

---

This is an electronic reprint of the original article.  
This reprint may differ from the original in pagination and typographic detail.

Kleyko, Denis; Hostettler, Roland; Lyamin, Nikita; Birk, Wolfgang; Wiklund, Urban; Osipov, Evgeny

**Vehicle classification using road side sensors and feature-free data smashing approach**

*Published in:*  
2016 IEEE 19th International Conference on Intelligent Transportation Systems, ITSC 2016

*DOI:*  
[10.1109/ITSC.2016.7795877](https://doi.org/10.1109/ITSC.2016.7795877)

Published: 22/12/2016

*Document Version*  
Peer-reviewed accepted author manuscript, also known as Final accepted manuscript or Post-print

*Please cite the original version:*  
Kleyko, D., Hostettler, R., Lyamin, N., Birk, W., Wiklund, U., & Osipov, E. (2016). Vehicle classification using road side sensors and feature-free data smashing approach. In *2016 IEEE 19th International Conference on Intelligent Transportation Systems, ITSC 2016* (pp. 1988-1993). Article 7795877 (Proceedings of the IEEE International Conference on Intelligent Transportation Systems). IEEE.  
<https://doi.org/10.1109/ITSC.2016.7795877>

---

This material is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

# Vehicle Classification using Road Side Sensors and Feature-free Data Smashing Approach

Denis Kleyko, Roland Hostettler, Nikita Lyamin, Wolfgang Birk,  
Urban Wiklund, and Evgeny Osipov

This is a post-print of a paper published in *IEEE International Conference on Intelligent Transportation Systems*. When citing this work, you must always cite the original article:

D. Kleyko, R. Hostettler, N. Lyamin, W. Birk, U. Wiklund, and E. Osipov, "Vehicle classification using road side sensors and feature-free data smashing approach," in *IEEE International Conference on Intelligent Transportation Systems*, Rio de Janeiro, Brazil, November 2016

## DOI

10.1109/ITSC.2016.7795877

## Copyright

© 2017 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

# Vehicle Classification using Road Side Sensors and Feature-free Data Smashing Approach

Denis Kleyko<sup>1</sup>, Roland Hostettler<sup>2</sup>, Nikita Lyamin<sup>3</sup>, Wolfgang Birk<sup>1</sup>, Urban Wiklund<sup>4</sup>, and Evgeny Osipov<sup>1</sup>

**Abstract**—The main contribution of this paper is a study of the applicability of data smashing – a recently proposed data mining method – for vehicle classification according to the “Nordic system for intelligent classification of vehicles” standard, using measurements of road surface vibrations and magnetic field disturbances caused by passing vehicles. The main advantage of the studied classification approach is that it, in contrast to the most of traditional machine learning algorithms, does not require the extraction of features from raw signals. The proposed classification approach was evaluated on a large dataset consisting of signals from 3074 vehicles. Hence, a good estimate of the actual classification rate was obtained. The performance was compared to the previously reported results on the same problem for logistic regression. Our results show the potential trade-off between classification accuracy and classification method’s development efforts could be achieved.

## I. INTRODUCTION

Vehicle classification is an important task in traffic monitoring and analysis. Rich information about the traffic composition provided by a classification analysis is commonly used for different purposes such as urban planning, road maintenance, traffic light scheduling, etc. For a long time, this kind of information has been obtained based on inductive loop detectors for permanent installations or pressure tubes for temporary installations [1], [2]. During the last decade, with the advance of cheap sensor technology, wireless communication, and electronics, sensor networks have started to replace those traditional systems [3], [4]. Some advantages of these novel approaches include the possibility of on-demand or real-time access to the data, and slower wear-off rate due to the possibility of non-invasive installations.

The Nordic Research and Development cooperation (NordFoU) [5] has started to formulate the “Nordic system for intelligent classification of vehicles” (NorSIKT) standard [6], a new vehicle classification standard. The purpose of the standard is to “establish a Nordic standard for classification of vehicles and thereby to be able to exchange and compare traffic data between the different countries” [5]. The standard is based on four different classification levels as illustrated

in Table I. Here, a higher classification level, represents a more fine-grained classification of the vehicles. For example, level one represents the coarsest classification that essentially corresponds to a detection of a vehicle only. Level four is on the other end of the scale, i.e. the most detailed level with a total of 14 classes for light and heavy vehicles including those with and without trailers and distinguishing between, for example, motorcycles and mopeds. The development of the standard also requires adaption of the existing as well as the development of new classification methods to fit both the newly available sensor hardware and classification standards. This newly developed standard will help national transportation authorities to easily compare different traffic counting equipments under consideration. Furthermore, it also provides a mean for certification of equipment in the future.

Traffic monitoring using wireless sensor nodes as a such is a rather mature field. In [4], it was shown that magnetometers can be used to count traffic, to estimate vehicles’ speed, and even to classify vehicles. Furthermore, in [7] it was shown how to use a road-surface mounted micro accelerometer and a neural networks-based algorithm to distinguish between diesel, gasoline, and heavy diesel engine vehicles using the frequency spectrum generated by a passing vehicle. The works in [8], [9] also used accelerometer-based vehicle detection. The authors developed a peak detection algorithm to detect individual vehicle axles followed by a table lookup. Authors in [10] proposed the bio-inspired classifier using vibration measurements from the road side sensor. The main disadvantage of the method is that it is not able to generalize to the previously unseen patterns of vibration signal changes. Finally, a setup using vibration measurements from under the roadway combined with a neural network classifier for perimeter surveillance was introduced in [11].

The main contribution of this study is the use of the recently proposed data mining method called data smashing [12] to the vehicle classification problem. The advantage of this classification approach is that it, in contrast to the most of traditional machine learning algorithms, does not require the extraction of features from raw signals. Extraction of relevant features from raw data is an engineering art on its own and usually requires a domain expert who understands the problem at hand. Moreover, the choice of features does significantly affect the classification performance. In contrast, when applying data smashing the requirements for the expert knowledge are minimized which positively affects development cost and time. Also, it was already shown that for several problems in different domains data smashing

<sup>1</sup>Denis Kleyko, Wolfgang Birk, and Evgeny Osipov are with Department of Computer Science Electrical and Space Engineering, Luleå University of Technology, Luleå, Sweden {denis.kleyko; wolfgang.birk; evgeny.osipov}@ltu.se

<sup>2</sup>Roland Hostettler is with Department of Electrical Engineering and Automation, Aalto University, Helsinki, Finland roland.hostettler@aalto.fi

<sup>3</sup>Nikita Lyamin is with School of Information Technology, Halmstad University, Halmstad, Sweden nikita.lyamin@hh.se

<sup>4</sup>Urban Wiklund is with Department of Biomedical Engineering and Informatics, Umeå University, Umeå, Sweden urban.wiklund@umu.se

TABLE I  
THE NORSIKT CLASSIFICATION SCHEME AND CLASSIFICATION OF THE MEASURED VEHICLES.

Level 1	Level 2	Level 3	Level 4
Motor Vehicle (MV) 3074	Light Motor Vehicles + Motorcycle (LMV) 2845	Motorcycles (LMV1): 53	Motorcycle (MC): 50
		Light Motor Vehicles (LMV2): 2792	Moped (MP): 3
			Passenger car without coupled vehicle (PC WOC): 2466
			Passenger car with coupled vehicle (PC WC)
			Light goods road motor vehicle without coupled vehicle: 326
			Light goods road motor vehicle with coupled vehicle
	Heavy Motor Vehicles (HMY) 229	Heavy Motor Vehicles (HMY1): 229	Light bus (LB)
			Other light road motor vehicle (LV)
			Heavy goods road motor vehicle without coupled vehicle: 86
			Heavy goods road motor vehicle with coupled vehicle: 118
			Heavy bus (HB): 27
			Road tractor without coupled vehicle (RT WOC): 2
			Road tractor with coupled vehicle (RT WC)
			Other heavy road motor vehicle (HV)

performs on a par with the traditional machine learning methods; while the traditional data mining methods were applied to these problems using the extensive knowledge of human experts e.g., for the extraction of relevant features from raw data [13].

The aims of this paper are threefold:

- to study the applicability of the data smashing approach for vehicle classification according to the NorSIKT classification standard.
- to evaluate the data smashing method on a large dataset, consisting of signals from 3074 vehicles, to get an understanding of the performance characteristics of the data smashing approach in the vehicle classification.
- to compare the results with the performance of traditional machine learning algorithms for the same dataset reported in [14].

The remainder of this article is structured as follows. Section II presents the data smashing process used for classification. A description of the measurement setup used for data collection, the preprocessing of signals, the classification pipeline, and the main contribution of the paper – the performance evaluation are described in Section III. We conclude the article in Section IV.

## II. DATA SMASHING

Data smashing is a data mining approach, which is suitable for comparison of two arbitrary data streams with each other. It was recently proposed by Chattopadhyay and Lipson in [12], [13]. The main goal of the data smashing process is to determine whether two compared data streams were produced by the same source (i.e. generating stochastic process) or not. Quantitatively it is done through the measurement of the casual similarity (distance) between data streams, which is represented by a real number between 0 (more similar) and 1 (less similar). Thus, data smashing is an algorithm consisting of several procedures, which outputs a single number characterizing how likely two input streams were produced by the same source. For example, in the considered application there are three possible generating sources of data streams: motorcycles, light motor vehicles, and heavy motor vehicles (see Level 3 in Table I).

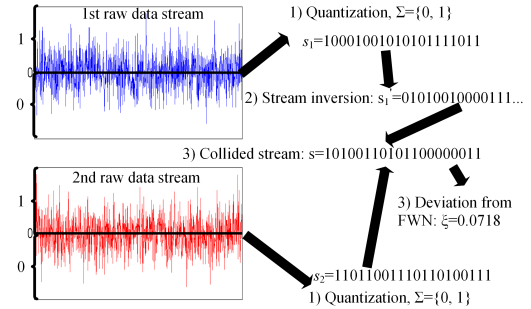


Fig. 1. An example of the data smashing process. First, two raw data input streams are quantized into binary streams. Next, one of the quantized streams is inverted using the corresponding algorithmic step. Finally, the inverted stream and another input stream are collided. The collided stream is then used to calculate the deviation from flat white noise (FWN), which characterizes similarity of two input streams.

The method is theoretically based on the assumption that each source is a stochastic process that can be described by a probabilistic automaton with a finite number of states. Due to the space limitations, this section gives only a high-level overview and the algorithmic steps involved in data smashing. Readers interested in the theoretical foundation and proofs are referred to [12] for the detailed description. Note that the method does not attempt to explicitly reconstruct the probabilistic automata from the data streams. Instead, it estimates the similarity of two streams purely through operations on the streams' values.

The whole process of data smashing consists of three steps and is exemplified in Fig. 1. The method works with data streams with a finite alphabet  $\Sigma$  where  $|\Sigma|$  denotes the size of the alphabet. The first step (unless streams are already symbolic) is a quantization of both input data streams. In other words, raw data, e.g. a stream of sensor measurements should be first quantized, thus, converting the original stream into the stream of symbols. There are no strict requirements to the usage of the particular quantization scheme. However, the general recommendation is that all symbols should have a relatively high or even similar frequency of appearance in a quantized data stream. There is also no restriction on the number of symbols in the alphabet, but it was theoretically

shown that for larger alphabet the reliable estimation of the similarity requires longer data streams. The simplest quantization scheme for real values is a binary alphabet, e.g. with symbols denoted as 0 and 1. This alphabet requires a setting of a single threshold. If the current value is above the threshold it is quantized to 1 otherwise to 0. Such quantization can be easily generalized to the case when the alphabet has  $|\Sigma|$  symbols; in this case,  $|\Sigma| - 1$  thresholds should be defined. The raw value is then always in a slice between two thresholds, and it is assigned a symbol corresponding to this slice. Note that concrete notations of symbols in the alphabet are not important (e.g. “0” or “a” or “#”) as long as they are consistent across all compared quantized data streams. The quantization schemes used in this paper for magnetometer and accelerometer streams are discussed in the next section.

In the second step, one of the quantized streams is chosen, and it is used to generate its anti-stream. The process is called stream inversion. Stream inversion in turn requires generation of independent stream copies from the quantized stream,  $s$ . The algorithmic procedure for the generation of the independent stream copy uses streams generated by flat white noise (FWN). According to [12] for the alphabet  $\Sigma$ , FWN generates the current symbol of a stream from the uniform distribution, i.e. the probability of appearance of each alphabet’s symbol is  $1/|\Sigma|$ . The pseudo code for the generation of an independent stream copy is as follows:

- 1) generate stream  $\omega_0$  from FWN
- 2) read current symbol  $\sigma_1$  from  $s$ , and  $\sigma_2$  from  $\omega_0$
- 3) if  $\sigma_1 = \sigma_2$ , then write  $\sigma_1$  to output stream  $s'$
- 4) read next symbol and go to step 1)
- 5) when done return stream copy  $s'$

The pseudo code for the stream inversion of a quantized input stream copy is as follows:

- 1) generate  $|\Sigma| - 1$  independent copies of  $s$ :  $s_1, \dots, s_{|\Sigma|-1}$
- 2) read current symbols  $\sigma_i$  from  $s_i$  ( $i = 1, \dots, |\Sigma| - 1$ )
- 3) if  $\sigma_i \neq \sigma_j$  for all distinct  $i, j$ , then write  $\Sigma \setminus \bigcup_{i=1}^{|\Sigma|-1} \sigma_i$  to output stream  $s'$
- 4) read next symbol and go to step 1)
- 5) when done return inverted stream  $s'$

In the final step the anti-stream from the second step is first summed (collided) with the second quantized input stream. The process is called information annihilation. Finally, we estimate information remaining in the collided stream via calculation of its deviation from FWN. The result of the calculation is a real number in the range  $[0, 1]$  which characterizes the similarity of two input data streams.

The summation of two streams  $s_1$  and  $s_2$  is done according to the following steps:

- 1) read current symbols  $\sigma_i$  from  $s_i$  ( $i = 1, 2$ )
- 2) if  $\sigma_1 = \sigma_2$ , then write  $\sigma_1$  to output stream  $s'$
- 3) read next symbol and go to step 1)
- 4) when done return collided stream  $s'$

The deviation of the collided stream from FWN is esti-

mated by

$$\hat{\xi}(s, l) = \frac{|\Sigma| - 1}{|\Sigma|} \sum_{x: |x| \leq l} \frac{\|\phi^s(x) - U_\Sigma\|_\infty}{|\Sigma|^{2|x|}}, \quad (1)$$

where  $|x|$  is the length of string  $x$ ,  $l$  is the maximum length of strings up to which the sum is evaluated. Thus, strings  $x$  are all possible combinations of up to  $l$  symbols in  $\Sigma$ . For a given threshold  $\epsilon$ ,  $l$  is chosen as  $l = \ln(1/\epsilon) / \ln(|\Sigma|)$ .  $U_\Sigma$  is uniform probability vector of length  $|\Sigma|$ . Finally, for  $\sigma_i \in \Sigma$ ,  $\phi^s(x)|_i = \frac{\text{number of occurrences of } x \cup \sigma_i \text{ in string } s}{\text{number of occurrences of } x \text{ in string } s}$ .

### Illustrative applications of data smashing

Authors in [12], [13] presented results of using data smashing for several machine learning problems from varied domains where the raw data can be represented as a stream. In particular, the method was used to solve two typical problems: clustering (unsupervised learning) and classification (supervised learning). Classification problems included identification of epileptic pathology using electroencephalographic (EEG) data series, biometric authentication using visually evoked EEG potentials, recognition of variable stars from light intensity series, and text independent speaker identification using speech recordings of individuals. Clustering was applied for detection of areas pertaining to heart murmur from noisy heart-sound recordings. For all the above mentioned problems data smashing showed high performance. The current work is to the best of our knowledge an unique attempt of applying the data smashing technique for feature-less vehicle classification.

## III. CLASSIFICATION WITH DATA SMASHING

### A. Experimental setup

The results reported in this paper are based on the measurements conducted in Amsberg, Sweden in August 2013 [15]. More detailed description of the setup can be found in [14]. The setup consisted of a pair of sensor nodes as in [16]. Each node was equipped with an accelerometer measuring the road surface vibrations and a magnetometer measuring the magnetic field disturbances. Measured signals were preprocessed (filtered and down-sampled) such that only signals’ magnitudes (envelopes) are stored. The resulting signals are both sampled at 256 Hz. In total, 3399 vehicles were registered. Out of these, 175 were not correctly detected. Additionally 150 vehicles were not used during the evaluation. Thus, 3074 passages with the corresponding ground truth are available in the dataset. Table I illustrates the distribution of the ground truth according to the NorSIKT standard.

Unfortunately, the dataset is heavily biased toward light motor vehicles in general and passenger cars in particular. This makes it difficult to evaluate classifiers that target level 4 classification with high confidence, especially since some classes are heavily underrepresented. Thus, in this study we focused on level three classification only. Examples of measured signals for classes in level three are presented in Fig. 2.

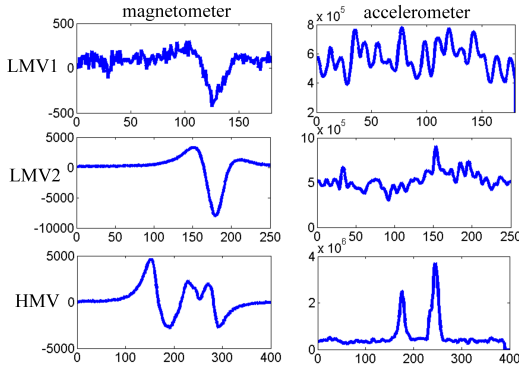


Fig. 2. Examples of magnetometer and accelerometer signals for each class at level three in the NorSIKT standard.

### B. Preprocessing and quantization of raw signals

The sensors used for the evaluation are an accelerometer and a magnetometer, both mounted on a single sensor node on the road side. The data sets were aggregated during evaluation trials, which were conducted by the Swedish Traffic Administration.

Once vehicle passage is detected by the roadside sensor node, there are two raw signals available: the magnetometer signal ( $y_m[n]$ ) and the accelerometer signal ( $y_a[n]$ ), see Fig. 2 for examples of signals for each vehicle class. First, both signals from each passage are narrowed to a sample window around the signal center where the signal's magnitude exceeds three standard deviations of the noise and a possible noise-offset is subtracted (solid box in the left part of Fig. 3). For reliable estimation of similarity between two data streams data smashing requires both streams to be relatively long. Therefore, when the sample window is extracted from the original signal, it is used to create an extended signal by cyclically repeating the sample window until the length of the signal achieves the predetermined length. The right part of Fig. 3 illustrates the extension of the accelerometer signal from HMV using the sample window extracted from the original vehicle passage (left part of Fig. 3). Our preliminary experiments showed that 5000 samples are enough for the robust estimation of similarity for both accelerometer and magnetometer signals. This length was used to get the classification results reported in this paper.

The extended signals are used as an input to data smashing. Recall, that these signals have continuous values, and, therefore, they first should be quantized to the finite alphabet. We have used the maximum-entropy quantization scheme from [12] for mapping of continuous measurements from magnetometer and accelerometer to the alphabet with four symbols. The number of symbols in the quantization scheme was chosen based on our preliminary tests. Such quantization requires the setting of three threshold values. Three thresholds define four areas of signal values. Each area corresponds to a symbol in the alphabet. Thresholds are allocated in such a way that each area includes approximately the same number (i.e. 25% for the chosen scheme) of continuous values of the given training signals. This is due to the

recommendation that each symbol should have a relatively high frequency of appearance in a symbolic data stream. Two sets of thresholds (one for each type of measurements – accelerometer and magnetometer) were estimated using a sample of training data. Estimated quantization schemes along with examples of continuous signals are presented in Fig. 4. Note, that once defined, the quantization schemes should be used to quantize all continuous value signals.

### C. Test and Training Datasets

Recall from Table I that the dataset consists of 53 passages in the class LMV1, 2792 passages of LMV2 and 229 passages of HMV1. Approximately 60% (32 passages) of the available LMV1 passages were used to form the training dataset. The rest (21 passages) was used for the test dataset to validate the trained models. In order to keep the training dataset more balanced, the presence of LMV2 and HMV classes was restricted to 69 and 84 passages respectively. Thus the size of the testing dataset is bigger than the training dataset (185 passages) and consists of 2889 passages being highly biased toward LMV2 class. To minimize the influence of the passages chosen for the training dataset and get the averaged performance of the classifier, the training and testing datasets were randomly generated 7 times from the initial dataset.

### D. Classification process

First, the magnetometer and the accelerometer signals of all 185 training passages are preprocessed as described above to get the extended continuous value signals. Next, these signals are used to estimate quantizations schemes for magnetometer and accelerometer measurements. Once schemes are estimated all training signals are quantized into symbolic streams with four possible states.

Both signals of a testing passage are first preprocessed and quantized according to the existing schemes. Next, data smashing is applied between the testing symbolic streams and the symbolic streams of each training passage. When comparing two vehicle passages data smashing is used four times: two times for magnetometer streams and two times for accelerometer streams. During the first usage the training stream is inverted while during the second usage the testing stream is inverted. This is necessary because the resultant similarity estimations are not necessarily the same (due to stochastic nature of the inversion process). Thus, the comparison of two vehicle passages is characterized by four real numbers, each in range  $[0,1]$ . Their average is taken to have a single number characterizing the similarity between two passages. Once all data smashing calculations for a single training passage are completed, there is a vector of 185 real numbers, which characterizes the similarity between the current testing passage and each passage in the training dataset. Because the values of the resultant vector act as distances between passages, the natural way to estimate the most probable class of the testing passage is via application of k-nearest neighbor (kNN) method; where k is a parameter representing the number of the closest neighbors used to

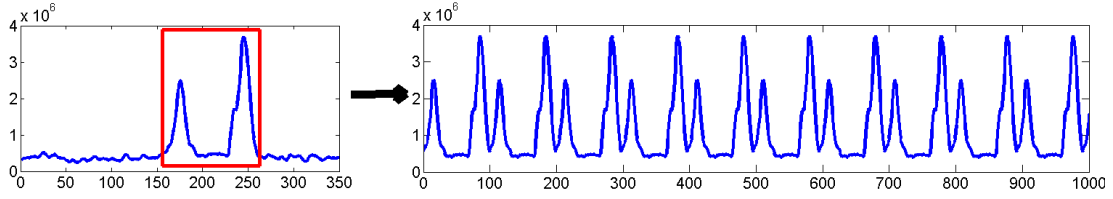


Fig. 3. An example of the extraction of a relevant part from an accelerometer signal for HMV and subsequent extension of the signal using this part.

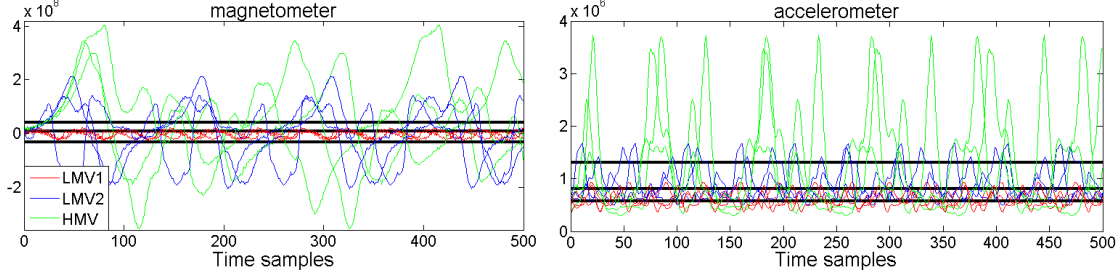


Fig. 4. Both magnetometer and accelerometer continuous signals are quantized into the alphabet with four symbols. Such quantization requires the setting of three threshold values. Left part presents used thresholds along with several examples of continuous signals from each class for magnetometer measurements. Right part illustrates the same information for accelerometer measurements.

predict the class label. The testing passage is assigned a class label corresponding to the highest number of the nearest neighbors. Finally, the performance of the classifier is assessed by comparing the predicted classes for testing passages with the ground truth.

#### E. Performance metrics

The main performance metric for the studied methods is the classification accuracy, which is a quantitative indication of how well a method predicts the correct class of a vehicle. It is defined as

$$\hat{r} = \frac{1}{N} \sum_{n=1}^N I_n(\hat{c}_n) \quad (2)$$

where  $N$  is the number of classified vehicles and  $I(\hat{c}_n)$  is the indicator function indicating whether the estimated class  $\hat{c}_n$  is the true class  $c_n^*$  or not. It is defined as

$$I(\hat{c}_n) = \begin{cases} 1 & \hat{c}_n = c_n^* \\ 0 & \hat{c}_n \neq c_n^* \end{cases} \quad (3)$$

#### F. Results

The performance of the classification with data smashing and kNN against the number of nearest neighbors is presented in Fig. 5. Plots in the upper part of Fig. 5 present the classification accuracy in each class: LMV1, LMV2, HMV. The lower part of the figure shows aggregated performance metrics, namely: weighted average recall (the ratio of correct predictions to the total number of predicted classes) and unweighted average recall, UAR (the average of class-specific prediction accuracies). As the dataset is heavily biased towards LMV2, the weighted average recall is nearly identical to the curve for LMV2. Therefore, the best number of nearest neighbors was identified as one providing the

TABLE II

CONTINGENCY TABLE FOR THE CLASSIFICATION BY DATA SMASHING.

		Ground truth		
		LMV1	LMV2	HMV
Forecast	LMV1	19.3	281.0	0.3
	LMV2	1.6	2079.1	15.1
	HMV	0.1	347.9	144.6

TABLE III

CONTINGENCY TABLE FOR THE CLASSIFICATION BY REGRESSION.

		Ground truth		
		LMV1	LMV2	HMV
Forecast	LMV1	20.0	23.3	0.3
	LMV2	0.9	2528.9	10.2
	HMV	0.1	155.8	149.5

highest unweighted average recall. Thus, the classification provides the best performance when  $k$  equals 6.

Table II presents the performance in the form of contingency table for this case. Values in the table were averaged across several runs using the training and testing datasets which were randomly generated from the initial one. The values in the main diagonal of the table indicate number of correct predictions for each class. The values outside main diagonal show misclassifications for the corresponding vehicle type.

In our previous study [14], traditional features-based machine learning methods (logistic regression, neural networks, and support vector machines) were applied to the same data. The best performance was demonstrated by the logistic regression (see Table III) with UAR 94.0%. Thus, the performance of the data smashing based classifier in terms of UAR (86.3%) is comparable with the benchmark from the feature-based methods. Both approaches are also comparable in terms of average accuracies for LMV1 (95.2%



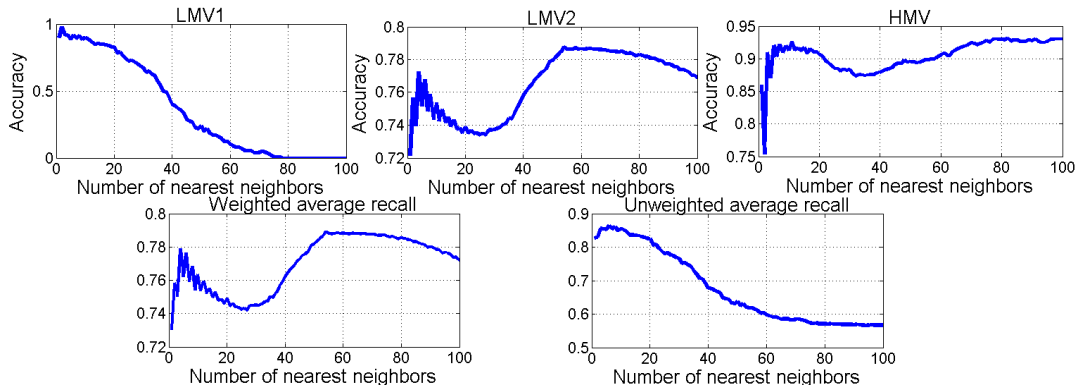


Fig. 5. The average performance of vehicle classification with data smashing and kNN against the number of nearest neighbors.

vs. 91.8%) and HMV (93.4% vs. 90.4%) classes. However, the data smashing based classification provides rather poor classification for LMV2 class (93.4% vs. 76.8%).

It is conjectured that the main reason for the low performance on LMV2, as can be seen from the table, is that LMV2 passages are often predicted as either LMV1 or HMV. The explanation appears to be the fact that LMV2 is a broad class, and it includes wide range of vehicles with different physical dimensions that can be quite close to the neighboring classes. At the same time the logistic regression was able to correctly predict larger number of LMV2 passages as its training process includes the optimization routine aimed at minimizing the number of incorrect predictions on the training dataset, while the approach studied here is based purely on the similarity calculation without any training routine.

#### IV. CONCLUSION

In this paper we applied the feature-free data smashing method to the problem of vehicle classification based on magnetometer and accelerometer measurements from road side sensors. The motivation for the data smashing usage comes from the fact that, in contrast to most of the traditional machine learning algorithms, it does not require the extraction of features from raw signals.

The method was evaluated on a large dataset, consisting of 3074 vehicle passages in total. As illustrated by experiments estimating the accuracy of the classification of vehicles passages, the main drawback of data smashing usage in this context is its moderate performance (76.8% vs. 93.4% for the state-of-the-art feature-based method) shown for the largest class.

In our future studies we aim at increasing the accuracy of the proposed classifications scheme to the level of state-of-the-art by tuning parameters of the classifier (e.g., quantization scheme) as well increasing available dataset by keeping feature-free nature of the approach. Another prospective direction is a design of an online learning architecture featuring zero configuration. Potentially this could be achieved by combining the proposed classification scheme with an additional detector (e.g. camera based) to enable an automatic collection of the initial knowledge.

Acknowledgements: This study was supported in part by the Swedish Research Council (grant no. 2015-04677).

#### REFERENCES

- [1] C. Sun, "An investigation in the use of inductive loop signatures for vehicle classification," California Path Program, Institute of Transportation Studies, University of California, Berkeley, Tech. Rep. UCB-ITS-PRR-2000-4, 2000.
- [2] C. Sun, S. G. Ritchie, and S. Oh, "Inductive classifying artificial network for vehicle type categorization," *Computer-Aided Civil and Infrastructure Engineering*, vol. 18, no. 3, pp. 161–172, May 2003.
- [3] E. Sifuentes, O. Casas, and R. Pallas-Areny, "Wireless magnetic sensor node for vehicle detection with optical wake-up," *Sensors Journal, IEEE*, vol. 11, no. 8, pp. 1669–1676, August 2011.
- [4] J. Ding, S.-Y. Cheung, C.-W. Tan, and P. Varaiya, "Signal processing of sensor node data for vehicle detection," October 2004, pp. 70–75.
- [5] T. Vestman, "NorSIKT: State of the art Norden," Nordic Research and Development Cooperation, Tech. Rep., 2012.
- [6] T. Vaa, P. Melen, D. Andersson, and B. Bach Nielsen, "Norsikt-nordic system for intelligent classification of traffic," *Procedia - Social and Behavioral Sciences*, vol. 48, pp. 1702–1712, 2012.
- [7] J. Lan, T. Lan, and S. Nahavandi, "A novel application of a microaccelerometer for target classification," *Sensors Journal, IEEE*, vol. 4, no. 4, pp. 519–524, August 2004.
- [8] R. Bajwa, R. Rajagopal, P. Varaiya, and R. Kavalier, "In-pavement wireless sensor network for vehicle classification," in *Information Processing in Sensor Networks (IPSN), 2011 10th International Conference on*, Chicago, IL, USA, April 2011, pp. 85–96.
- [9] R. Hostettler, W. Birk, and M. Lundberg Nordenvaard, "Feasibility of road vibrations-based vehicle property sensing," *IET Intelligent Transport Systems*, vol. 4, no. 4, pp. 356–364, 2010.
- [10] D. Kleyko and E. Osipov, "Brain-like classifier of temporal patterns," in *The Proceeding of the 2nd International Conference on Computer and Information Sciences - ICCOINS*, 2014, pp. 1–6.
- [11] M. Stocker, P. Silvonen, M. Rönkkö, and M. Kolehmainen, "Detection and classification of vehicles by measurement of road-pavement vibration and by means of supervised machine learning," *Journal of Intelligent Transportation Systems*, vol. 0, 2015.
- [12] I. Chattopadhyay and H. Lipson, "Data smashing: uncovering lurking order in data," *Journal of The Royal Society Interface*, vol. 11, no. 101, pp. 1–11, December 2014.
- [13] —, "Data smashing," in *Proceedings of the 28th AAAI Conference on Artificial Intelligence (AAAI-14)*, 2014, pp. 7–14.
- [14] D. Kleyko, R. Hostettler, W. Birk, and E. Osipov, "Comparison of machine learning techniques for vehicle classification using road side sensors," in *Proceedings of the 2015 IEEE 18th International Conference on Intelligent Transportation Systems*, 2015, pp. 572–577.
- [15] "Analysis of measured data for the calculation of accuracy in vehicle classification and various measurement data," [http://nordfou.org/documents/norsikt/Rapport\\_01v11\\_Report\\_Analysis\\_vehicle\\_data\\_Amsberg,%20August%202013.pdf](http://nordfou.org/documents/norsikt/Rapport_01v11_Report_Analysis_vehicle_data_Amsberg,%20August%202013.pdf).
- [16] W. Birk, E. Osipov, and J. Eliasson, "iRoad – cooperative road infrastructure systems for driver support," in *Intelligent Transportation Systems, 2009 16th ITS World Congress and Exhibition on*, 2009.