RESEARCH ARTICLE

Road Surface Classification with Images Captured From Low-cost Camera - Road Traversing Knowledge (RTK) Dataset

Classificação de Superfície de Estrada com Imagens Capturadas de Câmera de Baixo Custo - Road Traversing Knowledge (RTK) Dataset

Thiago Rateke^{1,2*}, Karla Aparecida Justen², Aldo von Wangenheim^{1,2}

Abstract: The type of road pavement directly influences the way vehicles are driven. It's common to find papers that deal with path detection but don't take into account major changes in road surface patterns. The quality of the road surface has a direct impact on the comfort and especially on the safety of road users. In emerging countries it's common to find unpaved roads or roads with no maintenance. Unpaved or damaged roads also impact in higher fuel costs and vehicle maintenance. This kind of analysis can be useful for both road maintenance departments as well as for autonomous vehicle navigation systems to verify potential critical points. For the experiments accomplishment upon the surface types and quality classification, we present a new dataset, collected with a low-cost camera. This dataset has examples of good and bad asphalt (with potholes and other damages) other types of pavement and also many examples of unpaved roads (with and without potholes). We also provide several frames from our dataset manually sorted in surface types for tests accuracy verification. Our road type and quality classifier was done through a simple Convolutional Neural Network with few steps and presents promising results in different datasets.

Keywords: Road Surface Classification — Road Quality Classification — Convolutional Neural Network — Road Surface Analysis

Resumo: O tipo de pavimento de uma estrada influência diretamente na forma como os veículos são conduzidos. É comum encontrarmos trabalhos que lidam com a detecção de caminho, mas que não levam em consideração grandes alterações nos padrões da superfície da estrada. A qualidade da superfície da estrada tem um impacto direto no conforto e principalmente na segurança dos usuários da estrada. Em países emergentes é comum encontrar estradas não pavimentadas ou com pavimentos sem manutenção. As estradas sem pavimento ou com defeitos impactam também em maiores gastos com combustível e manutenção dos veículos. Esse tipo de análise pode ser útil tanto para departamentos de manutenção de rodovias checarem possíveis pontos críticos como para sistemas autônomos de auxílio a navegação veicular. Para a realização dos experimentos de classificação de tipos e qualidade da superfície, apresentamos um novo conjunto de dados, coletados com câmeras de baixo custo. Essa base conta com exemplos de asfalto bons e ruins (com buracos e outros danos) outros tipos de pavimento e também muitos exemplos de estrada de terra (com e sem buracos). Disponibilizamos também grande parte dos frames de nossa base manualmente classificados para verificação de acurácia em testes. Nosso classificador de tipos e qualidade de estrada foi feito através de uma Rede Neural Convolucional simples, com poucas etapas e apresenta resultados promissores em diferentes conjuntos de dados. **Palavras-Chave:** Classificação de Superfície de Estrada — Classificação de Qualidade de Estrada — Redes

Neurais Convolucionais — Análise de Superfície de Estrada

¹ Graduate Program in Computer Science (PPGCC), Federal University of Santa Catarina (UFSC), Brazil

² Image Processing and Computer Graphics Lab (Lapix) at Brazilian Institute for Digital Convergence (INCoD), Federal University of Santa Catarina (UFSC), Brazil

*Corresponding author: thiago@incod.ufsc.br

DOI: http://dx.doi.org/10.22456/2175-2745.91522 • Received: 01/04/2019 • Accepted: 28/09/2019

CC BY-NC-ND 4.0 - This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

1. Introduction

The road pavement type, or the lack of pavement, is an important information in vehicular navigation, as it involves the safety and comfort for users, since the presence of potholes or other damages, sudden change to a more slippery surface, among others factors may represent a challenge to control the vehicle and can cause accidents. It's also important for the vehicles conservation (eg.: suspension, tires).

According to [1], which is a study about the condition and the quality of road surfaces in the European Union (EU), road maintenance conditions don't cause as many accidents as human behavior while driving a vehicle cause (eg: driving without attention, drunk driving, driving aggressively, etc). However, the lack of maintenance on the road is rather responsible for accidents and a large part is related to the drivers reaction to road damages or changes in terrain type.

Road surface types classification, as well as the quality of this surface, is an important task for both Advanced Driver-Assistance Systems (ADAS) and Autonomous Navigation Systems, and also to Road Infrastructure Departments for road maintenance checks. With the context of Advanced Driver-Assistance Systems or in Autonomous Navigation Systems, the surface type and quality classification may indicate a better way of driving for user safety and comfort. With the context of Road Infrastructure Departments, can accelerate and improve the identification of critical points process, automatically identifying portions that require more attention and maintenance.

It's common to find in the literature studies focused on path detection that identify the path on different surfaces, but which don't extract relevant road features, such as surface type and also if there is some damage on that surface. ([2], [3], [4], [5], [6] and several other examples as shown in this Systematic Literature Review about road detection [7]).

There's approaches which classify the type of road related to the size of the road, such as: [8], which classifies into: "*Highway*" or "*Minor Roads*". The authors have used Convolutional Neural Networks (CNN) to perform the road classification.

In [9] the authors have also made a road classification using Convolutional Neural Networks (CNN), like [8] is not based on surface type, focusing on classifying the road as: *highway*, *freeway*, *urban road*. or as a *parking area*.

In emerging countries it's more common to find unpaved or poorly maintained roads. For example, in Brazil, from time to time a survey ([10]) is made on the federal and state roads to examine the roads quality. One of the aspects checked is precisely the floor, which takes into account the visible surface and also the base below that surface. The visible surface is the only layer perceivable by road users and it's on this layer that our work is based. In the material of [10] the analysis of the pavements brought a result where 37.0% of the roads were classified as "*Regular*", 9.5% as "*Bad*" and 4.4% as "*Poor*".

As another example, according to [11], in East Timor country 50% of the roads are unpaved. In this study the

authors made the classification if the road is paved or unpaved analyzing data obtained from sensors such as: accelerometer, gyroscope, GPS and compass. That is, the analysis is done as a mapping task, not as a prediction.

The work from [12] also focus in road quality classification as a mapping task, that is, mapping the quality of the roads after passing. Useful to road infrastructure departments. In [12] also no images were used, the authors used accelerometers and GPS.

Also related to the road quality, other approaches ([13] and [14]) make the pothole detection through laser sensors, such as LIDAR. However, according to material presented in [15] and [16] even LIDAR which is the most safety laser, can cause damage to the human eye in longer exposure (eg.: cataracts and burn the retina). In a hypothetical future where smart vehicles can turn out to be a reality in human life, this *lasersmog* can be a problem. With this in mind, we aim to work with techniques that focus on passive vision.

There are recent works that are focusing on the detection of damages in the pavement using passive vision (images), such as [17], [18] which use Stereo Vision to do the pothole detection (also called as: negative obstacles). And also [19] which use Deep Neural Network to do a bounding box detection and classification of damages on the asphalt road surface. But none of them classify the road surface type.

Other approaches perform the detection of potholes with very close and vertical view images to the pothole (eg.: [20], [21], [22], [23], [24]), which differs from the way and angle as a driver or a camera in a vehicle could see the road and the damages, doesn't giving time enough to avoid them. But still, they don't classify the type of surface or the quality.

In [7] it's possible to notice that there was a great advance in the use of Convolutional Neural Networks (CNN) together with Computer Vision (VC) to do the vehicular navigation tasks. In addition, the datasets available are mostly from developed countries where roads have good quality, no damage and/or no major variations in the road surface types.

In addition to the two works cited above, [8] and [9], that make the classification of road using CNN, in a different context from our propose (taking into account the road surface type). It's possible to find many other recent works in the literature that use CNN in image classification in different tasks (eg.: [25], [26] and [27]).

These are the reasons why we propose this work with an approach to classifying the road surface type and quality, which can be useful for different scenarios such as: ADAS, autonomous vehicles and road maintenance analysis. In this approach our focus is with passive vision perception (only images obtained by cameras) and with the use of Convolutional Neural Networks (CNN) for the surfaces classification task.

We use the same simple CNN structure, based on [28], to create four different models. A specific model to the surface type classification task. The other three models, to do the surface quality classification task, specific for each surface type resulting from the first model (Figure 1).



Figure 1. Our approach sequence

We also present a new dataset, collected with low-cost cameras and with large quantities of frames that vary in: different qualities of asphalt, different qualities of other types of pavement (eg.: cobblestones, concrete blocks, etc), different qualities of unpaved road and with brightness variation.

The remaining of this article is organized as follows: In Section 2 we provide a brief description of two datasets, already published, that we also used in our experiments. Our dataset is presented in Section 3. Our approach is detailed in Section 4. In Section 5 we present the results about our approach. We conclude this paper with a discussion in Section 6.

2. Datasets

In addition to our own images, we used in our experiments images from two different datasets already published in the literature: KITTI dataset¹ [29] and CaRINA dataset² [30].

Both bases are with high-quality image captures and with higher cost cameras. KITTI used 4 PointGray Flea2 cameras, one pair for grayscale capture and another pair for color capture. And CaRINA, used one Bumblebee XB3. Both projects used other sensors such as LIDAR (active vision). But our focus on this work is just with images (passive vision).

There are other bases which focus on Driver Assistance or Autonomous Navigation systems such as: CamVid dataset³ [31] from Cambridge, England. CityScapes dataset⁴ [32] from Germany. DIPLODOC ⁵ [33] near Trento city in Italy. However, they have few surface variations, without unpaved road options, and few (or none) examples of damaged surface. An exception would be the OffRoadScene dataset ⁶ [34]

⁵http://tev.fbk.eu/databases/diplodoc-road-stereo-sequence

from China, which is presented as having only unpaved and low quality roads (which would not help to check surface types variations). Also the link given in their paper is offline and we tried to get in touch with the authors, but didn't get any response.

More recently, the RoadDamageDetector dataset ⁷ [19] also presented a dataset with asphalt damage images, filmed in Japan, but in this case, only asphalt images were used, nevertheless, shows the interest in the surface quality verification and is another set available for experiments.

The KITTI dataset [29], from the Karlsruhe Institute of Technology, is an example with scenario in a developed country, its data were collected in Karlsruhe, Germany with rural, urban and highways. We maintained in our experiments, as it's currently one of the most used and cited dataset in the literature. And although it doesn't contain much variation in surfaces and also without damages and unpaved, there is enough lighting variation and also changes between asphalt and other pavement types.

The CaRINA dataset [30] was filmed in São Carlos city, São Paulo, Brazil, created by the Mobile Robot Laboratory (ICMC/USP - São Carlos). It's a recent dataset focused on the emerging countries scenario and shows more surface variation, having examples like: asphalt, other types of pavement and few unpaved road examples. They also give examples of damaged roads. The CaRINA dataset is composed by highly quality images.

3. Road Traversing Knowledge (RTK) Dataset

Here we present a new dataset, called Road Traversing Knowledge (RTK), filmed in Brazil, more specifically in Águas Mornas and Santo Amaro da Imperatriz cities, Florianópolis neighboring cities, in the Santa Catarina state. In this dataset, we collect road images with different surfaces.

In general, all images were collected with good weather conditions during the daytime period. There are lighting variations, shadows and a few frames with solar glare. Even within each road surface category there are differences in surface patterns, such as: older asphalts with lighter color patterns and new asphalts with darker color, other pavements different than asphalt and also with variations in pattern and color (eg.: cobblestone, concrete, etc), and also differences in unpaved surface types.

In Figure 2 we present part of our coverage roads for the RTK dataset. The area with the light red color is composed of a new and good quality asphalt without speed bumps, potholes or other damages. In the dark red color it's an area inside the city, with speed bumps, occurrences of potholes, cracks and patches on the road. In the orange color areas are occurrence of dirt road, which may have potholes, undulations and stones. And in the blue areas are different pavements.

⁻⁻⁻⁻⁻

¹http://www.cvlibs.net/datasets/kitti/raw_data.php

²http://www.lrm.icmc.usp.br/dataset

³http://mi.eng.cam.ac.uk/research/projects/VideoRec/CamVid/

⁴https://www.cityscapes-dataset.com/

⁶http://www.cim.mcgill.ca/li-jian/roaddatabase.html

⁷https://github.com/sekilab/RoadDamageDetector/



All frames were obtained with a low-cost camera attached to a moving vehicle. The camera used was a HP Webcam HD-4110 (Figure 3). The dataset can be accessed in: (http://www. lapix.ufsc.br/pesquisas/projeto-veiculo-autonomo/datasets?lang= $en\rangle$.



Figure 3. HP Webcam HD-4110

The dataset has 77547 frames from four collection sequences, with different surfaces types under different lighting conditions, consisting of: Asphalt roads, Paved roads and Unpaved roads.

Many of these frames contain transitions between surface types, which can be: asphalt to paved, paved to asphalt, paved to unpaved, unpaved to asphalt, which also allows transition identification.

We also provide part of the dataset sorted by road surface type, as a Ground Truth (GT), enabling accuracy testing. This GT with classified data consists of 62662 frames:

- Total classified frames as asphalt: 43263
- Total classified frames as paved: 13697
- Total classified frames as unpaved: 5702

The GT frames were manually separated by the authors. As the focus in this paper is to verify the classification of surface types, we avoid frames with transitions between different surface types. Thus, GT frames contain only one type of surface in each frame: just asphalt or just paved or just unpaved.

This dataset is intended to serve as an experiment for driver assistance and autonomous navigation systems projects and to the road conditions analysis for the road infrastructure departments. And show that low-cost sensors can be used in these projects.

In addition to road surface and quality classifying, this dataset allows image segmentation projects, as example: road shape detection. Although it was filmed in a calm and quiet region, inside the great amount of frames, there are enough vehicles and some situations with pedestrians and/or bikers on the road allowing the detection of vehicles, pedestrians or obstacles in general.

As previously stated, the dataset contains different surface types, including many unpaved road frames, and in different lighting conditions, as well as speed bump and different surface damages (Figure 4).



Figure 4. Samples from the RTK dataset

4. Our approach

In the Figure 5 we present our approach steps. The training of our model and our experiments were performed on a machine with a NVIDIA GeForce GTX 1060 GPU in an Ubuntu 18.04 system with 16 GB RAM. Our approach consists in the use of Convolutional Neural Networks (CNN) to do the road surface classification task.

Firstly, a Region of Interest (ROI) is defined as a preprocessing step for each input frame. With this pre-processing we aim to keep only the part of the image most relevant to the classification, which is the lower half of the image. We also cut few lines from the bottom of the image to avoid situations where part of the vehicle used to capture the images appears (Figure 6).

This ROI is hard-coded, and takes into account the image

dimensions of the 3 datasets used in our experiments. The same ROI was used for the 3 datasets (proportionally to the images resolution). Because they all have a similar placement, with front and nearly horizontal images, ROI was able to separate regions where more pixels belong to the road surface. This way we avoid most of the background and non-road pixels. We define the ROI in a hard-coded way because a possible adaptive ROI, which should find the horizon line automatically, if it fails in this step, can compromise the model training.

After this pre-processing, a data augmentation step is performed. This data augmentation consists of increasing and decreasing brightness in each frame. Trying to simulate the brightness and illumination which may vary in frames captured in different times of the day or different weathers. Some frames may be darker depending on the time of day they







Figure 6. Region-of-Interest. First row: our dataset. Second row: KITTI dataset. Third row: CaRINA dataset.

were captured or if it's a cloudy day, and likewise other frames may be lighter if it's on a sunny day. In this way, the training contains, in addition to the original ROI's, data augmentation from these ROI's.

Finally, the input images are passed to a convolution layer, with 32 filters with size of 3x3, followed by another equal convolution layer (32 filters with size of 3x3). and a final convolution layer with 64 filters with size of 3x3. All the strides were rated as 1 and Padding as 0. A Normal distribution is used for weights initialization. In the first convolution layer we pass all images selected to the training step with information about width, height and number of channels. A max-pooling is applied in all the convolution layers, in order to reduce the inputs dimensionality which helps to analyse features information in the input sub-regions. At the end of each convolution layer, after the max-pooling function, a ReLU is used as an activation function.

After the convolutional layers, the flatten layer is used to transform the convolution multi-dimensional tensor into a one-dimensional tensor. Two fully connected layers are added at the end. In the first fully connected layer a ReLU activation function is applied. The second fully connected layer has the possible outputs, the desired classes, which for road surface type model, road asphalt quality model and road

paved quality model will be 3 classes as outputs, and in road unpaved quality model will be with only two classes (this is because in our scope we define that an unpaved road should under no circumstances be considered as good quality).

We used the softmax function to achieve the probabilities of each class. This is because we consider that softmax makes more sense in situations where there is no subjectivity, such as the definition of pavement type. We also use the Adam optimizer at the end, which update the network weights based in the input data used in the training.

4.1 Road Surface Classification

To the road surface classification, we have trained our model with 6264 frames from our dataset, separated manually in 3 classes: asphalt, paved and unpaved. The number of frames selected for each class were proportional to the quantity in the dataset. Being around 70% asphalt (4344 frames), 20% paved (1337 frames) and 10% unpaved (585 frames).

We also did experiments with some images from KITTI (509 frames) dataset and CaRINA (156 frames) dataset in conjunction with our training data (from RTK dataset) checking different training models. From KITTI we have used 509 frames, being 478 of asphalt and 31 of paved. From CaRINA we have used 156 frames, being 97 asphalt, 49 paved and 10 unpaved. This allows us to check the results without occurrence of overfitting. We did previous experiments without ROI, and the validation result was 96.9%. With ROI definition the validation reached 100%.

We have separated the training data, being 80% for training and 20% randomly chosen for validation. The training step only with our data achieve the following numbers: Training Accuracy: 100%; Validation Accuracy: 100% and Validation Loss: 0.044. When we did the training adding the KITTI and CaRINA frames, we got the following numbers: Training Accuracy: 100%; Validation Accuracy: 93.8% and Validation Loss: 0.146.

4.2 Road Surface and Quality Classification

In the surface quality classification, we used three models equal to that used for the surface type classification, being a quality model for each type of surface. The quality models for asphalt and paved maintains the 3 outputs being: Good, Regular and Bad. The only difference is in the unpaved quality model, which contains only two outputs: Regular and Bad. We understand that this task is more subjective, but we try to define some criteria.

We consider based on [10] that the unpaved surface is the worst case scenario, so we did not create the Good class for unpaved roads. In this case, we use *Regular* for unpaved roads without potholes and other damages and *Bad* when there are potholes, stones or other damages.

The asphalt surfaces is considered *Good* if the road doesn't have damages and it's clean as possible. It's classified as *Regular*, if the road has patches or dirt. And it's classified as *Bad*, if it has damages, like potholes, water puddles, cracks, etc.

Paved surfaces are classified as *Good*, if the surface is clean as possible and without dirt or damage. It's considered *Regular* if the road is dirty. And if the road has damages or patches, it's classified as *Bad*.

We have used only few frames in the training step: 1579 frames of asphalt quality model; 630 frames of paved quality model; 497 frames of unpaved quality model. We have separated 80% of the data for training and 20% was randomly chosen for validation. For the asphalt quality model the training and validation results are: Training Accuracy: 100%; Validation Accuracy: 100% and Validation Loss: 0.011. For the paved quality model the training results are: Training Accuracy: 100%; Validation Accuracy: 96.9% and Validation Loss: 0.190. For the unpaved quality model the training results are: Training Accuracy: 100%; Validation Accuracy: 96.9% and Validation Loss: 0.207.

5. Results

First, we did tests exclusively for the road surface classification model. In these tests, we checked the accuracy of both the model: trained only with our dataset (Table 1) and also with the model that we used KITTI and CaRINA data together with our dataset (Table 2).

We ran this tests in our manually classified set, consisting of 62662 frames. More 5093 frames from KITTI and more 1401 frames from CaRINA, we also did the manual sorting of these frames to verify the accuracy of the model. Our results accuracy in the test data are presented in Table 1 and in Table 2. In these test data, we avoided frames that contained transitions between surface, and also avoided frames where the solar glare was too strong causing great reflection.

It's possible to realize that training with images from other datasets, in addition to improving the results for those datasets, did not have a negative impact on our own dataset.

Table	1. Results accuracy
First model ((trained only with RTK)

	Asphalt	Paved	Unpaved	Total
RTK	98.58%	85.75%	93.84%	95.73%
KITTI	54.70%	23.40%	-	53.07%
CaRINA	58.82%	6.61%	100.00%	45.30%

Table 2. Results accuracy

 Second model (trained with 3 datasets)

	Asphalt	Paved	Unpaved	Total
RTK	98.45%	86.24%	93.67%	95.72%
KITTI	94.22%	73.21%	-	93.13%
CaRINA	97.66%	100.00%	100.00%	98.36%

We also tried training with the balanced dataset in the data augmentation step, trying to leave each surface type with a similar amount of samples. As shown in Table 3 there was a slight improvement overall, as well as a slight worsening from our dataset results, and our dataset has more images of other surface types, mainly unpaved.

 Table 3. Results accuracy

 Third model (3 datasets and with balanced data)

	Asphalt	Paved	Unpaved	Total
RTK	96.81%	86.05%	94.41%	94.57%
KITTI	94.49%	76.23%	-	93.54%
CaRINA	99.90%	100.00%	99.00%	99.86%

With these results it's possible to notice that paved road frames obtained the worst results, with exception of CaRINA dataset. Even so, we got reasonable values: 73.21% or 76.23% in KITTI and 86.24% or 86.05% in RTK. We also present the road type classification confusion matrices results for each dataset: Table 4 for RTK dataset, Table 5 for KITTI dataset and Table 6 for CaRINA dataset.

Table 4. Surface type confusion matrix for RTK dataset

	Asphalt	Paved	Unpaved
Asphalt	96.81%	10.24%	4.12%
Paved	2.68%	86.05%	1.47%
Unpaved	0.51%	3.71%	94.41%

Table 5. Surface type confusion matrix for KITTI dataset

	Asphalt	Paved	Unpaved
Asphalt	94.49%	22.64%	-
Paved	5.12%	76.23%	-
Unpaved	0.39%	1.13%	-

Table 6. Surface type confusion matrix for CaRINA dataset

	Asphalt	Paved	Unpaved
Asphalt	99.90%	0.00%	1.00%
Paved	0.10%	100.00%	0.00%
Unpaved	0.00%	0.00%	99.00%

Also in Figures 7, 8 and 9 we present some results respectively in the RTK, CaRINA and KITTI images.



Figure 7. Road surface classification results from RTK dataset. First row only asphalt. Second row only paved. Third row only unpaved. Fourth row with some of the classification faults, in the first column where it should be paved it was classified as asphalt, in the middle column it should be unpaved and was classified as asphalt and in the third column it should be asphalt and classified with unpaved.



Figure 8. Road surface classification results from CaRINA dataset.



Figure 9. Road surface classification results from KITTI dataset.

It's important to emphasize that in the training images selection step we avoid many images with vehicles, because this is a very common occurrence in asphalt roads, usually with more movement, could induce the network to errors, such as when there are cars the model assumes it's an asphalt.

Even taking this care, there were more cars on the asphalt surfaces in training step, some of the errors may have this as cause, such as the middle column in the fourth row of Figure 7. However, in the same Figure, in row three, also in the middle column, we can notice a correct classification even with the presence of vehicles.

For the surface quality classification we used the previously road surface classification trained model as the first step. Based on the output of the surface type model we then call the specific quality model. Obviously, with this approach, if the road surface classification results comes with an error from the first classification model all the remainder process is compromised. On the other hand, as a great number of hits from the first model, the quality classification models is able to specialize for the surface quality of its specific surface type.

We ran this tests in our manually classified set, consisting of 6327 frames and also from KITTI were 606 frames and from CaRINA were 721 frames. Our results accuracy in the test data are presented in Table 7.

In Table 7 we can see that the results for surface quality classification are still better for Asphalt types. And with a lot of variations for Paved and Unpaved types. For the Paved(Good) type the model had reasonable accuracy for KITTI data, and a poor accuracy for Paved(Regular), Paved(Bad) and Unpaved(Bad) at RTK and Paved(Bad) at CaRINA. Otherwise, CaRINA's Paved(Good) went very well, as well as Unpaved(Bad). Paved(Good) and Unpaved(Regular) types on RTK achieved reasonable results.

The confusion matrices, for each dataset, can help to understand the results obtained in Table 7. The confusion matrix for RTK dataset is presented by Table 8, for KITTI dataset in Table 9, and for CaRINA dataset in Table 10. In bold are the correct values expected for each class, in blue the errors in lateral classes, closest to the expected, in red the most distant classes. The other values represent errors that come from the previous model, ie from the surface type classification.

In the Figures 10, 11 and 12 we present some results from the surface quality classification, respectively for RTK, CaRINA and KITTI images. In these Figures it's possible to see the quality classification results for different pavement types.

In Figure 10, the first three rows present the expected results. In the fourth row, three examples of error are presented. From the left column the error came from the first model, even though it was *Bad* unpaved. In the middle and right columns, the quality model was wrong (inverted).

6. Conclusions, Discussions and Future Works

Many studies and advances have been made focused on autonomous vehicles, since road detection and segmentation to obstacle detection. It's certainly that new challenges and scenarios arise, such as damaged and/or unpaved roads.

Identifying the road surface type is a important task for both Advanced Driver-Assistance Systems (ADAS) and Autonomous Vehicles, as depending on the road surface the driver or the autonomous vehicle must adapt its driving behavior. Also assists in identifying roads that require pavement or maintenance on the pavement.

Deep Learning advances in last years makes possible the ability to train a model to do the image classification, automating the road surface type detection task and also the quality of that surface has become more feasible and faster.

In this work we presented a new dataset with roads images containing different types of surface, with and without damages. In addition, the images from the dataset were obtained through a low-cost camera. In this way, enabling experiments in a more challenging scenario in an emerging country.

We also presented our experiments with a simple Convolutional Neural Network (CNN) to perform the surface types classification between: asphalt, paved and unpaved. We reached good results in a large amount of data in different datasets.

The difference obtained by the test results, seen between the Table 1 and Table 2 raise some questions about network generalization. This difference is due to overffiting, because the first table shows the tests with network trained only with our dataset, with images containing different resolutions compared to the other two datasets, different cameras used, and also different camera settings. Most likely, if we try to run the trained model with images from other datasets that were not used in our training, the result tends not to be so positive.

To have a more generic network, we should use images from more datasets. Increasing the chance of network learning regardless of resolution or capture sensor settings. But selecting this data and classifying it for training can be a time-consuming task.

However, if we think about the applications that are part of our motivation in this work (ADAS, Autonomous Vehicles and Road Maintenance), it usually has a specific configuration, previously defined (vehicle and cameras), and it's only necessary to perform training containing the data of that specific settings. In this case it may be better even if they don't generalize the network, performing the training only with the settings that will be used. This doesn't change the fact that in this work we show that it's possible to perform this surface analysis, even with a simple neural network structure together with low cost camera images.

We also tested a structure for road surface quality classification. In this structure we used the first model, the surface type classification, as a first step, which already had good results. Then, the specific quality model for a given surface type is used. The surface quality evaluation has a more subjective aspect, but our experiments showed that it's possible to obtain reasonable results (mainly to the asphalt patterns) with the use of images and deep learning.

6.1 Future Works

From this work we can visualize some future works, improvements and possibilities. For example in relation to the dataset we can try to obtain images that increase the level of difficulty and verify if it's possible to obtain results with passive vision at night, in rainy conditions or with fog situations. A classification of weather based on images can be useful as a pre-processing step as this can help define if camera sensors will be advantageous in any given climate and whether to trigger different sensors to ensure safety.

For the surface types classification, we can try to specify the model in more classes, considering different asphalt conditions, different pavements (eg.: cobblestone, concrete). We can also have transition classes, generating a warning in advance when a surface type will changes, this situation forces an adaptation in our approach, because our pre-processing step with ROI generation would cut part of the image that contains the road transition. In this case, a segmentation network that differentiates the surface types should give a more precise answer, although it will have to appeal to a more robust network structure and consequently greater computational power and also more effort on annotating GT images.

Finally, regarding to the road quality classification, a specialist in road analysis and maintenance helping to train the quality model would allow greater robustness and reliability. Allowing a specific Ground Truth creation to the road quality in the dataset.

	Asphalt(Good)	Asphalt(Regular)	Asphalt(Bad)	Paved(Good)	Paved(Regular)	Paved(Bad)	Unpaved(Regular)	Unpaved(Bad)
RTK	98.23%	94.04%	96.77%	73.71%	45.37%	9.68%	74.37%	19.73%
KITTI	97.85%	ı	I	83.67%	I	I	I	I
CaRINA	98.35%	97.50%	94.12%	%06	I	%0	I	100%

Table 7. Surface quality results accuracy

dataset
RTK
for
matrix
confusion
quality
Surface
ω
Table

		Iable	o. ourrace quant	y comusion mar	ITY TOT VIN DALASCI			
	Asphalt(Good)	Asphalt(Regular)	Asphalt(Bad)	Paved(Good)	Paved(Regular)	Paved(Bad)	Unpaved(Regular)	Unpaved(Bad)
Asphalt(Good)	98.23%	1.79%	0.22%	3.56%	7.41%	9.68%	1.51%	4.72%
Asphalt(Regular)	0.35%	94.04	2.16%	2.63%	2.78%	1.61%	0.38%	1.35%
Asphalt(Bad)	0.15%	0.83%	96.77%	0.59%	1.85%	1.61%	0%0	0.67%
Paved(Good)	0.40%	0%0	0%0	73.71%	33.95%	29.84%	%0	1.85%
Paved(Regular)	0.05%	3.10%	0%	4.33%	45.37%	43.55%	0%0	0%
Paved(Bad)	0%0	0%0	0%	13.57%	7.72%	9.68%	%0	0.84%
Unpaved(Regular)	0.81%	0%0	0%0	1.19%	0.31%	4.03%	74.37%	50.59%
Unpaved(Bad)	0%0	0.24%	0%0	0%	0.62%	0%	14.45%	19.73%

[dataset
KITTI
for
matrix
confusion
quality
Surface
Table 9.

	Asphalt(Good)	Asphalt(Regular)	Asphalt(Bad)	Paved(Good)	Paved(Regular)	Paved(Bad)	textbfUnpaved(Regular)	Unpaved(Bad)
Asphalt(Good)	97.85 %	1	I	8.16%	I	I	I	I
Asphalt(Regular)	%0	•	I	0%	I	I	I	I
Asphalt(Bad)	%0	1	I	0%	I	I	I	I
Paved(Good)	1.97%	I	Ĩ	83.67 %	-	-	I	I
Paved(Regular)	0%0	I	Ĩ	6.12%	•	I	I	I
Paved(Bad)	0%0	I	Ĩ	2.04%		I	I	I
Unpaved(Regular)	0.18%	I	-	0%0	I	I	•	I
Unpaved(Bad)	%0	I	I	0%	I	I	I	1

aset
, dat
NA
ıRI
Ű
for
atrix
1 mí
sio
onfu
ty c
quali
ace (
urf
s.
10
e
q
Ta

			outlace quality	COLLASION IIIAUL	A TUL CAINEN AND VALA	301		
	Asphalt(Good)	Asphalt(Regular)	Asphalt(Bad)	Paved(Good)	Paved(Regular)	Paved(Bad)	Unpaved(Regular)	Unpaved(Bad)
Asphalt(Good)	98.35%	%0	0%0	%0	I	0%	I	0%
Asphalt(Regular)	0%0	97.50%	5.88%	0%0	I	0%0	I	0%0
Asphalt(Bad)	0.21%	0%0	94.12%	%0	I	0%	I	0%0
Paved(Good)	1.44%	960	0%	<i>2</i> /2006	1	100%	I	0%0
Paved(Regular)	%0	2.50%	0%0	%0	·	%0	I	0%
Paved(Bad)	%0	0%0	0%	10%	1	0%0	I	0%
Unpaved(Regular)	%0	%0	0%	%0	I	0%	•	0%0
Unpaved(Bad)	200	260	0%0	20		260	-	100%



Figure 10. Road surface quality classification results from RTK dataset.



Figure 11. Road surface quality classification results from CaRINA dataset.



Figure 12. Road surface quality classification results from KITTI dataset.

Acknowledgements

This study was financed in part by the Coordenação de Aperfei - çoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. CAPES (Brazilian Federal Agency for Support and Evaluation of Graduate Education).

Author contributions

- *Thiago Rateke*: Collected the images for the RTK dataset. Designed and developed the experiments. Wrote the paper.
- *Karla Aparecida Justen*: Helped to plan, organize, and collect the images for the RTK dataset. Participated in the analysis of the results with ideas and discussions about the experiments. Revised the text.
- *Aldo von Wangenheim*: Supervised the whole work. Perform the critical review of the manuscript.

References

[1] FRISONI, R. et al. *EU Road Surfaces: Economic and Safety Impact of the Lack of Regular Road Maintenance*. [S.1.], 2014. Url: (http://www.europarl.europa.eu/thinktank/ en/document.html?reference=IPOL_STU(2014)529059).

[2] WANG, H.; REN, M.; YANG, J. Capitalizing on the boundary ratio prior for road detection. *Multimedia Tools and Applications*, v. 75, n. 19, p. 11999–12019, Oct 2016. Doi: (https://doi.org/10.1007/s11042-016-3280-y).

[3] VALENTE, M.; STANCIULESCU, B. Real-time method for general road segmentation. In: 2017 IEEE Intelligent Vehicles Symposium (IV). [S.l.: s.n.], 2017. p. 443–447. Doi: (https://doi.org/10.1109/IVS.2017.7995758).

[4] YUAN, Y.; JIANG, Z.; WANG, Q. Video-based road detection via online structural learning. *Neurocomputing*, v. 168, p. 336 – 347, 2015. Doi: (http://dx.doi.org/10.1016/j.neucom. 2015.05.092).

[5] ZU, Z. et al. Real-time road detection with image texture analysis-based vanishing point estimation. In: 2015 *IEEE International Conference on Progress in Informatics and Computing (PIC)*. [S.l.: s.n.], 2015. p. 454–457. Doi: (http://dx.doi.org/10.1109/PIC.2015.7489888).

[6] SHI, J. et al. A novel path segmentation method for autonomous road following. In: 2016 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC). [S.1.: s.n.], 2016. p. 1–6. Doi: (http: //doi.acm.org/10.1109/ICSPCC.2016.7753701).

[7] RATEKE, T. et al. Passive vision region-based road detection: A literature review. *ACM Comput. Surv.*, ACM, New York, NY, USA, v. 52, n. 2, p. 31:1–31:34, mar. 2019. Doi: (http://doi.acm.org/10.1145/3311951).

[8] TEICHMANN, M. et al. Multinet: Real-time joint semantic reasoning for autonomous driving. 2018 IEEE Intelligent Vehicles Symposium (IV), p. 1013–1020, 2018. Doi: (https://doi.org/10.1109/IVS.2018.8500504).

[9] SEEGER, C. et al. Towards road type classification with occupancy grids. In: *IEEE Intelligent Vehicles Symposium 2016 Workshop: DeepDriving - Learning Representations for Intelligent Vehicles*. [S.l.: s.n.], 2016. p. 1–4. Url: (http://iv2016.berkeleyvision.org/papers/seeger.pdf).

[10] CNT. *Pesquisa CNT de Rodovias 2016. Relatório Gerencial.* [S.l.], 2018. v. 20. Url: (https://pesquisarodovias.cnt.org. br/Home).

[11] Cabral, F. S. et al. An automatic survey system for paved and unpaved road classification and road anomaly detection using smartphone sensor. In: 2018 IEEE International Conference on Service Operations and Logistics, and Informatics (SOLI). [S.l.: s.n.], 2018. p. 65–70. Doi: (https://doi.org/10.1109/SOLI.2018.8476788).

[12] TARIQ, H.; MAZHAR, S.; HAMEED, H. Road quality classification for road repair authorities and regular drivers, using an on-board data logger: Poster abstract. In: *Proceedings of the 17th ACM/IEEE International Conference on Information Processing in Sensor Networks*. Piscataway, NJ, USA: IEEE Press, 2018. (IPSN '18), p. 142–143. Doi: (https://doi.org/10.1109/IPSN.2018.00034).

[13] Kang, B.; Choi, S. Pothole detection system using 2d lidar and camera. In: 2017 Ninth International Conference on Ubiquitous and Future Networks (ICUFN). [S.I.: s.n.], 2017. p. 744–746. Doi: (https://doi.org/10.1109/ICUFN.2017. 7993890).

[14] Yu, X.; Salari, E. Pavement pothole detection and severity measurement using laser imaging. In: 2011 IEEE INTER-NATIONAL CONFERENCE ON ELECTRO/INFORMATION TECHNOLOGY. [S.l.: s.n.], 2011. p. 1–5. Doi: (http://doi. acm.org/10.1109/EIT.2011.5978573).

[15] COMMISSION, I. E. SAFETY OF LASER PRODUCTS. Part 1: Equipment classification, requirements and user's guide. [S.l.]: Geneva : International Electrotechnical Commission, 2001, 2001.

[16] STANDARD, A. N. American National Standard for Safe use of Lasers Outdoors. Orlando, FL, 2005.

[17] HERGHELEGIU, P.; BURLACU, A.; CARAIMAN, S. Negative obstacle detection for wearable assistive devices for visually impaired. In: 2017 21st International Conference on System Theory, Control and Computing (ICSTCC). [S.l.: s.n.], 2017. p. 564–570. Doi: (http://doi.acm.org/10.1109/ICSTCC. 2017.8107095).

[18] KARUNASEKERA, H. et al. Stereo vision based negative obstacle detection. In: 2017 13th IEEE International Conference on Control Automation (ICCA). [S.l.: s.n.], 2017. p. 834–838. Doi: (https://doi.org/10.1109/ICCA.2017. 8003168). [19] MAEDA, H. et al. Road damage detection and classification using deep neural networks with smartphone images. *Computer-Aided Civil and Infrastructure Engineering*, v. 33, n. 12, p. 1127–1141, 2018. Doi: (https://doi.org/10.1111/mice. 12387).

[20] ERIKSSON, J. et al. 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. New York, NY, USA: ACM, 2008. (MobiSys '08), p. 29–39. Doi: (http://doi.acm.org/10.1145/ 1378600.1378605).

[21] KOCH, C.; BRILAKIS, I. Pothole detection in asphalt pavement images. *Advanced Engineering Informatics*, v. 25, n. 3, p. 507 – 515, 2011. Special Section: Engineering informatics in port operations and logistics, doi: (https://doi.org/10.1016/j.aei.2011.01.002).

[22] HUIDROM, L.; DAS, L. K.; SUD, S. Method for automated assessment of potholes, cracks and patches from road surface video clips. *Procedia - Social and Behavioral Sciences*, v. 104, p. 312 – 321, 2013. 2nd Conference of Transportation Research Group of India (2nd CTRG), doi: (https://doi.org/10.1016/j.sbspro.2013.11.124).

[23] TEDESCHI, A.; BENEDETTO, F. A real-time automatic pavement crack and pothole recognition system for mobile android-based devices. *Advanced Engineering Informatics*, v. 32, p. 11 – 25, 2017. Doi: (https://doi.org/10.1016/j.aei. 2016.12.004).

[24] BANHARNSAKUN, A. Hybrid abc-ann for pavement surface distress detection and classification. *International Journal of Machine Learning and Cybernetics*, v. 8, n. 2, p. 699–710, Apr 2017. Doi: (https://doi.org/10.1007/ s13042-015-0471-1).

[25] WANG, X. et al. Facial expression recognition with deep learning. In: *Proceedings of the 10th International Conference on Internet Multimedia Computing and Service*. New York, NY, USA: ACM, 2018. (ICIMCS '18), p. 10:1–10:4. Doi: (http://doi.acm.org/10.1145/3240876.3240908).

[26] Vo, A. T.; Tran, H. S.; Le, T. H. Advertisement image classification using convolutional neural network. In: 2017 9th International Conference on Knowledge and Systems Engineering (KSE). [S.l.: s.n.], 2017. p. 197–202. Doi: (https://doi.org/10.1109/KSE.2017.8119458).

[27] Li, Y. et al. Sleep posture classification with multistream cnn using vertical distance map. In: 2018 International Workshop on Advanced Image Technology (IWAIT). [S.I.: s.n.], 2018. p. 1–4. Doi: (https://doi.org/10.1109/IWAIT. 2018.8369761).

[28] SACHAN, A. Tensorflow Tutorial 2: image classifier using convolutional neural network.
2017. Url: (https://cv-tricks.com/tensorflow-tutorial/training-convolutional-neural-network-for-image-classification/).

[29] GEIGER, A. et al. Vision meets robotics: The kitti dataset. *Int. J. Rob. Res.*, Sage Publications, Inc., Thousand Oaks, CA, USA, v. 32, n. 11, p. 1231–1237, sep 2013. Doi: (http://dx.doi.org/10.1177/0278364913491297).

[30] SHINZATO, P. Y. et al. Carina dataset: An emergingcountry urban scenario benchmark for road detection systems. In: 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC). [S.l.: s.n.], 2016. p. 41–46. Doi: (http://dx.doi.org/10.1109/ITSC.2016.7795529).

[31] BROSTOW, G. J.; FAUQUEUR, J.; CIPOLLA, R. Semantic object classes in video: A high-definition ground truth database. *Pattern Recognition Letters*, v. 30, n. 2, p. 88 – 97, 2009. Doi: (https://doi.org/10.1016/j.patrec.2008.04.005).

[32] Cordts, M. et al. The cityscapes dataset for semantic urban scene understanding. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). [S.I.: s.n.], 2016. p. 3213–3223. Doi: (https://doi.org/10.1109/CVPR. 2016.350).

[33] ZANIN, M. et al. *diplodoc road stereo se-quence*. 2013. Url: (https://tev.fbk.eu/databases/ diplodoc-road-stereo-sequence).

[34] SHANG, E. et al. Offroadscene: An open database for unstructured road detection algorithms. In: 2013 International Conference on Computer Sciences and Applications. [S.l.: s.n.], 2013. p. 779–783. Doi: (http://dx.doi.org/10.1109/CSA. 2013.186).