

In evaluating the simulation results, one of the comparable criteria was the system error rate. This includes the error rate of Face detection, Vehicle detection techniques and finally classification algorithm. The model of trained data which have been used in KNN classification is shown in figure 6. This model has been built on 30 inputs and the KNN algorithm classifies the input parameters obtained from face detection and vehicle detection processes based on this model. Figure 7 depicts the test data after applying the KNN classification based on this model. The results of classification on test data shows a high percentage of accuracy.

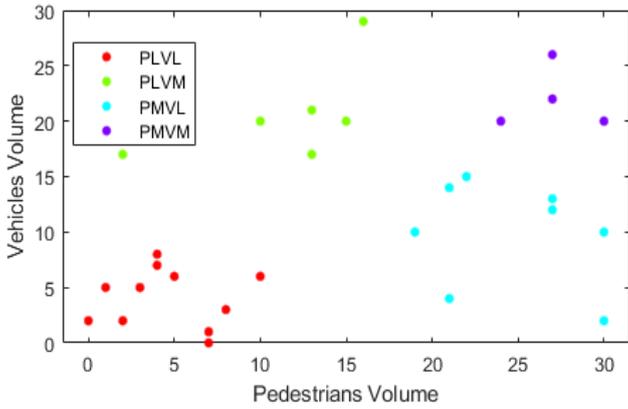


Fig. 6. Train Model after applying KNN Classification

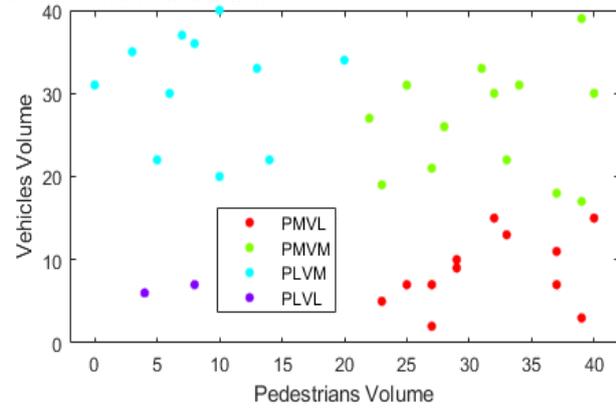


Fig. 7. Test Data after applying KNN Classification

5. Conclusion

In this paper, a new approach for controlling the traffic lights of highways was proposed. In order to do this, an image processing model was presented to detect the number of pedestrians and vehicles waiting for traffic lights. This model, first, detects and counts the number of pedestrians and then determines and counts the number of vehicles using efficient and accurate image processing techniques. Following that, the parameters obtained from face detection and vehicle detection phases are classified using KNN classification algorithm. This part classified the input parameters into 4 different classes. According to the result of the classification part the system controls the status of the traffic lights. The method proposed in this paper was tested on various inputs for which the results was outstanding and accurate.

References

- [1] Zheng, Jianyang, et al. "Detecting cycle failures at signalized intersections using video image processing." *Computer-Aided Civil and Infrastructure Engineering* 21.6 (2006): 425-435.
- [2] Reyes, Mac Michael. "Traffic Light Control System Simulation Through Vehicle Detection By Image Processing." (2008).
- [3] Wang, Kunfeng, et al. "An automated vehicle counting system for traffic surveillance." *Vehicular Electronics and Safety, 2007. ICVES. IEEE International Conference on. IEEE, 2007.*
- [4] Lee, Daeho, and Youngtae Park. "Measurement of traffic parameters in image sequence using spatio-temporal information." *Measurement Science and Technology* 19.11 (2008): 115503.
- [5] Bhaskar, Lala, et al. "Intelligent traffic light controller using inductive loops for vehicle detection." *Next Generation Computing Technologies (NGCT), 2015 1st International Conference on. IEEE, 2015.*
- [6] Hsu, W-L., et al. "Real-time traffic parameter extraction using entropy." *IEE Proceedings-Vision, Image and Signal Processing* 151.3 (2004): 194-202.
- [7] Sivakumar, R., et al. "Automated traffic light control system and stolen vehicle detection." *Recent Trends in Electronics, Information & Communication Technology (RTEICT), IEEE International Conference on. IEEE, 2016.*
- [8] Mirchandani, Pitu, and Larry Head. "A real-time traffic signal control system: architecture, algorithms, and analysis." *Transportation Research Part C: Emerging Technologies* 9.6 (2001): 415-432.
- [9] Hu, Peiyun, and Deva Ramanan. "Finding tiny faces." *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2017.*
- [10] Zamani, Zahra, Mahmoud Pourmand, and Mohammad Hossein Saraee. "Application of data mining in traffic management: case of city of Isfahan." *Electronic Computer Technology (ICECT), 2010 International Conference on. IEEE, 2010.*
- [11] Thakare, Vishakha S., et al. "Design of smart traffic light controller using embedded system." *ISOR-JE* 10.1 (2013): 30-3.
- [12] Kotsiantis, Sotiris B., I. Zaharakis, and P. Pintelas. "Supervised machine learning: A review of classification techniques." (2007): 3-24