cogent
engineering

CrossMark

CIVIL & ENVIRONMENTAL ENGINEERING | RESEARCH ARTICLE

Multiple vehicle fusion for a robust road condition estimation based on vehicle sensors and data mining

Julia Hofmockel[1§], Johannes Masino[2*§], Jakob Thumm[2], Eric Sax[3] and Frank Gauterin[2]

Abstract: The road condition is an important factor for driving comfort and has impact on safety, economy and health. Delayed detection of defects lead to renewals which yields to complete roadblocks or traffic jams. Therefore, an early identification of road defects is desirable. Novel condition monitoring systems employ vehicles as sensor platforms and apply machine learning methods to predict the road condition based on the sensor data. The paper addresses the question how to combine the classification results from different vehicles to improve the final prediction. Different fusion strategies are investigated in various scenarios in a novel simulation. It is demonstrated that the performance of the classification can be improved compared to a majority vote or only considering one vehicle by taking the probability for the prediction of each vehicle into account. The probabilities follow a multinomial distribution and the precision matrix of the classifiers provide the best parameters. Overall, the results show that the application of the presented fusion strategies on road condition estimation greatly improve the performance and guarantee a robust detection of defects.

Subjects: Artificial Intelligence; Automotive Technology & Engineering; Intelligent & Automated Transport System Technology; Transportation Engineering

*Corresponding author: Johannes Masino, Institute of Vehicle System Technology, KIT, Karlsruhe, Germany
E-mail: johannes.masino@kit.edu

ABOUT THE AUTHORS

Julia Hofmockel received the MSc degrees in Operations Research from Maastricht University in 2014. She joined the Institute for Information Processing Technologies at Karlsruhe Institute of Technology and the Audi Electronics Venture GmbH as PhD student in 2015. Research interests are machine learning, data mining and operations research.

Johannes Masino received the BSc and MSc degrees in Industrial Engineering and Management from Karlsruhe Institute of Technology (KIT) in 2011 and 2014. He joined the Institute of Vehicle System Technology at KIT as a PhD student in 2014. His research interests are signal processing, pattern recognition and machine learning with applications on vehicle sensors.

PUBLIC INTEREST STATEMENT

The condition of the road infrastructure has influence on the safety, economy and health. Therefore, there is a big interest in monitoring the road condition to detect defects and repair them as soon as possible. Modern vehicles combined with machine learning methods can be used as sensor platforms to monitor the road condition automatically. This paper addresses the problem how to combine the prediction of the road condition from multiple vehicles to increase the accuracy of the prediction.

cogent ·· oa

1. Introduction

1.1. Motivation

The condition of the road infrastructure has influence on the safety, economy and health. Moreover, studies have shown that the road condition has an impact on driving comfort (Ihs, 2005). Many researchers have proven the relation of the road surface condition to tire road noise (Sandberg, 1987). Subsequently, tire road noise leads to general annoyance, sleep disturbances or speech interference (Öhrström, Skånberg, Svensson, & Gidlöf-Gunnarsson, 2006; Schwela, 2000). Furthermore, roads with a rough surface and a high level of tire road noise lead to an increase in costs for noise abatement measures, such as sound absorbing barriers. It also yields to a decrease in value of houses and land close to noisy road transport infrastructure (Taylor, Breston, & Hall, 1982; Theebe, 2004; Wilhelmsson, 2000). Rough surfaces and road damages have an impact on the rolling resistance and therefore increase carbon dioxide emission and the vehicle operational costs, such as the fuel consumption (Descornet, 1990). Furthermore, if the road condition exceeds a critical value due to delayed detection of road defects, the erosion of the road substance accelerates. Subsequently, the road needs a complete renewal and the costs increase over the life cycle of the road (Forschungsgesellschaft für Strassen- und Verkehrswesen, 2004). A renewal often yields to a complete roadblock and to traffic jams. Overall, the sustainment of the road infrastructure is important for the economy, quality of life and traffic noise. The road network in many countries is large and the monitoring of the road infrastructure is mainly performed manually and is therefore time-consuming and expensive and scheduled in fixed intervals of 1 to 4 years (Koch & Brilakis, 2011).

1.2. Literature review

Researchers have addressed the problem of a manual road infrastructure monitoring and present concepts and methods for an automated data acquisition and analysis to estimate the road roughness or to detect road hazards.

For example, Eriksson et al. (2008) introduces a system to autonomously detect potholes. They equipped seven taxicabs in Boston with accelerometers and GPS. The authors applied several filters on the data to recognize potholes. The accuracy of the training data was 92.4% and 39 out of 48 potholes from the testing data were correctly classified.

Based on this approach, Chen, Lu, Tan, and Wu (2013) proposes a method to both detect potholes and rate the quality of the road surface. For this purpose, the authors collected acceleration and GPS data from 100 taxicabs. They distinguished between anomalies and normal road segments using a threshold on the vertical acceleration of the vehicle body. The value of the threshold was calculated separately for each vehicle with a Gaussian mixture distribution. The data with the identified anomalies were sent to a server to apply the same filters as described in (2008). In addition, the overall road roughness was determined with the standard deviation of the vertical acceleration. The accuracy to detect potholes was 90%.

Seraj et al. (2014) collected acceleration, gyroscope and GPS data in two cities from five different vehicles. The authors distinguished between three classes, namely severe (e.g. potholes, heavily patched road sections), mild (e.g. cracks) and span (e.g. speed bumps, pedestrian crossings) events. They used a machine learning approach and extracted features from the data in time and frequency domain with windows of 2.5 s with an overlapping factor of 0.66. The authors applied a two-stage support vector machine (SVM) algorithm. In the first step, they separated the windows with events and other windows and determined the class of event in the seconds step. The accuracy of the training with a ten-folded cross-validation and of the testing was 91 or 86%, respectively.

Overall, previous literature has proven the feasibility of an automated and inexpensive road condition monitoring with standard sensors. Either, the studies used similar vehicles with the same suspension system or adjusted the thresholds for anomaly detection. The combination of the classification results of each vehicle can be seen as a multiple expert problem.

The combination of multiple experts has been discussed in previous studies (e.g. Kittler, Hojjatoleslami, & Windeatt, 1997; Raykar et al., 2009). An interesting application of the multiple expert problem is the medical diagnosis from multiple medical experts. For example Mavandadi et al. (2012) improves the accuracy of diagnosis of malaria by combining decisions of multiple experts. Chu and Hwang (2008) cooperates multiple experts for Severe Acute Respiratory Syndrome diagnosis. Furthermore, the combination of multiple experts for medical diagnosis was tested by Hwang (1992) and Valizadegan, Nguyen, and Hauskrecht (2012).

The subjects robust estimation or robust observers gain more and more attention by researchers in the field of automotive engineering and signal processing. For example Dahmani, Chadli, Rabhi, and El Hajjaji (2013) use a robust observer to estimate the vehicle dynamic states in the presence of road bank angle. A robust fault tolerant tracking controller design for vehicle dynamics is presented in Aouaouda, Chadli, Boukhnifer, and Karimi (2015). Hu, Jing, Wang, Yan, and Chadli (2016) introduce a robust H_∞ output-feedback control strategy for path following of autonomous ground vehicles without the information of the lateral velocity. Furthermore, Bououden, Chadli, and Karimi (2016) proposes the design of a robust predictive control design for non-linear active suspension systems.

Overall, no approach has been presented to combine and merge the classification of multiple vehicles to get a robust and more precise estimation for the road condition estimation. We close this research gap and investigate different strategies to combine the prediction from different vehicles. The benefit of multiple vehicle fusion is shown with the following example. One vehicle might drive through a pothole at one specific position, whereas an other vehicle might slightly touch the pothole and subsequently predicts only a minor damage of the road. The classifier of one vehicle might tend to predict a wrong ground truth. Through fusing the output with other vehicles, the chance to classify the ground through correctly might increase. Further examples and scenarios are discussed in this paper.

1.3. Our contribution

Previous studies have shown, that it is possible to learn a classifier to predict the current state of the road with vehicles, using sensor data (Section 1.2). Connected vehicles can transmit the classification result over the air to a global data base. Here, the information is collected to create an overall representation of the road network.

The disadvantage of a machine learning approach is the uncertainty of the prediction since the classifier is usually not perfectly trained. It means, trusting the prediction of only one vehicle could lead to a wrong representation of the current road infrastructure condition. In this paper, we present a novel solution to merge the output of multiple vehicles to provide a robust estimation of road condition. The paper addresses the question of how often and by how many vehicles a road section needs to be passed, to be confident about the combined prediction. Section 2 describes and Figure 1 shows the required steps to classify the road condition: after training a classifier with sufficient accuracy, the model is applied in vehicles to categorize the road while passing. The results are transmitted and stored in a global back-end. It is a data collector where the data from multiple vehicles are saved and are available for real-time evaluations.

The contribution of the paper begins with the fusion of the data on the back-end and the analysis of the certainty about road condition estimations while applying a multiple expert contribution. The less vehicles are necessary to be confident about the predictions, the faster decisions can be made. An example is the warning of drivers before they pass a section with potholes. In the future, also autonomous vehicles benefit from the knowledge about the current road condition to reduce risks. Especially in real-time applications fast decisions are desirable but only reliable ones are valuable. Therefore, the paper has the aim to provide accurate estimates of the road condition while requiring as few observations as possible.

For the investigation of our proposed combination strategies, we developed a simulation in Matlab to analyze the amount of vehicles or overruns, which is needed to provide a prediction with a sufficient

cogent ·· engineering

Figure 1. Overview of the data processing in the vehicle to predict the road condition.
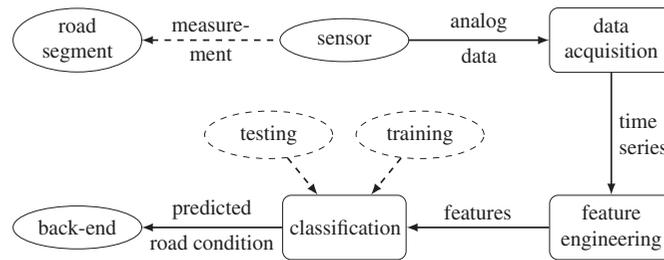
high accuracy. Furthermore, we identify the combination strategy with the most precise estimate. We also simulate different scenarios, for example, in which one sensor of a vehicle breaks down and the classifier outputs a lot of misclassifications or, in which the ground truth changes from one point of time to an other, e.g. when the road is maintained or the road condition gets worse after winter.

Hereby, our developed methods can be adapted to any road condition monitoring problem based on vehicle sensors and machine learning approaches, as long as a precision matrix can be derived from the output.

In Section 3, we describe our experiment to estimate the road condition with an inertial sensor in the vehicle body by applying a support vector machine. From the output of the classifier from the testing phase, we derive the precision matrix, which is the basis of our proposed methods.

2. Methodology

Modern vehicles are equipped with multiple sensors to measure their operation condition as well as the surrounding of the vehicles, for example the road condition. The relevant sensors to measure the road condition are inertial sensors, which determine the motion and vibration of the vehicle, and cameras, which scan the road. Compared to physical modelling, a condition estimation based on machine learning has the advantage that it does not require extensive system characterization.

Therefore, a classifier can be trained for each vehicle to predict the road condition based on vehicle sensor data. Our data processing method in the vehicle is shown in Figure 1. The measurement data are acquired and processed in the vehicle. From the time series data, features are calculated and selected to reduce the complexity and dimension of the data-set. After a classifier is trained and tested, it can be applied on new data. The predicted road condition is then sent to a back-end via V2X-communication interface. Section 3 describes our data acquisition and classification in more detail.

When using an inertial sensor in the vehicle for road condition estimation, the following points must be incorporated.

- Due to the different vibration behaviour of various vehicles, every vehicle might need an individual trained classifier.
- The performance and accuracy might vary among these individual classifiers.

When using a camera in the vehicle, the following aspect has to be considered.

- The accuracy and performance of the classifier might vary with the weather condition and the lighting condition.

Our proposal is to provide a more precise and robust prediction of the road condition by combining the prediction of various vehicles. These vehicles have individual classifiers with different characteristics. Each vehicle predicts the road condition online with the method shown in Figure 1. Afterwards, the prediction Class K_s for vehicle k is transmitted to a back-end, where the predictions of multiple vehicles is combined for a specific road segment, represented by its GPS position (longitude, latitude), as summarized in Figure 2. For the combination strategy, a believe factor $\pi_s^{(k)}$ for the prediction

Figure 2. Our approach to combine the predictions of the condition of road segments from multiple vehicles. The final prediction has the highest confidence by incorporating the believe factor π of the predictions of each vehicle.
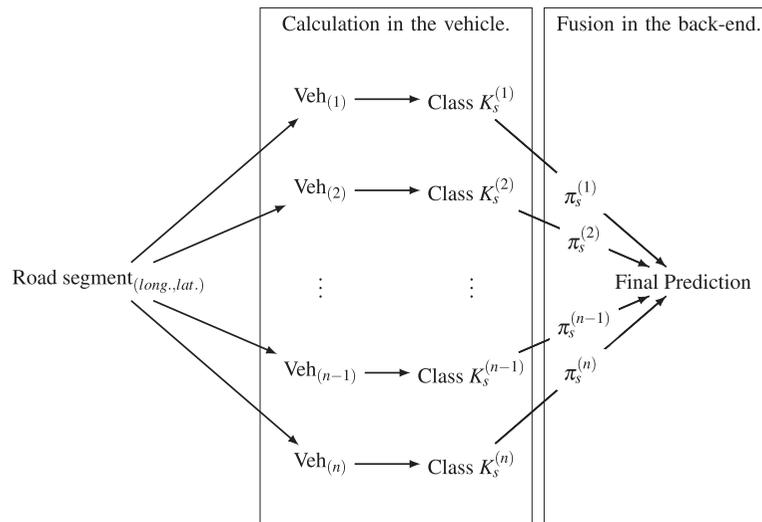
of vehicle k can be incorporated. The combination of the predictions can be realized with the following methods, which are explained in Section 2.2 in more detail.

- Majority vote,
- Vote with precision diagonal matrix,
- Vote with precision matrix.

The precision diagonal matrix and precision matrix for the latter two methods are derived in Section 2.1.

The focus of this study is to investigate the performance of the different methods to combine the predictions. Furthermore, different assumptions are formulated in scenarios and tested in a simulation to quantify their impact on the performance to predict the road condition. For example, in Scenario 1, the same precision matrix for each vehicle is assumed. Scenario 2 is more realistic, where each vehicle has an unknown precision matrix. The impact of a broken sensor for one vehicle is investigated with Scenario 3. Scenario 4 simulates the change of the road condition and introduces a method to improve the combination strategy.

2.1. Precision of a classifier for multiple vehicle fusion

The output of a classifier for each vehicle k from the training and testing process is a confusion matrix $M = (m_{ij}) \in \mathbb{N}^{l \times l}$ for classes K_i, $i = 1, \ldots, l$, as shown in Table 1. In the confusion matrix, m_{ii} presents the *true positives* for class i. The other elements in column j are called *false negatives*, in row i *false positives* and in the diagonal *true negatives*.

From the confusion matrix, one can calculate multiple performance measures, such as the recall with $\frac{m_{ii}}{\sum_{j=1}^{n} m_{ji}}$ for class K_i or the overall accuracy of the classifier with $\frac{\sum_{i=1}^{n} m_{ii}}{\sum_{i=1}^{n} \sum_{j=1}^{n} m_{ij}}$. An overview for performance measures for different calculation problems can be found in Sokolova and Lapalme (2009).

The precision presents the fraction of retrieved instances that are relevant and can be seen as the probability π_{ij} of the classifier to predict class i as class j for $i, j = 1, \ldots, l$.

We calculate the precision as

$$\text{Precision}_{ij} = \frac{m_{ij}}{\sum_{j=1}^{n} m_{ij}} = \pi_{ij}, \tag{1}$$

and can derive a precision matrix Π for each vehicle k as shown in Table 2.

cogent ·· engineering

Table 1. Confusion matrix of a classifier from the training or testing process

Target class	Output class					
	K_1	K_2	$\cdots$	K_{l-1}	K_l	Total
K_1	m_{11}	m_{12}	$\cdots$	$m_{1(l-1)}$	m_{1l}	$\sum_{j=1}^{l} m_{1j}$
K_2	m_{21}	m_{22}	$\cdots$	$m_{2(l-1)}$	m_{2l}	$\sum_{j=1}^{l} m_{2j}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$
K_{l-1}	$m_{(l-1)1}$	$m_{(l-1)2}$	$\cdots$	$m_{(l-1)(l-1)}$	$m_{(l-1)l}$	$\sum_{j=1}^{l} m_{(l-1)j}$
K_l	m_{l1}	m_{l2}	$\cdots$	$m_{l(l-1)}$	m_{ll}	$\sum_{j=1}^{l} m_{lj}$
Total	$\sum_{j=1}^{l} m_{j1}$	$\sum_{j=1}^{l} m_{j2}$	$\cdots$	$\sum_{j=1}^{l} m_{j(l-1)}$	$\sum_{j=1}^{l} m_{jl}$	$\sum_{i=1}^{n} \sum_{j=1}^{n} m_{ij}$

Table 2. Precision matrix derived from the confusion matrix

Target class	Output class					
	K_1	K_2	$\cdots$	K_{l-1}	K_l	Total
K_1	π_{11}	π_{12}	$\cdots$	$\pi_{1(l-1)}$	π_{1l}	1
K_2	π_{21}	π_{22}	$\cdots$	$\pi_{2(l-1)}$	π_{2l}	1
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$
K_{l-1}	$\pi_{(l-1)1}$	$\pi_{(l-1)2}$	$\cdots$	$\pi_{(l-1)(l-1)}$	$\pi_{(l-1)l}$	1
K_l	π_{l1}	π_{l2}	$\cdots$	$\pi_{l(l-1)}$	π_{ll}	1

The precision vector for Class K_s for vehicle k is then defined as

$$\vec{\pi}_s^{(k)} = (m_{s1}, \ldots, m_{sl})^{(k)}. \tag{2}$$

As soon as the performance of the classifier is below 100%, the task can be regarded as a multiple expert problem because each prediction has an uncertainty. The final decision can then be made by the fusion of the individual predictions (Dawid & Skene, 1979). The classifier predicts one out of l possible classes and its precision matrix Π expresses the accuracy. Hereby, entry π_{ij} describes the probability of classifying class i as class j for $i, j = 1, \ldots, l$.

The collected predictions for one road segment, named observation $\mathbf{o} \in R^l$, are distributed according to a multinomial distribution. The observation $\mathbf{o} \in R^l$ counts the number of predictions per class, with class i predicted o_i-times. Equation (3) describes the probability of observation $\mathbf{o}$, when the segment has class i as ground truth:

$$p_{o|i} = \frac{||o||_1!}{\prod_{j=1}^{l} o_j!} \prod_{j=1}^{l} \pi_{ij}^{o_j}. \tag{3}$$

2.2. Multiple expert decision
To make a decision about the property of one road segment, the different vehicle classifications are merged. Strategies under consideration are majority vote, vote with diagonal vector and vote with the precision matrix.

2.2.1. Majority vote
One possibility is to count how often each attribute is predicted. The one with most votes is assumed to be the ground truth:

$$\arg\max_{i \in 1, \ldots, l} o \tag{4}$$

✳ cogent ·· engineering

2.2.2. Vote with precision diagonal matrix

An extension to the majority vote approach is the vote with the diagonal vector from the precision matrix. Here the counts for the predictions are weighted with the precision vector:

$$w = (\pi_{11}, \pi_{22}, \ldots, \pi_{ll}). \tag{5}$$

That leads to an adapted observation o':

$$o' = ow^T. \tag{6}$$

As before, the final decision falls on the attribute with most votes:

$$\arg\max_{i\in 1,\ldots,l} o' \tag{7}$$

2.2.3. Vote with precision matrix

Each of the n vehicles might use an individually optimized classifier. This leads to precision matrix $\Pi^{(k)}$ and observation $o^{(k)}$ for vehicle k. Due to the assumption of independent predictions, the probability that class i is the ground truth, is calculated as shown in Equation (8) (Dawid & Skene, 1979).

$$p_{O|i} = \frac{\prod_{k=1}^{n} \prod_{j=1}^{l} (\pi_{ij}^{(k)})^{o_j^{(k)}}}{\sum_{q=1}^{l} \prod_{k=1}^{n} \prod_{j=1}^{l} (\pi_{qj}^{(k)})^{o_j^{(k)}}}, \tag{8}$$

with $O = \{o^k\}$ for $k = 1, \ldots, n$.

As final estimate, the class with highest probability is selected:

$$\arg\max_{i\in 1,\ldots,l} p_{O|i} \tag{9}$$

2.3. Scenarios

The presented methods to estimate the ground truth are evaluated by simulating different scenarios. We propose scenarios, which might occur in reality, e.g. that the road condition changes after maintenance or that a classifier of one vehicle is over-fitted and outputs random predictions. The point of interest is how many vehicles or overruns are necessary to reach a very small proportion of misclassifications for the prediction of the road segment and how to react to latter described scenarios.

2.3.1. Scenario 1: Classification with one global precision matrix

First of all the assumption of one global precision matrix Π_b for vehicles and for the back-end is made. Therefore, the classification of the vehicles and the evaluation on the back-end follow the same distribution. For simulation, a road with l segments, each segment having a different property, is assumed. The outputs of multiple vehicles lead to the evaluation of the different fusion strategies. Additionally, the prediction quality of the different classes is analyzed.

2.3.2. Scenario 2: Classification with unknown vehicle precision matrices

In reality, not all vehicles use the same distribution for prediction and the real precision matrix is unknown, since it might be different to the precision matrix calculated during training or testing. Therefore, in the second scenario each vehicle classifies the road segments by its own classifier while the final decision is made based on the global precision matrix Π_b. Hereby the vehicle precision matrix $\Pi_v^{(k)}$ is a slightly adapted version of the global one. In the simulation, vehicle k generates a random number to classify the road segments. For segment having property i, it follows the multinomial distribution described by the entries $\pi_{ij}(k)$, $j = 1, \ldots, l$ from the adapted precision matrix $\Pi_v^{(k)}$.

2.3.3. Scenario 3: Classification with equally distributed precision matrix

Another scenario under consideration is the classification of a vehicle totally independent of the ground truth. An example in real life is a vehicle with a classifier, which is over-fitted, a change of the vehicle vibration behaviour or a broken sensor. For this scenario, the effect on the fusion strategies, the possibilities to identify the vehicle making troubles and to react with adjustments of our fusion method are analyzed.

Table 3. Road features and corresponding class	
Class	**Road feature**
1	No visible defects
2	Minor defects (e.g. small cracks)
3	Major defects (e.g. potholes)
4	Off-road or cobblestone
5	Concrete joints in bad condition

2.3.4. Scenario 4: Change of the road condition

As last scenario the change of the ground truth from one time to another is simulated. It characterizes situations like a pothole that gets repaired, or a hard winter that destructs the road. It is verified if rules of forgetting or less weighting the past can improve the outcomes. The focus of this scenario is to analyze the impact of the change of ground truth over time on the results of our combination methods and not the impact of amounts of vehicles.

3. Experiment

We have performed experiments with vehicles to estimate the road condition and to determine a precision matrix for our simulation. We use an inertial sensor to measure the vehicle motion and vibration due to road damages. Modern vehicles have inertial sensors on-board.

The goal is to distinguish between classes listed in Table 3. The road classes were selected in consultation of local civil and transportation engineering departments to support the current process of the manual inspection of the road infrastructure. However, the classes can be customized and the proposed methods to combine the output of multiple vehicles are independent from the classes.

The data processing of the inertial sensor data to classify the road condition is applied according to Figure 1. The sensor data are annotated with the class of the road damage (Table 3), which is necessary for a supervised machine learning approach.

Firstly, time series data are transferred into a features space to reduce the complexity. We calculate representative and useful individual features from the complete or partial series of measurements, which is referred to feature extraction.

The features are calculated for sliding windows from time series as well as band-pass filtered signal of the vertical acceleration, pitch and roll rate. The features are

- standard deviation,
- energy density,

Figure 3. Visualization of the classified road condition in the urban area of Karlsruhe.
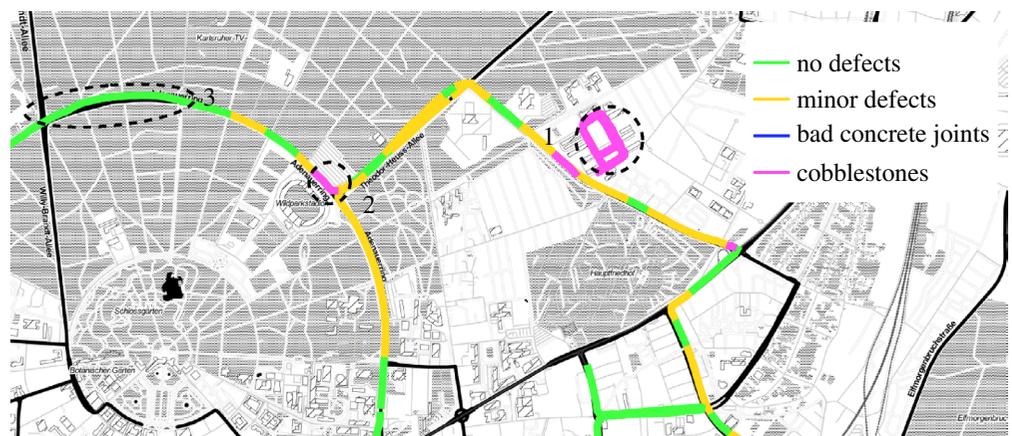
Figure 4. Visualization of the classified road condition of inter-urban roads.
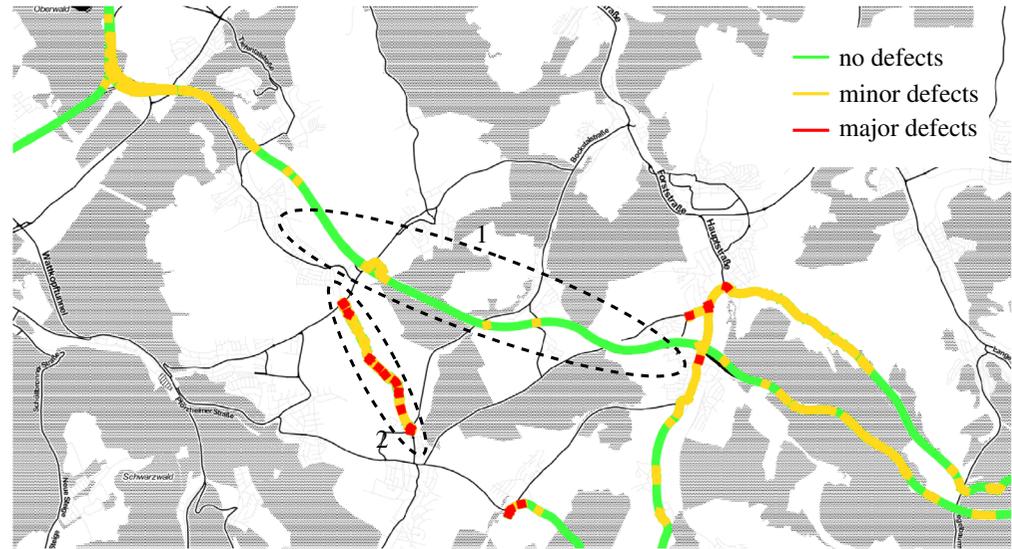
- maximum,
- minimum,
- peak-to-peak.

The calculated features are evaluated and the most suitable are selected with Multivariate Analysis of Variance Analysis (MANOVA), which is referred to as feature selection. From the selected features, which are annotated with the class of the road damage, a support vector machine with standard set-up is trained and tested. In the training process, the decision function of the support vector machine is determined and the performance is evaluated in the testing process. A portion of the collected data is used for training to determine the decision function to make predictions of sparse data.

The other portion of data is used for testing. Since these data are labelled as well, we examine the correctly and wrongly predicted classes. From the amount of correctly or wrongly predictions, we determine the confusion matrix and afterwards the precision matrix illustrated in Table 4 for one vehicle. The experiment is explained in more detail in Masino et al. (1979).

The trained classifier can be used to predict new data online in the vehicle. The feature calculation and classification can run on an electronic control unit of modern vehicles. The classification of such a road condition estimation method is exemplary visualized in Figure 3. The figure shows the results of data acquired in the urban area of Karlsruhe, Germany. Exemplary, there are two correctly predicted areas of *cobblestone* (Labels 1 and 2). The remaining road segments are correctly classified as segments with *minor defects* or *no defects*. Especially, latter class was correctly predicted for a road segment, which was recently renewed (Label 3). One miss-classification of *cobblestone* can be found in the south-east from Label 1. However, this road segment is highly damaged with multiple potholes, which have a high impact on the vehicle vibration similar to cobblestones or off-road.

Table 4. Global precision matrix for simulation						
Target class	**Output class**					
	K_1	K_2	K_3	K_4	K_5	**Total**
K_1	0.9	0.02	0.03	0.01	0.04	1
K_2	0.1	0.8	0.05	0.02	0.03	1
K_3	0.07	0.03	0.82	0.04	0.04	1
K_4	0.001	0.067	0.023	0.822	0.087	1
K_5	0.13	0.02	0.03	0.09	0.73	1

An other example for the classification of the road condition is shown in Figure 4. The inter-urban road marked with 1 is a freshly renovated asphalt with close to *no defects*. *On the other hand, the road marked with 2 is a poorly patched asphalt road with a lot of medium to severe defects, referred to as* minor *and* major *defects. The figure indicates, that our classifier predicts both road conditions correctly. Moreover, the condition of the other road segments of the map also accords with the ground truth.*

With the results of this experiment and the comparison of the predictions and the ground truth, a precision matrix for this classifier could be determined (Table 4). This precision matrix is bases for the investigation of the combination strategy of predictions from different vehicles in our simulation.

4. Simulation

We have developed a simulation environment in Matlab to calculate and compare the proportion of wrong predictions of different methodologies of the described scenarios. In our simulation, the road consists of s segments and each segment represents one of l classes as ground truth with $l = 5$ for all scenarios. In reality, theses classes could respond to types of damages, such as potholes. Therefore, the number of wrong predictions ranges from 0 to s for the created road. We set the values of the global precision matrix as shown in Table 4. The method to determine the precision matrix is explained in Section 3. Each vehicle drives over the road and classifies the segments with its individual precision matrix, which is unknown to the back-end. For classification in the simulation, a multinomial random number is generated. The values of the row in the precision matrix where the target class corresponds to the ground truth are the parameters.

Algorithm 1: estimate and adjust precision matrix in the back-end

Input : n – max. number of vehicles
s – max. number of road segments
$\Pi_v^{(k)}$ – precision matrix of vehicle $k \in 1,\dots,n$
Π_b – precision matrix stored in the back-end
l – number of classes
t – threshold for using $\Pi_e^{(k)}$ instead of Π_b

Output: number of misclassifications with adjusted precision matrix of vehicle k

1 **for** $v \leftarrow 1$ **to** n **do**
2   **for** $i \leftarrow 1$ **to** s **do**
3   **for** 1 **to** v **do**
4   vehicles drive over road segment (ground truth) with Π_v
5   **end**
6   combine predictions of all vehicles with Π_b
7   count number of misclassifications from combined predictions
8   **if** $i \mod 5,000 = 0$ **and** $v \geq 3$ **then**
9   combine predictions from vehicles $1\dots v$ with Π_b without vehicle k from last 5,000 road segments
10   estimate precision matrix $\Pi_e^{(k)}$ from the comparison of predictions of vehicle k with combined predictions
11           $d = \frac{\sum\sum |\Pi_e^{(k)} - \Pi_b|}{l^2}$
12   **if** $d \geq t$ **then**
13   use $\Pi_e^{(k)}$ for vehicle k instead of Π_b for further calculations
14   display: *'confusion matrix for vehicle k was adjusted, d, $\Pi_e^{(k)}$'*
15   **end**
16   **end**
17   **end**
18 **end**

For Scenario 1 to 3, we simulate 5 segments and perform 10,000 iterations of the drives to reduce stochastic effects. In Scenario 1 the precision matrix of the vehicles corresponds with the global precision matrix. We count the total number of wrong predictions and analyze the proportion of wrong classifications of all classes as well as the impact of each class for the number of vehicles.

The wrong predictions in scenarios 1 and 2 are calculated with the fusion based on majority vote, vote with precision diagonal matrix and vote with precision matrix.

In Scenario 2, the vehicle precision matrices are slightly randomly different to the global precision matrix and the fusion of the outputs is performed with the global one. The difference of the vehicle and global precision matrix could correspond to differences in the vehicle mass or various environment conditions and therefore a slightly different vibration behaviour of the vehicles. The proportions of wrong classifications are calculated for the number of drives as well as for the number of vehicles.

For the simulation of the further scenarios, we perform the output fusion only with the vote with precision matrix since the results of the previous scenarios have demonstrated the highest accuracy with this method.

In Scenario 3, we simulate one vehicle with a problem of the classifier and refer to it as vehicle *with broken sensor*. It classifies the segments totally random and the question is, how much impact it has on the final prediction. To decrease the negative impact of such a scenario, we apply the following method. After 5,000 overdriven segments and at least 3 vehicles, the simulation compares the output of one vehicle with the output of the other vehicles, which are fused with the system precision matrix as explained in earlier scenarios. Lines 8 to 16 in Algorithm 1 show the pseudocode of this method, which is embedded in the simulation algorithm. Based on these comparisons, we estimate the precision matrix of the one vehicle and check if the values differ greatly from the back-end precision matrix. If the difference exceeds a value, which is set to 0.15 in our simulation, we have identified the vehicle with the equally distributed precision matrix and wrong predictions. Furthermore, we use the estimated vehicle precision matrix for further calculations to reduce the belief of the output of this vehicle and therefore reduce the overall number of misclassifications. For the simulation of this method and for scenario 4, we set the number of segments to $s = 10,000$ and perform one iteration of the simulation per vehicle. The properties of the segments, representing one of the five classes, are nearly uniformly distributed.

In contrast to the previous scenarios, we simulate the time in Scenario 4. In the simulation, the ground truth changes with time stamp five for every third road segment. It depicts situations like the fix of road damages or the formation of new ones over time. The scenario should clarify that the back-end needs a kind of forgetfulness of previous classifications to guarantee a fast reaction. As benchmark it is counted how many of the $s = 10,000$ segments are wrongly classified when using the vote with precision matrix without forgetting. Additionally, we implemented a modified version, which has the ability to forget. Therefore, the vehicles observations $o^{(k)}$, with $k = 1, \ldots, 3$, are scaled down. For example, the classification of vehicle k from time step $t - i$ has only the weight f^i in time step t with $f \in [0, 1]$. A forget factor of $f = 0.5$ is used for the simulation. A higher factor means less forgetting because the observations are weighted stronger.

5. Results
The simulations have been performed according to the set-ups described in Section 4. The left side in Figure 5 shows the difference between the estimators for Scenario 1. The vote with precision matrix is the one with fewest wrong predictions. Already with one vehicle, the proportion of wrong predictions is below 0.2. That means that in average each segment is classified correctly. With increasing number of vehicles, the number of wrong predictions decreases faster compared to majority vote or diagonal vote. Between these two estimates no big difference can be identified. The right side in Figure 5 represents the results of the prediction based on the vote with precision matrix split

Figure 5. Results of Scenario 1.

Note: Each vehicle predicts the ground truth based on the global precision matrix.
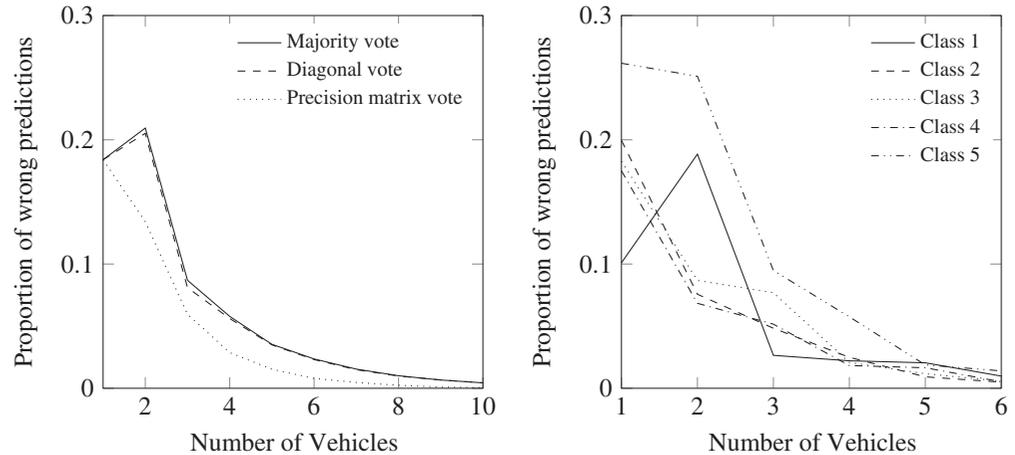
into the classes. The result shows a high impact of the class 5 on the number of wrong predictions. A look at the precision matrix clarifies that the low true positive rate $\pi_{55} = 0.73$ is the reason.

In Scenario 2, the simulation is extended such that each vehicle has a slightly adapted, individual version of the original precision matrix. However, for the fusion and the final decision the original one from Table 4 is used. As shown in Figure 6, again the vote with precision matrix is the most promising estimate. When comparing one vehicle, which drives multiple times over the road and classifies the segments (Figure 6, left), and multiple vehicles (Figure 6, right), the latter one achieves better results. The reason is that the fixed precision matrix of only one vehicle cannot compensate the difference to the global one. It can be seen that in this case the proportion of wrong predictions converges towards 0.1. This border is surpassed when using the classifications from five or more different vehicles recognizable by the right side in Figure 6.

Figure 7 shows the influence of a vehicle with *broken sensor* from Scenario 3 on the number of wrong predictions. In the left plot of the figure, the first vehicle has or has not an equally distributed precision matrix and therefore randomly predicts the classes. After five vehicles, in average less than 10% of the road is classified wrongly. After 10 vehicles, the impact of the bad precision matrix can be ignored. The right plot of Figure 7 shows the results of a simulation, in which the third vehicle has an equally distributed precision matrix, which is referred to as *broken sensor*. The solid line shows the proportion of misclassifications simulated the third vehicle with a *broken sensor* and no adjustments. As soon as the output of the third vehicle is included in the simulation, the number of wrong classifications increases greatly. With increasing the number of vehicles with good precision matrices, the number of wrong predictions converges again to 0. The dashed line shows the results of the

Figure 6. Results of Scenario 2. Unknown distribution of classification due to different precision matrices for vehicles.

Note: Left: one vehicle with more drives. Right: more vehicles each one drive.

cogent ·· engineering
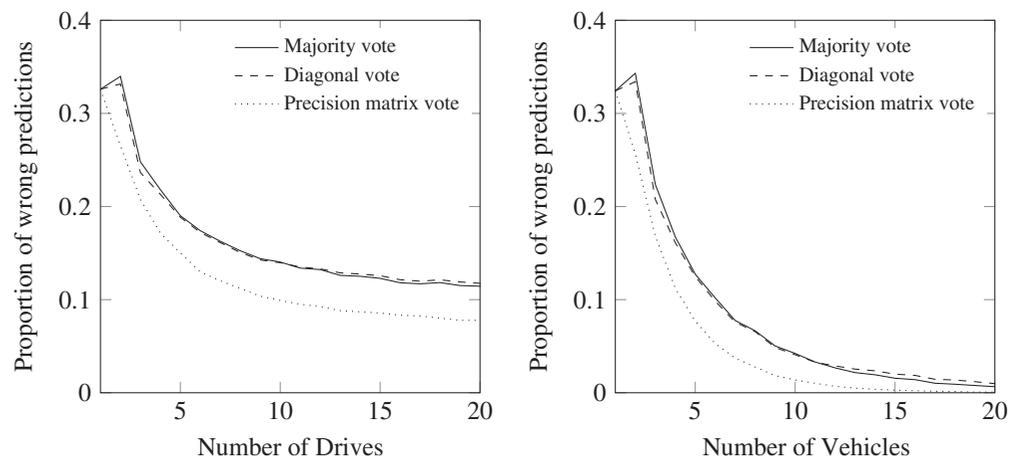
Figure 7. Results of Scenario 3. Classification with and without broken sensor.

Notes: Left: 1st vehicle with and without broken sensor. Right: third vehicle with broken sensor with and without automated adjusting precision matrix for this vehicle.
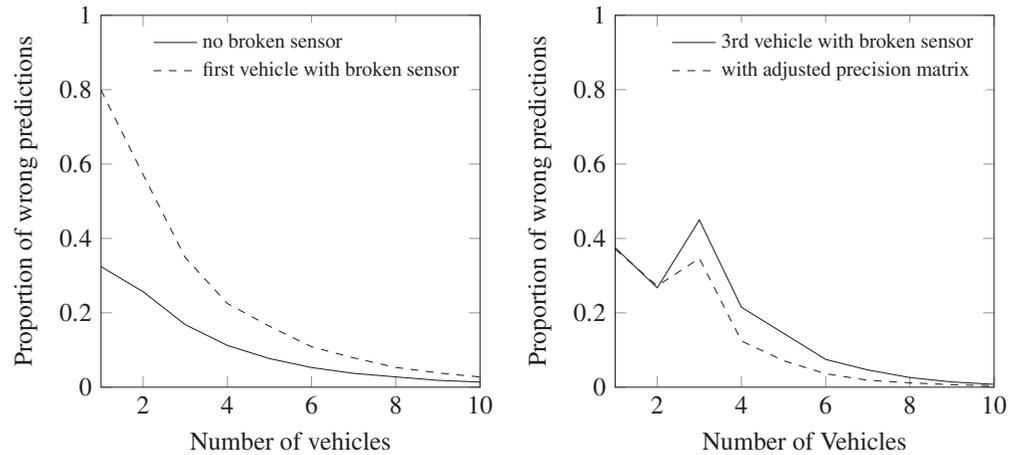
Figure 8. Results of Scenario 4. Change of the road conditions and simulated with and without forgetting of previous classifications.
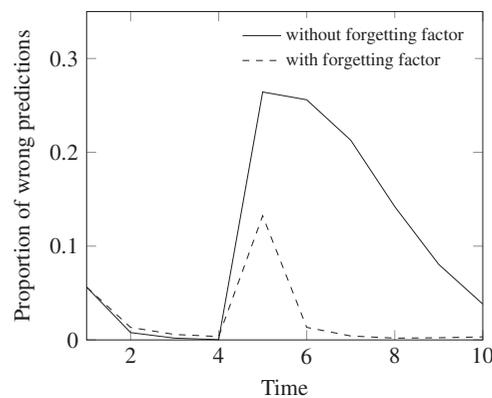
simulation with the identification of the vehicle with the bad precision matrix and the adjustment of the precision matrix in the back-end for this vehicle. With this adjustment, the increase in number of wrong predictions is not as great as without adjustment and the number of misclassifications faster converges to 0. Summarized, this method improves the results of our multiple vehicle fusion and can identify vehicles with bad classification outputs.

The results of Scenario 4 can be seen in Figure 8. After the change of the ground truth at time stamp $t = 5$, the vote with precision matrix without forgetting needs 4 more time stamps to reduce the proportion of wrong classified segments below 10%. When using the forget factor, already in the next time stamp an error rate below 2% can be reached. With forgetting, the accuracy does not only decrease faster. It is also higher immediately after changing the ground truth.

6. Conclusion

The paper addresses the problem of manual road infrastructure monitoring. In literature, the possibilities of training a classifier is already introduced. It is able to predict the road roughness or different road damages or conditions using vehicle sensor data. This knowledge is used to transfer the application of automated road infrastructure monitoring to a multiple expert problem.

Multiple vehicles predict the class for identical positions and the fusion leads to a robust estimate of the ground truth. Under consideration are different estimates and scenarios representing imaginable situations. The comparison of the estimates demonstrate that the results are more precise when the probability of the observations is calculated and the class with highest probability is selected. Only counting the predictions and vote for the class with majority leads to a lower accuracy.

Regardless of which estimate is used, the more vehicles classify one position, the better is the prediction of the ground truth. It can be concluded, that the aggregation of multiple predictions is a good way to increase the accuracy of the estimate of the road conditions.

For the next steps, the precision matrix, fundamental for the simulation, will be fine tuned and the proposed combination strategies and scenarios will be tested in reality. More data will be collected with multiple vehicles and under different conditions, such as rain and snow. In addition, a weighting up between fast reaction and certainty is necessary to determine a border.

Author details
Julia Hofmockel[1]
E-mail: julia.hofmockel@audi.de
Johannes Masino[2]
E-mail: johannes.masino@kit.edu
ORCID ID: http://orcid.org/0000-0002-4259-6311
Jakob Thumm[2]
E-mail: jakob.thumm@web.de
Eric Sax[3]
E-mail: eric.sax@kit.edu
Frank Gauterin[2]
E-mail: frank.gauterin@kit.edu

[1] Audi Electronics Venture GmbH, Gaimersheim, Germany.
[2] Institute of Vehicle System Technology, KIT, Karlsruhe, Germany.
[3] Institute for Information Processing Technologies, KIT, Karlsruhe, Germany.
[§] MATLAB code can be find on Hofmockel, Thumm, and Masino (2018), J. Hofmockel and J. Masino equally contributed to the paper.

Citation information
Cite this article as: Multiple vehicle fusion for a robust road condition estimation based on vehicle sensors and data mining, Julia Hofmockel, Johannes Masino, Jakob Thumm, Eric Sax & Frank Gauterin, *Cogent Engineering* (2018), 5: 1449428.

References
Aouaouda, S., Chadli, M., Boukhnifer, M., & Karimi, H.-R. (2015). Robust fault tolerant tracking controller design for vehicle dynamics: A descriptor approach. *mechatronics, 30*, 316–326.

Bououden, S., Chadli, M., & Karimi, H. R. (2016). A robust predictive control design for nonlinear active suspension systems. *Asian Journal of Control, 18*(1), 122–132.

Chen, K., Lu, M., Tan, G. & Wu, J. (2013). CRSM: Crowdsourcing Based Road Surface Monitoring. *High Performance Computing and Communications & 2013 IEEE International Conference on Embedded and Ubiquitous Computing (HPCCEUC), 2013 IEEE 10th International Conference On* (pp. 2151–2158). Cham: Springer.

Chu, H. C., & Hwang, G.-J. (2008). A delphi-based approach to developing expert systems with the cooperation of multiple experts. *Expert Systems with Applications, 34*(4), 2826–2840. Retrieved from http://www.sciencedirect.com/science/article/pii/S0957417407001777

Dahmani, H., Chadli, M., Rabhi, A., & El Hajjaji, A. (2013). Vehicle dynamic estimation with road bank angle consideration for rollover detection: theoretical and experimental studies. *Vehicle System Dynamics, 51*(12), 1853–1871.

Dawid, A. P., & Skene, A. M. (1979). Maximum likelihood estimation of observer error-rates using the EM algorithm. *Applied Statistics*, 20–28.

Descornet, G. (1990, January). Road-Surface influence on tire rolling resistance. In *Surface characteristics of roadways: international research and technologies*. ASTM International.

Eriksson, J., Girod, L., Hull, B., Newton, R., Madden, S. & Balakrishnan, H. (2008). The Pothole Patrol: Using a mobile sensor network for road surface monitoring. In *Proceedings of the 6th International Conference on Mobile Systems, Applications, and Services* (pp. 29–39). ACM.

Forschungsgesellschaft für Strassen- und Verkehrswesen. (2004). *Merkblatt über den Finanzbedarf der Straßenerhaltung in den Gemeinden*. FGSV-Verlag.

Hofmockel, J., Thumm, J. & Masino, J. (2018). *Multiple vehicle fusion for a robust road condition estimation based on vehicle sensors and data mining, Matlab files*. doi:10.5281/zenodo.1187068

Hu, C., Jing, H., Wang, R., Yan, F., & Chadli, M. (2016). Robust h output-feedback control for path following of autonomous ground vehicles. *Mechanical Systems and Signal Processing, 70*, 414–427.

Hwang, G. J. (1992). Knowledge elicitation and integration from multiple experts. In *Proceedings ICCI '92: Fourth International Conference on Computing and Information* (pp. 208–208).

Ihs, A. (2005). The influence of road surface condition on traffic safety and ride comfort. *Statens väg- och transportforskningsinstitut*, 11–21.

Kittler, J., Hojjatoleslami, S. & Windeatt, T. (1997). *Weighting factors in multiple expert fusion*. BMVC.

Koch, C., & Brilakis, I. (2011). Pothole detection in asphalt pavement images. *Advanced Engineering Informatics, 25*(3), 507–515.

Masino, J., Levasseut, G., Frey, M., Gauterin, F., Mikut, R. & Reischl, M. (1979). Charakterisierung der Fahrbahnbeschaffenheit durch Data Mining von gemessenen kinematischen Fahrzeuggrößen. *65*(12), 867–877.

Mavandadi, S., Feng, S., Yu, F., Dimitrov, S., Nielsen-Saines, K., Prescott, W. R., & Ozcan, A. (2012). A mathematical framework for combining decisions of multiple experts toward accurate and remote diagnosis of malaria using tele-microscopy. *PloS one, 7*(10), e46192.

Öhrström, E., Skånberg, A., Svensson, H., & Gidlöf-Gunnarsson, A. (2006). Effects of road traffic noise and the benefit of access to quietness. *Journal of Sound and Vibration, 295*(1–2), 40–59.

Raykar, V. C., Yu, S., Zhao, L. H., Jerebko, A., Florin, C., Bogoni, L., & Moy, L. (2009). Supervised learning from multiple experts: Whom to trust when everyone lies a bit. In *Proceedings of the 26th Annual International Conference on Machine Learning, ser. ICML '09* (pp. 889–896). New York, NY: ACM. doi:10.1145/1553374.1553488

cogent •• engineering

Sandberg, U. (Jan. 1987). Road traffic noise—The influence of the road surface and its characterization. *Applied Acoustics, 21*(2), 97–118.

Schwela, D. H. (2000). The World Health Organization guidelines for environmental noise. *Noise News International, 8*(1), 9–22.

Seraj, F., van der Zwaag, B. J., Dilo, A., Luarasi, T., Havinga, P. (2014). RoADS: A road pavement monitoring system for anomaly detection using smart phones. *Big data analytics in the social and ubiquitous context* (pp. 128–146).

Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing Management, 45*(4), 427–437.

Taylor, S. M., Breston, B. E., & Hall, F. L. (1982). The effect of road traffic noise on house prices. *Journal of Sound and Vibration, 80*(4), 523–541.

Theebe, M. A. J. (2004). Planes, trains, and automobiles: The impact of traffic noise on house prices. *The Journal of Real Estate Finance and Economics, 28*(2), 209–234.

Valizadegan, H., Nguyen, Q., & Hauskrecht, M. (2012). Learning medical diagnosis models from multiple experts. In *AMIA Annual Symposium Proceedings* (Vol. 2012, pp. 921). Rockville Pike, Bethesda MD: American Medical Informatics Association.

Wilhelmsson, M. (2000). The impact of traffic noise on the values of single-family houses. *Journal of Environmental Planning and Management, 43*(6), 799–815.

Cogent Engineering (ISSN: 2331-1916) is published by Cogent OA, part of Taylor & Francis Group.

Publishing with Cogent OA ensures:

- Immediate, universal access to your article on publication
- High visibility and discoverability via the Cogent OA website as well as Taylor & Francis Online
- Download and citation statistics for your article
- Rapid online publication
- Input from, and dialog with, expert editors and editorial boards
- Retention of full copyright of your article
- Guaranteed legacy preservation of your article
- Discounts and waivers for authors in developing regions

Submit your manuscript to a Cogent OA journal at www.CogentOA.com