A Fuzzy C-Means Image Segmentation Approach for Axle-based Vehicle Classification

A Revised Paper Submitted to TRB for Review for Presentation & Publication At the 95th TRB Annual Meeting, January 10-14, 2016, Washington, DC

Zhuo Yao, Ph.D. Candidate
ART-Engines Transportation Research Laboratory
College of Engineering and Applied Science, 735 ERC,
University of Cincinnati, Cincinnati, OH 45221-0071
Tel: (513) 556-5420; Fax: (513) 556-2599; E-mail: yaozo@mail.uc.edu

Heng Wei*, Ph.D., P.E.
Professor, Director, ART-Engines Transportation Research Laboratory
College of Engineering and Applied Science, 792 Rhodes Hall,
University of Cincinnati, Cincinnati, OH 45221-0071
Tel: (513) 556-3781; Fax: (513) 556-2599; E-mail: heng.wei@uc.edu

Zhixia (Richard) Li, Ph.D.
Assistant Professor in Transportation Engineering
Department of Civil and Environmental Engineering, University of Louisville
W.S. Speed Bldg., Room 111, Louisville, KY 40292
Tel: (502)-852-2307; E-mail: richard.li@louisville.edu

Jonathan Corey, Ph.D.
Assistant Professor, College of Engineering and Applied Science, 796 Rhodes Hall
University of Cincinnati, Cincinnati, OH 45221-0071
Tel: (513) 556-6554; Fax: (513) 556-2599; E-mail: coreyjn@ucmail.uc.edu

Submitted on: November 15, 2015

Word Count: 5,451 + 8(figures) * 250 = 7,451

*Corresponding Author
ABSTRACT

Vehicle classification information is vital to almost all types of transportation engineering and management applications such as pavement design, signal timing and many other safety applications. Though vehicular length-based classification scheme is widely used by state Departments of Transportation, it lacks capability of accurately produce axle-based classification data. Limited by the capital cost, axle-based vehicle classification data sources are very limited. This paper presents an image segmentation-base vehicle classification system with an attempt to increase the efficiency of the axle-based vehicle classification. The video-base vehicle classification system: Rapid Video-based Vehicle Identification System (RVIS) tool is developed to automatically identify number of axles extracted from ground-truth videos. Through testing of individual vehicle images, it is shown that RVIS is indeed capable of accurately detect all FHWA 13 class vehicles. Large scale testing the RVIS running with a predetermined set of morphological parameters produce less accurate results. However, the comparison of two testing hours shows that through more efforts in calibration, the results can be improved. The advantages of the RVIS are its robust and fast algorithm, flexibility to be applied either at a mobile video source or locations with the traffic monitoring videos are available. And therefore, expand the locations where vehicle classification data can be available at network scale level.

KEYWORDS: Axle-based Vehicle Classification, Video-based Vehicle Identification System, Fuzzy-C Means Classification,
INTRODUCTION

Any state using traffic data for the allocation of Federal funds are required to maintain a traffic monitoring system that meets the Federal Highway Administration (FHWA) requirements. Many transportation agencies (i.e. State DOTs, MPOs, etc.) have recognized that traffic data supports a growing variety of functions and critical decision making processes. The need for data and the benefits that result from the required data must be balanced against available and potential resources to implement an effective and efficient traffic monitoring program. As part of a traffic monitoring program, state Departments of Transportation (DOTs) are required to collect vehicle count, classification, weight, and speed data. Since participation in federally funded programs is essential to the integrity of a State’s highway systems, the accurate, efficient collection of traffic data becomes a critical component of transportation infrastructure management [1]. The traffic programs serve a variety of traffic engineering purposes include planning, design, calibration, collection, distribution, analysis, reporting, and maintenance [2]. The 2013 edition of the Traffic Monitoring Guide (TMG) defaults the application of axle-based FHWA 13 class scheme in the vehicle classification data format requirements. Common sources of axle-based vehicle classification data sources are intrusive in nature. They include pneumatic rubber tubes, piezoelectric sensors, active infrared sensors, etc. Pneumatic tube traffic counters usually provide the total number of axles at the end of a count, but an adjustment factor is usually required to convert the total number of axles into total number of vehicles [3]. The adjustment factor itself is usually an additional source of inaccuracy aside from its under-counting problem. A survey of state Departments of Transportation [4] shows that Pneumatic rubber tube has been reported on issues in data accuracy, weather interference and limitations in lanes monitored. The piezoelectric sensor is capable of counting axles by sensing the passage of the vehicle’s individual axles but may not be able to tell how many axles belong to one single vehicle. In practice, lots of the automatic traffic recording (ATR) stations use a P-L-P (piezo-loop-piezo) or L-P-L configuration in the freeway for better performance. Comparative study [5] results show that active infrared device such as The Infra-Red Traffic Logger (TIRTL) provides adequate accuracy for the FHWA axle-based classifications comparing to other commercial products such as radar detectors. However, active infrared devices are usually expensive to own, install and maintain.

In recent years, video and image processing techniques have shown to be cost effective in various traffic data collection and traffic control applications [6, 7]. In spite of limits of its dependence on lighting conditions, video-based systems have some advantages over intrusive axle-based vehicle methods. For example, low impact on the road infrastructure, low maintenance costs, robustness of feature detection from images and readily available from regional intelligent transportation systems (ITS) centers. It is of the state DOTs’ interest to periodically check the performance of available automatic traffic recording (ATR) stations. However, it would very labor intensive and almost impossible to manually collect data for quality control at every desirable location. An efficient and fast method and associate tool for processing video to produce accurate traffic information is therefore needed to improve the quality and increase the sources for vehicle classification data acquisition. This paper hence presents an image segmentation approach to improve the accuracy of the axle-based vehicle classification.

Yao et al [8] proposed a prototype of the Rapid Video-based Vehicle Identification System (RVIS) with system function designs and preliminary results. This paper extends previous work
on generating traffic classification data in three ways: (a) to improve the previously developed computer vision-based algorithm to extract axle-based vehicle classification, and (b) to introduce the theoretical foundations of the proposed methodology, and (c) by building the capability of automatically extract traffic classification data from a larger video dataset. The remainder of the paper is organized as follows. The state-of-the-art computer vision applications in traffic data extraction study is summarized as a result of literature review, followed by a description of the methodology and data source used in this study. Then, case studies covering the FHWA scheme F class vehicles results are presented. Finally, a summary, conclusions, and recommendations for further research are presented.

SUMMARY OF EXISTING STUDIES

Much research has been conducted research using data extracted from video and image-based tools [9, 10, 11] at the University of Cincinnati. Zhang et al. [12] used virtual dual loop detectors setup within the videos screen to mimic the functionality of dual-loop detectors. It filled the gap of traffic data extraction where dual-loop detectors are not available. However, the method is still model-based and adapted the modeling errors pertaining in the system. Kanhere [13] attempted to develop video based vehicle classification system using pattern recognition. The method considered using wheels as a feature, but did not provide reference to classification using the axle and its configuration. Hsieh et al. [14] used size and linearity features to dynamically classify vehicles from a built library. This machine learning-based method is very library templates dependent and usually requires a considerable amount of time and efforts for the calibration process. Besides, the classification only has four bins which representing car, minivan, truck and van truck which may not satisfy the needs of axle-based classification scheme. Ma and Grimson [15] attempted to classify vehicles from edge points of a detected vehicle object from video. They used classification techniques to extract the vehicle shape. This effort is also library based and requires a long time for learning and recognizing vehicles. Recently, Yao et al. [16] developed a computer vision-based software tool, namely, Rapid traffic Emission and energy Consumption Analysis (REMCAN) system. It enables a rapid vehicle operating mode distribution profiling for MOtor Vehicle Emission Simulator (MOVES) model from video data. Clearly, the availability of ground truth vehicle classification data will help to maximize the MOVES model capacity.

Less research efforts have been reported regarding generation of axle-based vehicle classification data by using recent development in computer vision and image processing techniques. Frenzel [17] proposed a video-based system in a heuristic study focused on truck detection and axle counting rather than vehicle classification. Their results shows only 56% axle detection rate which is very low. Nevertheless, no vehicle classification study was carried out within this study. A couple of studies [18, 19] applied 3D model-based computer vision system for vehicle classification. The vehicles are modeled at 3-dimentional level and then classified into classes. This method usually requires extensive expertise in computer vision techniques and computer programming yet produced misclassifications since one 3D model may have multiple axle configurations according to the FHWA scheme F.

METHODOLOGY
The goal of this research is to explore an axle-based vehicle classification method using existing image processing techniques to fulfill the identified research gap. To achieve the goal, two objectives are designated to fulfill: (1) to design a vehicle classification system, Rapid Video-based Vehicle Identification System (RVIS), based on axle numbers of a vehicle; and (2) to test and validate the proposed axle parameter based vehicle classification system using a FHWA scheme F classes. The proposed research addresses the challenges and identified research gap through the development and testing of the proposed RVIS with a case study. The advantage of the proposed RVIS system is that it is a ground truth video data-based, non-intrusive classification method. The ground-truth based method is reliable since it bypasses the modeling and malfunctioning errors which conventional sensors might have. The video source is from home entertainment grade camcorder that produces 1024 x 768 videos at 30 frames per second.

Figure 1 shows the system flow of the RVIS. The RVIS has four modules, namely, video to image, vehicle axle extraction, axle-based vehicle classification and the calibration and validation module. The video to image module enhances, splits, and resizes raw video data into a series of images of individual vehicles with common standard size. The vehicle axle extraction module is the core module in this system. It detects, segments the vehicle axle pixels using the fuzzy C-Mean clustering algorithm. The vehicle classification module, which contains predefined FHWA vehicle classification axle configurations, matches and classifies the outputs from the axle parameter extraction module. The classification is based on number of axles. Finally, a calibration and validation module is designed to guarantee the performance of the proposed RVIS system. This module uses the ground truth video to check and correct any possible misclassifications and errors.

**Background Segmentation and Foreground Extraction**

A robust background subtraction algorithm should be able to handle lighting changes and long-term scene changes. To achieve this, the method of Mean filter is used. Let V(x,y,t) represent a video sequence where t is the time of the video frame, x and y are the 2-dimension pixel location variables. To estimate the background image, a series of sequential frames are averaged in a sliding time window. The background image at the time t is calculated as

\[ B(x, y) = \frac{1}{N} \sum_{i=1}^{N} V(x, y, t-i*m) \]  

(1)

where N is the number of sequential images taken for averaging, m is the number of consecutive frames between two neighboring sequential images. N and m should be picked carefully to capture the instant change of the environment like the light change. After estimating the background B(x,y), the foreground at time t can be obtained by subtracting B(x,y) from V(x,y,t) and threshold it. The foreground is

\[ |V(x, y, t) - B(x, y)| = Th \]  

(2)

where Th is threshold. The selection of Th would affect the accuracy of foreground extraction. To remove noise in the foreground image, a set of morphological operations including opening, closing and dilation are used in this project. Figure 2 further illustrates the background segmentation and foreground extraction process.

**Fuzzy C-Mean Clustering-based Image Segmentation**
Image segmentation is the process of segmenting a group of image pixels into a set of disjoint regions that have similar characteristics such as intensity, color and texture. A considerable number of image segmentation techniques are available. The clustering-based methods, which segment the feature space of image into several clusters and derive a sketch of the original image, include K-means [20], Fuzzy C-means (FCM) [21, 22] and mean-shift [23] algorithms. The FCM algorithm is widely applied to image segmentation because it has robust characteristics for ambiguity and ability to retain much more information comparing to threshold-based segmentation methods. FCM clustering is an unsupervised clustering technique applied to segment images into clusters with similar spectral properties. It utilizes the distance between pixels and cluster centers in the spectral domain to compute the membership function. The pixels on an image are highly correlated, and this spatial information is an important characteristic that can be used to develop the clustering. This technique was originally introduced by Jim Bezdek in 1981 [24] as an improvement on earlier clustering methods. It provides a method that shows how to group data points that populate some multidimensional space into a specific number of different clusters.

The FCM algorithm is an iterative optimization that minimizes the cost function $J_{FCM}$ defined as follows:

$$J_{FCM} = \sum_{k=1}^{K} \sum_{i=1}^{M \times N} u_{ik}^m d_{ik}^2 = \sum_{k=1}^{K} \sum_{i=1}^{M \times N} u_{ik}^m \| x_i - \mu_k \|^2$$

where $m$ is any real number greater than 1, $M \times N$ is the number of pixels in image, $u_{ik}$ is the degree of membership of $x_i$ in the cluster $k$, $x_i$ is the $i^{th}$ element of $d$-dimensional measured data, $\mu_k$ is the center of the cluster with $d$-dimension (for images $d = 2$), $d_{ik}^2$ is a distance measure between object $x_i$ and cluster center $\mu_k$, and $\| \ast \|$ is any norm expressing the similarity between any measured data and the center $\mu_k$. The cost function is achieved when pixels with short distance to the centroid of their clusters are assigned with high membership values, while pixels far from the centroid are assigned to low membership values. The membership function represents the probability (fuzziness) that a pixel belongs to a specific cluster to some degree. The probability is dependent solely on the distance between the pixel and each individual cluster center in the feature domain.

The Fuzzy partitioning is carried out through an iterative optimization of the objective function shown previously, with the update of membership $u_{ik}$ and the cluster centers $\mu_k$ by

$$u_{ik} = \frac{1}{\sum_{i=1}^{K} \left( \frac{d_{ik}^2}{d_{ij}^2} \right)^{\frac{1}{m-1}}} \left( \frac{1}{\sum_{i=1}^{MN} u_{ik}} \right)$$

$$\mu_k = \frac{\sum_{i=1}^{MN} u_{ik} x_i}{\sum_{i=1}^{MN} u_{ik}}$$
This iteration stops when \( \max_i \left\{ \| u_{ik}^{(n+1)} - u_{ik}^{(n)} \| \right\} < \varepsilon \), where \( \varepsilon \) is a termination criterion between 0 and 1, \( d_j \) is the \( j \)th of \( d \)-dimensional measured data, whereas \( n \) is the number of iteration. This procedure converges to a local minimum or a saddle point of error.

The FCM algorithm is directly applied on the pixels of an image. The degree of membership of pixels in each class is therefore calculated. Starting with an initial guess at each cluster center, the FCM converges on a solution for \( \mu_k \) representing the local minimum or a saddle point of the cost function. Convergence can be detected by comparing the changes in the membership function or the cluster center of two successive iteration steps.

The algorithm is composed of the following steps [25]:

1. Set values for \( M, N, m \) and \( \varepsilon \).
2. Initialize \( U = [u_{ik}] \) matrix, \( U^{(0)} \)
3. Set the loop counter \( b = 0 \).
4. At \( n \)-step: calculate the centers vectors \( \mu^{(n)} = [\mu_k] \) with \( U^{(n)} \)
   \[ \mu_k = \frac{\sum_{i=1}^{MN} u_{ik}^n x_i}{\sum_{i=1}^{MN} u_{ik}^n} \]  
   \[ u_{ik} = \frac{1}{\sum_{i=1}^{k} \left( \frac{d_{ik}}{d_{ij}} \right)^{m-1}} \]
5. Update \( U(n) \), \( U(n+1) \)
6. If \( ||U(n+1) - U(n)|| < \varepsilon \) then STOP; otherwise set \( b = b + 1 \) and return to step 4.

In each iteration of the algorithm, \( M \times N \times k \) probability functions are calculated (\( k \) is the number of classes). Figure 3 shows an illustration on the FCM clustering results. In this case, the pixels composing the vehicle tires are highlighted in the crosshair (Figure 3-a) and clustered into a group as shown in the RGB color space (Figure 3-b). For the below specific image, the tire pixels have an average RGB value of 58, 58, 58 (grayish tire colors tend to have similar RBG values) and share 4.23% of among the total number of pixels within the image.

**Axle Detection**

The axle parameter extraction module fine tunes the axle segmentation and is further illustrated using an example in figure 4. This module takes in an input image and start with a Fuzzy C-means clustering to segment out the tires from the image (figure 4.2). Then, the color information of the image can be discarded and it is converted into grayscale image (figure 4.3). Figure 4.4 shows the result of morphological opening (filtering out the small objects in image) on the grayscale or binary image with the structuring element. The morphological open operations, erosion followed by dilation, use the same structuring element for both operations. After this, a hole filling process is
performed where a hole is defined as a set of background pixels that cannot be reached by filling in the background from the edge of the image (figure 4.5). The difference between figure 4.5 and 4.4 (mathematical operation), therefore, representing the wheel entity is shown in figure 4.6. The final step (figure 4.7) simply extracts the segmented wheels’ parameter. The further axle-based classification will be performed based on results from this step.

To further extract vehicle tires and reduce noisy, two morphological operators are used: erosion and dilation. With applying these two operators if a foreground object has some holes in it this technique will fill the holes. If some small foregrounds detected which were not connected to a big foreground object it will be eliminated.

Erosion on a binary image can be expressed as:

$$A \ominus B = \bigcap_{b \in B} A_b$$

where A is a binary image. Note that in a binary image, 1 represent as white and 0 represents, and B is a $3 \times 3$ matrix with a center anchor and value of one for each element and gets convolved with the entire image. Figure 5 shows the result of this operation on a binary image. It can be seen that this operator was able to eliminate the noisy and unwanted objects.

Dilation can be expressed as:

$$A \oplus B = \bigcup_{b \in B} A_b$$

where A is a binary image. Note that in a binary image 1 represent as white and 0 represents and B is a $3 \times 3$ matrix with a center anchor and value of one for each element and gets convolved with the entire image. Figure 6 shows the result of this operation on a binary image. It can be seen that this operator was able to connect the relevant parts.

This technique is employed after the Fuzzy C-means image pixel classification. It has significant impact on the accuracy because it identifies the axles.

RESULTS

The case study of RVIS used video data collected on October 19th, 2013 at I-275 mile marker 45-46, Cincinnati, Ohio. The case study results focused on covering the full spectrum of FHWA 13 class rather than single or several classes. A selection of FHWA 13 class vehicle images from the RVIS video to image module is used to test the FCM clustering algorithm. Figure 4 shows the axle detection and parameter extraction results from the FHWA 13 vehicle classes. The first image line shows the original images as input for the RVIS. The second image line shows the result using fuzzy c-mean segmented axles. Third image line shows the hole filling result from the rim of the axle. The binary images (fourth image line) shows the difference of the hole filled image and the segmented image. The last image line shows the axle blobs where the count of axles.

Note for class 1 (i.e. motorcycle) and class 4 (i.e. bus) vehicles, RVIS used methods were not axle-based. For class 1- motorcycles, all the rear axle of the motorcycles are covered and it is impossible for the program to segment out. Therefore, in the below case study, class 1 vehicles are only detected as vehicle with one-axle. For class 4 vehicles, our sample is an image of a yellow school bus (there are no other types of bus running on the stretch of freeway in our data collection).
It is a more simply way to just detect the color yellow by its RGB values at 255, 255, 0 range. This is a faster and more efficient way to determine the vehicle class if it is a yellow school bus.

Figure 7 shows the result of RVIS detecting the axle-based FHWA 13 classes with success. The large scale, automatic running of two-hours of video data is presented below. After processing the video data, the results were compared to the ground truth images extracted from the video dataset by RVIS.

Table 1 summarizes the RVIS results after conducting a two-hour testing. Due to the fact that vehicle with higher axles are less on the road, the sample size distribution is extremely loaded to two-axle vehicles. RVIS was able to sample out 1,397 vehicles from the video data and generated 842 correctly identified for the first hour. The detection rate is 60.27%. For the second hour, the sample size is 1300 with 803 correctly identified. The detection rate is 65.31%. Averaging the two hours RVIS detection, the overall detection rate is 62.79%. The morphological operators (erosion and dilation) for hour 1 and hour 2 are calibrated to improve the detection rate. The table 1 results show that the RVIS accuracy can be improved when the morphological operators are better calibrated to adapt the changes to the lighting environment.

Although the relative detection rate is approximately 63%, RVIS still have some advantages comparing to other automated vehicle classification methods. First of all, RVIS is cheap to deploy, run and generate results since it bypasses the intrusive nature of traditional automated vehicle classification methods such as loop detectors. RVIS could be easily deployed at any desired locations that is in need of vehicle classification data just using a home camcorder. Since it is so easy to deploy video cameras and record the video data, RVIS can potentially provide a lot bigger spatial territory without investing the cost of automatic vehicle classification devices. In addition, RVIS could provide data at locations such as truck terminals, bus terminals, air ports and logistic centers to generate axle-based vehicle classification data for the purposes of freight modeling, travel demand model calibration and validation, and safety studies etc. This along saves lots of capital, time and avoids the safety risks of installing the intrusive devices. Nevertheless, RVIS could generate data when traditional data collection method fails in conditions such as congestion. Under conditions such as congestion, since the traffic is moving very slow or not moving at all, it brings error to the data collection methods such as loop detectors. RVIS, on the other hand, could capture the true vehicle axles since it is ground truth-based rather than modeled algorithm-based.

CONCLUSION

This research set out an alternative vehicle classification data source enabling low-cost performance check over the tradition data sources such as ATR stations. RVIS helps transportation agencies’ traffic data programs to achieve accurate vehicle classification data which is critical to transportation engineering, safety and management applications. The proposed vehicle classification approach made its contribution to currently vehicle classification data collection of the following: (1) provides a low-cost way of collecting vehicle classification to as performance checking over existing data sources such as loop detectors, Radar, etc.; (2) provides an alternative way of vehicle classification data source when traditional classification method fails; (3) complimenting the existing vehicle classification data sources spatially and temporally. The
preliminary results showed that RVIS is capable of providing vehicle classification information and is potentially capable of providing supportive grounds for decision makers to confront the challenge of invest wisely in rehabilitation, maintenance, materials and process. The low-cost vehicle classification method will expand the vehicle classification data coverage and performance and return cheaper and more location available results.

The case study testing the FHWA 13 classes of vehicles shows that the RVIS can be an additional vehicle classification data source on top of existing vehicle classification data collection methods. Although the results are preliminary, RVIS is indeed capable of accurately detect all FHWA 13 class vehicles from the testing images as shown in figure 7. Large scale testing the RVIS running with a set of morphological parameters produce less accurate results. However, the second hour produce higher accuracy through a better set of calibrated morphological parameters. The comparison of two testing hours shows that through more efforts in calibration, the results can be improved. The advantages of the RVIS are its robust and fast algorithm, flexibility to be applied either at a mobile video source or locations with the traffic monitoring videos are available. And therefore, expand the locations where vehicle classification data can be available at network scale level. Future work includes continuously improving the algorithm in the detection to adapt the changes in lighting environment, develop a calibration protocol to produce more accurate results, testing with larger datasets, and the RVIS performance evaluation against other automated vehicle classification methods. Issues related with low-lighting or lighting environment changes should also be explored using alternative video sources such as thermal videos. Since all video-based data collection systems are what you see is what you got (WYSIWYG) type of application, another potential improvement is solve the vehicle occlusion issues. In addition, a thorough analysis of economic benefit and cost of the proposed RVIS comparing to other types of automated vehicle classification methods would be beneficial.
REFERENCES


1. **Motorcycles** (2 Axles, 2 or 3 tires)
2. **Passenger Cars** (2 Axles, Can have 1 or 2 axle trailers)
3. **Pickups, Panes, Van** (2 Axles, 4 tire single units can have 1 or 2 axle trailers)
4. **Buses** (2 or 3 axles, full length)
5. **Single Unit 2 Axle Trucks** (2 axles, 6 tires, single unit)
6. **Single Unit 3 Axle Trucks** (3 axles, single unit)
7. **Single Unit 4 or More Axle Trucks** (4 or more axles, Single Unit)
8. **Single Trailer 3 or 4 Axle Trucks** (3 or 4 axle trucks)
9. **Single Trailer 5 Axle Trucks** (5 axles, single trailer)
10. **Single Trailer 6 or More Axle Trucks** (6 or more axles, single trailer)
11. **Multi-trailer 5 or Less Axle Trucks** (5 or less axles, Multiple Trailers)
12. **Multi-trailer 6 Axle Trucks** (6 axles, multiple trailers)
13. **Multi-trailer 7 or More Axle Trucks** (7 or more axles, multiple trailers)

**FIGURE 1** System flowchart for the Rapid Video-based Vehicle Identification System (RVIS).
Image 1: Background Image.

Image 2: New Video Frame.

Image 3: Grayscale New Video Frame.
Figure 2 Sample background segmentation and foreground extraction.
Figure 3. Fuzzy C-Mean clustering on axle pixels-RGB (58, 58, 58), percentage: 4.23%.
(3-a: Original Image with Marked Tire Pixels; 3-b: Classified Tire Pixel Cluster based on RGB Value)
Image 1: Input Image

Image 2: Fuzzy-C-Means Clustering

Image 3: Convert Image to Grey scale

Image 4: Filtering Out Small Blobs (binary large object)
Image 5: Filling Out Closed Areas in Image

Image 6: Subtract Image 5 and 4, then Convert to Binary Image

Image 7: Feature Detection and Extraction

FIGURE 4 The axle detection and extraction process.
FIGURE 5 Result of erosion operator on a binary image.
FIGURE 6 Result of dilation on a binary image.
<table>
<thead>
<tr>
<th>Original</th>
<th>Class 1</th>
<th>Class 2</th>
<th>Class 3</th>
<th>Class 4</th>
<th>Class 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Axle Segmentation</td>
<td><img src="image" alt="Axle Segmentation" /></td>
<td><img src="image" alt="Axle Segmentation" /></td>
<td><img src="image" alt="Axle Segmentation" /></td>
<td><img src="image" alt="Axle Segmentation" /></td>
<td><img src="image" alt="Axle Segmentation" /></td>
</tr>
<tr>
<td>Hole Filling</td>
<td><img src="image" alt="Hole Filling" /></td>
<td><img src="image" alt="Hole Filling" /></td>
<td><img src="image" alt="Hole Filling" /></td>
<td><img src="image" alt="Hole Filling" /></td>
<td><img src="image" alt="Hole Filling" /></td>
</tr>
<tr>
<td>Binary Image</td>
<td><img src="image" alt="Binary Image" /></td>
<td><img src="image" alt="Binary Image" /></td>
<td><img src="image" alt="Binary Image" /></td>
<td><img src="image" alt="Binary Image" /></td>
<td><img src="image" alt="Binary Image" /></td>
</tr>
<tr>
<td>Axle Detection</td>
<td><img src="image" alt="Axle Detection" /></td>
<td><img src="image" alt="Axle Detection" /></td>
<td><img src="image" alt="Axle Detection" /></td>
<td><img src="image" alt="Axle Detection" /></td>
<td><img src="image" alt="Axle Detection" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Original</th>
<th>Class 6</th>
<th>Class 7</th>
<th>Class 8</th>
<th>Class 9</th>
<th>Class 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Axle Segmentation</td>
<td><img src="image" alt="Axle Segmentation" /></td>
<td><img src="image" alt="Axle Segmentation" /></td>
<td><img src="image" alt="Axle Segmentation" /></td>
<td><img src="image" alt="Axle Segmentation" /></td>
<td><img src="image" alt="Axle Segmentation" /></td>
</tr>
<tr>
<td>Hole Filling</td>
<td><img src="image" alt="Hole Filling" /></td>
<td><img src="image" alt="Hole Filling" /></td>
<td><img src="image" alt="Hole Filling" /></td>
<td><img src="image" alt="Hole Filling" /></td>
<td><img src="image" alt="Hole Filling" /></td>
</tr>
<tr>
<td>Binary Image</td>
<td>Axle Detection</td>
<td>Class 11</td>
<td>Class 12</td>
<td>Class 13</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>----------------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td></td>
</tr>
<tr>
<td>Original</td>
<td>Axle Segmentation</td>
<td>Hole Filling</td>
<td>Binary Image</td>
<td>Axle Detection</td>
<td></td>
</tr>
</tbody>
</table>

FIGURE 7 Axle detection results for FHWA 13 classes.
## TABLE 1 RVIS Detection Rate Summary

<table>
<thead>
<tr>
<th>Vehicle Axle Type</th>
<th>Ground Truth</th>
<th>RVIS Result</th>
<th>Detection Rate</th>
<th>Ground Truth</th>
<th>RVIS Result</th>
<th>Detection Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hour 1</td>
<td></td>
<td></td>
<td>Hour 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-axle</td>
<td>1,322</td>
<td>813</td>
<td>61.50%</td>
<td>1,228</td>
<td>759</td>
<td>61.81%</td>
</tr>
<tr>
<td>3-axle</td>
<td>9</td>
<td>1</td>
<td>11.11%</td>
<td>12</td>
<td>8</td>
<td>66.67%</td>
</tr>
<tr>
<td>4-axle</td>
<td>8</td>
<td>2</td>
<td>25.00%</td>
<td>7</td>
<td>1</td>
<td>57.14%</td>
</tr>
<tr>
<td>5-axle</td>
<td>52</td>
<td>24</td>
<td>46.15%</td>
<td>51</td>
<td>14</td>
<td>54.90%</td>
</tr>
<tr>
<td>6-axle</td>
<td>5</td>
<td>1</td>
<td>20.00%</td>
<td>2</td>
<td>1</td>
<td>50.00%</td>
</tr>
<tr>
<td>7-axle</td>
<td>1</td>
<td>1</td>
<td>100.00%</td>
<td>3</td>
<td>1</td>
<td>66.67%</td>
</tr>
<tr>
<td>School bus</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>100.00%</td>
</tr>
<tr>
<td><strong>Detection Rate</strong></td>
<td><strong>1,397</strong></td>
<td><strong>842</strong></td>
<td><strong>60.27%</strong></td>
<td><strong>1,300</strong></td>
<td><strong>803</strong></td>
<td><strong>65.31%</strong></td>
</tr>
<tr>
<td><strong>Overall Detection Rate</strong></td>
<td></td>
<td></td>
<td><strong>62.79%</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>