Preprint of the paper to appear in Proc. IEEE International Conference on Intelligent Transportation Systems, 2021.

Self-Supervised Learning of Camera-based Drivable Surface Friction

David Vosahlik, Jan Cech, Tomas Hanis, Adam Konopisky, Tomas Rurtle, Jan Svancar, Tomas Twardzik
Faculty of Electrical Engineering, Czech Technical University in Prague

***Abstract*— The visual predictor of a drivable surface friction ahead of the vehicle is presented. The image recognition neural network is trained in self-supervised fashion, as an alternative to tedious, error-prone, and subjective human annotation. The training images are labelled automatically by surface friction estimates from vehicle response during ordinary driving. The Unscented Kalman Filter algorithm is used to estimate tire-to-road interface friction parameters, taking into account the highly nonlinear nature of tire dynamics. Finally, the overall toolchain was validated using an experimental subscale platform and real-world driving scenarios. The resulting visual predictor was trained using about 3 000 images and validated on an unseen set of 800 test images, achieving 0.98 cross-correlation between the visually predicted and the estimated value of surface friction.**

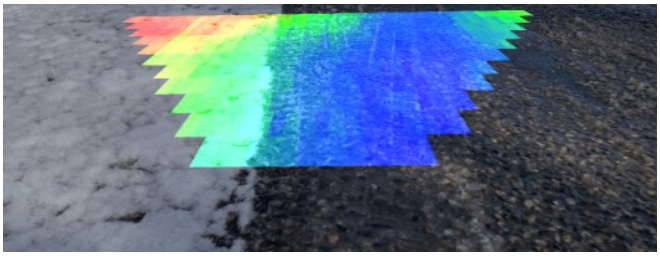


Fig. 1: Color-coded friction of the surface found by a visual predictor that was trained only by self-supervision from vehicle response data, *without any manual annotation*. Colder colors encodes higher friction (wet tarmac), while warmer colors lower friction (snow).

I. Introduction and Related Work

The road condition has a severe impact on vehicle performance, maneuver safety, and passenger comfort. The unsteady nature of environmental conditions and road quality is known to every human driver. Therefore, the instant and continuous assessment of the terrain the vehicle is about to negotiate is a common and valuable practice for any experienced driver. The camera and image recognition systems are essential building blocks and technology-enabling factors for many automotive functionalities, especially the Advance Driver Assistance Systems. The critical part of image recognition tool development, the neural network training phase, heavily relies on the training dataset and labeling process accuracy. Manual image labeling is a tedious and time-costly part of a machine learning process. On top of that, a human annotator is prone to subjective evaluation and errors. The large-scale driving data collection, both visual and road properties, combined with automatic and objective road condition evaluation, enables a cheap and reliable alternative to manual data labeling.

The algorithms presented in this paper provide three principal functionalities. Furthermore, all parts combined are providing an unique Visual-Advance-Drive-Assistance-System (V-ADAS) functionality.

First, the vision-based algorithm is trained using real-world data. The presented approach brings the possibility of continuous data collection during ordinary driving, without any dedicated equipment or vehicle instrumentation needed. This is demonstrated in section III, where the experimental subscale platform is put in use to collect the visual and vehicle response data for training of Neural Networks (NN).

Second, the collected visual data are then annotated in an automatic and objective manner, based on vehicle dynamics and traction system response. The road friction is estimated based on Unscented Kalman Filter (UKF) algorithm, see section II.

Finally, once all functionalities are combined, the predictive visual information about riding conditions the car is about to drive is provided. The proposed V-ADAS functionality provides essential information to the traction system and thus increases vehicle active safety. The merit of visual assisted driving systems was already shown e.g. in [1].

A. Related Work

The road condition properties like the friction or skid resistance coefficient are accurately measured by dedicated devices and almost in a laboratory-like fashion. The specialized equipment like the British pendulum [2] evaluates the road condition locally. The continuous and larger-scale road measurement is performed by grip testers [3]. The grip testers measure a generated friction force using a measuring wheel with defined loading and braking torque. The device is typically towed in a trailer or uses a fifth wheel extended from a test vehicle. Finally, a smart infrastructure provides in-road sensors, e.g., [4], or a road condition has been estimated from static traffic cameras [5]. The dedicated equipment usually provides just isolated location information or is limited to a one-time measurement ride. On top of that, the dedicated equipment is expensive and not widely available.

Besides the dedicated devices and infrastructure, road friction from a vehicle has been estimated by various techniques. We refer to a recent survey [6]. The taxonomy roughly splits into two groups: (1) Effect-based methods that estimate the friction properties by measuring vehicle response, typically using wheel torque and slip ratio, and (2) Predictive methods that estimate the surface friction ahead of the vehicle before it

travels on it. The latter methods typically use a front-looking camera and machine learning. Other modalities used to estimate surface type or friction include laser [7], sound [8], ultrasonic [9], or polarized light [10].

The effect-based methods are mainly built upon the wheel and vehicle dynamics. Most of the works estimate the peak friction coefficient μ_{max}, which is shown in Fig. 3 (Maximum of the slip curve). Papers use various complexity models of tire, wheels, and vehicle overall dynamics. In [11], the authors propose Kalman filter for force estimation and recursive least squares for the friction coefficient estimation. Polynomial slip curve approximation fitting is proposed in [12]. The fitted polynomial function is parameterized by μ_{max}. Therefore, the estimated μ_{max} parameter is obtained already for low slip ration. The Unscented Kalman Filter (UKF) is utilized in [13] to estimate vehicle states that are used as inputs for a neural network algorithm providing the μ_{max} parameter estimation. The Lyapunov method is used in [14] for the maximum friction coefficient estimation. In [15], authors propose a combined vehicle, wheel, and tire-to-road interface model used in UKF, which then provides the desired estimate of the friction coefficient. The method also incorporates the sensitivity analysis of the information contained in the measured data.

The camera-based friction estimation methods typically translate into a problem of image classification into a set of discrete semantic classes, e.g., dry, wet, snowy, or tarmac, cobblestone, gravel, etc. A recent survey on these methods can be found in [16]. Both classical [17], [18] or deep learning techniques are harnessed [19], [20], [21], [22]. The surface friction prediction has recently been employed in the improved Anti-lock Braking System (ABS) [23], [1]. A combination of vehicle response and image recognition methods can be found in [24].

The models mentioned above were learned from manually labeled datasets. We propose a method where the friction recognition model is learned from images labelled automatically. A work similar in spirit is the master's thesis [25], where a combination of manual labels, interpolated and filtered friction estimates using a device of NIRA dynamics [26] was used in training. The NIRA system, as reported relying on ABS and TCS, is not open and the details on friction estimates are unknown. Moreover, the friction is discretized into 6 classes or 3 classes (low/middle/high) of friction levels. On the other hand, our method provides a continuous estimate of the surface friction, besides the details on the vehicle response friction estimates.

For a different problem, a drivable surface roughness, a similar self-supervision method was used to train a visual predictor by associating the surface roughness measured by an accelerometer with the camera images [27].

II. Drivable Surface Properties Estimation

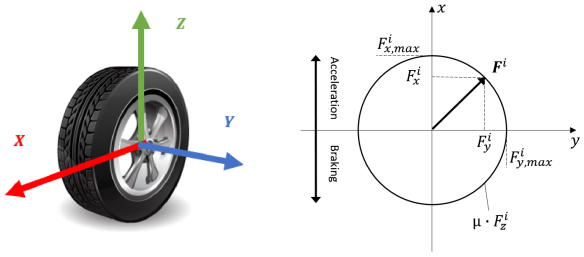
Identification of surface friction ϕ will be described in this section. The goal of such an algorithm is to provide a measurable value which will label the driven surface. The surface label will be later on used for training (and experimental validation) of the proposed neural network. Tire-to-road interface properties, including friction, are in this paper modeled by well-known Pacejka magic formula introduced in [28]. To properly use the formula for the force computation, vehicle dynamical models are to be introduced. Using mathematical models and measurements, the Unscented Kalman filter (UKF) algorithm is developed to estimate the surface friction ϕ. Detailed design and validation by the simulations are described in the following subsections. Experimental validation is presented at the end of this section.

A. Models

The high fidelity nonlinear mathematical models used for simulations and for estimator derivation will be described in this subsection. Twin-track nonlinear model [29] is used as the simulation model. Implementation of the model is in [30]. The twin-track model used is composed of the following parts:

- Nonlinear vehicle dynamics with suspension.
- 4 wheels where each is composed of
  - Pacejka magic formula tire model.
  - Wheel dynamics.
  - Friction ellipse (Kamm's circle).

Vehicle rigid body dynamics and suspension models are not presented as those are not crucial parts of the model in this paper. Details about these models can be found in [29]. Remaining model parts will be described further in the following subsections.



(a) Wheel coordinate system. (b) Wheel traction ellipse.

Fig. 2: Wheel model.

1) Wheel dynamics: Coordinate system connected to the wheel is shown in figure 2a. Wheel dynamics is characterized by the wheel torque equation

$$J_i \cdot \dot{\omega}_i = \tau_i + F^i_x \cdot r_{w_i} + \tau_{drag}, \tag{1}$$

where the subscript i represents i-th wheel, J_i is lumped powertrain and wheel moment of inertia, ω_i stands for wheel angular speed along the wheel hub axis, τ_i is applied wheel traction torque, F^i_x is traction force generated by the particular wheel in x direction, r_{w_i} is wheel effective radius and τ_{drag} represents all resistant torques such as rolling resistance induced torque, friction torque and others if applicable. For different road conditions, the variation in wheel normal force and longitudinal force reduction due to

the lateral force effect are modeled by different shapes of the friction ellipse (see Fig. 2b) and the slip curve (see Fig. 3). All these effects combined are sketched as

$$F_x^i = f(\lambda_i, F_z^i, \mu, F_y^i), \tag{2}$$

where λ_i is wheel longitudinal slip ratio (introduced below), F_z^i is wheel normal/load force, μ is friction coefficient of tire to road interface and F_y^i is wheel lateral force (perpendicular to wheel traveling direction – wheel y direction).

For the model completeness, the longitudinal slip ratio λ and wheel slip angle α are introduced. Wheel longitudinal slip ratio λ is defined as

$$\lambda_i = \frac{\omega_i \cdot r_{w_i} - v_x^i}{\max(|\omega_i| \cdot r_{w_i}, |v_x^i|)}, \tag{3}$$

where i represents i-th wheel, v_x^i stands for wheel pivot point travelling speed in x direction.

Wheel slip angle α is defined as

$$\alpha_i = -\arctan(\frac{v_y^i}{v_x^i}), \tag{4}$$

where v_y^i is wheel pivot point lateral velocity (in y direction).

2) Tire-to-road traction model: The Pacejka magic formula is used to describe the interaction of tire-to-road interface and evaluate the force generated by the interface. Forces in this formula depend on the so-called slip variables. In the longitudinal direction x, the slip variable is slip ratio λ (3). Slip variable in lateral direction y is the wheel slip angle α (4). The formula is parameterized by a set of parameters that are specific for given set of tires, road surface, and more. Examples of the Pacejka's slip curves for different road conditions are shown in the Fig. 3. The Pacejka magic formula itself is

$$F_x(\lambda) = \mu F_z D \sin(C \arctan(B\lambda - E(B\lambda - \arctan(B\lambda)))), \tag{5}$$

where the slip ratio λ is used as the slip variable, μ is road friction coefficient, B, C, D and E are Pacejka coefficients characterizing the slip curve shape, and F_z stands for wheel normal force. The lateral variant follows the same structure as in Eq. (5). The only difference is in the usage of the lateral slip variable – slip angle α and sets of Pacejka coefficients that are applicable for lateral tire direction.

3) Friction ellipse: The traction capacity of a tire is limited by the normal force acting on the tire and the magic formula parameters. To respect this limitation, the so-called friction ellipse is implemented in the simulation model. The graphical representation of a friction ellipse is depicted in Fig. 2b. The friction ellipse constrains the longitudinal and lateral force generated by the Pacejka magic formula to be lower than the wheel normal force

$$\mu F_z \geq \sqrt{F_x^2 + F_y^2}. \tag{6}$$

The implementation of the friction ellipse algorithm is adopted from [31], [32], [28] and its description is not presented here for the sake of brevity.

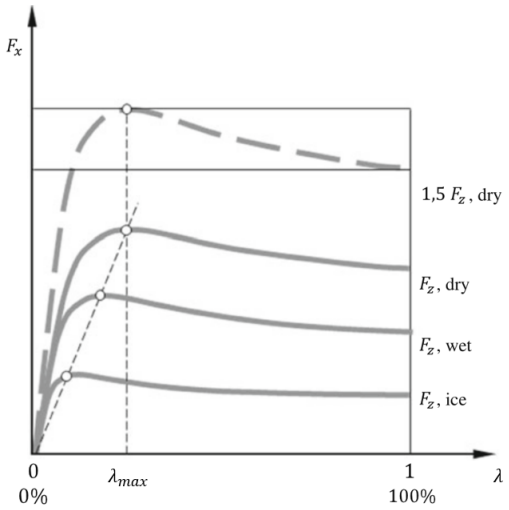


Fig. 3: Longitudinal slip curve. Adopted from [29].

B. Force estimator

It was already mentioned earlier that the UKF filter, which is the actual heart of the whole estimation algorithm, requires the longitudinal force as an input. As force measurement is a challenging problem and force sensors are not cheap, a force estimation algorithm based on the wheel dynamics Eq. (1) is proposed in this work. The estimator is inspired by the work presented in [33]. The discretized wheel dynamics is used in force estimation and presented below

$$\omega[t+1] = \omega[t] + \frac{T_s \cdot r_w}{J} F_x^i[t] + \frac{T_s}{J}(\tau[t] + \tau_{drag}[t]), \tag{7}$$

where T_s is the sampling time. Based on the equation, the force estimator is proposed as

$$F_x^i[t-1] = \frac{J}{T_s \cdot r_w}(f(\omega[t] - \omega[t-1])) - \frac{1}{r_w}(\tau[t] + \tau_{drag}[t]), \tag{8}$$

where $f(\omega[t] - \omega[t-1])$ represents the low pass filtered wheel speed difference.

C. UKF estimator

In the Pacejka magic formula, the term $\mu F_z D$ represents friction ϕ, Eq. 11. The ultimate goal of the Unscented Kalman filter (UKF) algorithm is to estimate this value. The dual filtering problem arises here from the need of parameter estimation. To overcome the problem, the Kalman-Schmidt algorithm is used, see [34]. System model in the form $x[t+1] = f(x[t], u[t], p[t])$ and $y[t] = g(x[t], u[t], p[t])$ is needed for the UKF (p represents parameters, x states and u inputs). Extension of the state is introduced to fit into the UKF state estimation framework

$$\tilde{x}[t] = \begin{bmatrix} x[t] \\ p[t] \end{bmatrix}. \tag{9}$$

The vector of parameters p in this case represent the Pacejka magic formula parameters B, C, D, E and the state x represents wheel angular speed ω. The overall dynamics equation

used in UKF is presented below as

$$f(\hat{\tilde{\mathbf{x}}}[t], \tau_i[t]) = \begin{pmatrix} \frac{T_s}{J}(\tau_i[t] + r_{w_i} \cdot F_x(\hat{\tilde{\mathbf{x}}}[t])) + \hat{\tilde{x}}_1[t] \\ \min(\max(\hat{\tilde{x}}_2[t], B_{min}), B_{max}) \\ \min(\max(\hat{\tilde{x}}_3[t], C_{min}), C_{max}) \\ \min(\max(\hat{\tilde{x}}_4[t], \phi_{min}), \phi_{max}) \\ \min(\max(\hat{\tilde{x}}_5[t], E_{min}), E_{max}) \end{pmatrix}, \quad (10a)$$

$$g(\hat{\tilde{\mathbf{x}}}[t]) = \begin{pmatrix} \hat{\tilde{x}}_1[t] \\ F_x(\hat{\tilde{\mathbf{x}}}[t]) \end{pmatrix}, \quad (10b)$$

$$F_x(\hat{\tilde{\mathbf{x}}}, \lambda) = \hat{\tilde{x}}_4 \sin(\hat{\tilde{x}}_3 \arctan(\hat{\tilde{x}}_2\lambda - \hat{\tilde{x}}_5(\hat{\tilde{x}}_2\lambda - \arctan(\hat{\tilde{x}}_2\lambda)))). \quad (10c)$$

It is seen that the parameters (states x_2 to x_5) are bounded (min and max values). This is to keep the parameters in a meaningful range. It can be seen from the equations (10a), (10c) and (5) that the state $\hat{\tilde{x}}_4 = \phi$. The F_z term is assumed constant which does not bring a significant error into the estimation as confirmed by the simulations, see Fig. 5. Under these assumptions the $\hat{\tilde{x}}_4$ variation represents what is intuitively understood under friction. Friction ϕ is expressed as

$$\phi = \mu F_z D. \quad (11)$$

The slip-ration λ is reduced to reflect the friction ellipse effect, as described in Sec. II-A.3.

Summarized assumptions on the overall estimation algorithm are:

- The wheel angular velocity is measured/estimated
- The wheel pivot point velocity vector is measured/estimated
- Parameters of the wheel dynamics (1) are known or identified
- The τ_{drag} from (1) is known/identified

D. Simulation results

The proposed algorithm was at first validated in simulations using the nonlinear high fidelity twin track model described in Sec. II-A.

The simulation results of the force estimator, described in Sec. II-B, are presented in Fig. 4. Force estimation performance shown in the first subfigure is effected by the measurement noise. The noise is visible in the second subfigure where the wheel speed measurement is shown.

Results of the simulations are presented in Fig. 5. Both figures are from the same scenario. The vehicle is driving the whole time on the same surface.

Relative error of the estimated slip curve force for slip ratio values in [-0.15, 0.15] range is plotted in Fig. 5. The computation of the estimation integral error e is

$$e\left(\hat{F_x}(\lambda), F_x(\lambda), \lambda_{max}, \lambda_{min}\right) = \int_{\lambda_{min}}^{\lambda_{max}} |F_x(\xi) - \hat{F_x}(\xi)| d\xi. \quad (12)$$

The plotted error in Fig. 5 is normalised with respect to the maximum error value in the time interval. Actual slip ratio absolute value is plotted in the same figure for illustration. The slip ratio excites the wheel dynamics and therefore the relative estimation error drops rapidly as the slip ratio rises. Without the excitation, no information about the surface friction is available and the estimation value and error remain unchanged.

III. Visual Predictor Training

The goal of our method is to train a visual-only predictor, using a convolutional neural network (CNN), that takes an image captured by a front-looking camera as an input and provides an estimate of the corresponding surface friction. To train the CNN, a dataset is collected by freely driving while recording the camera and vehicle response signals. The training images are labeled automatically by associating with the surface friction estimates ϕ computed by the method described in the previous section.

A. Image labeling

The overview of the visual predictor is sketched in Fig. 6. A raw image captured by a camera is first orthographically rectified to a bird's-eye view. The homography mapping is used to warp the raw image of a scene rectangle on the ground plane in front of a vehicle into a rectified image, such that the corners of the scene rectangle and the corners of the rectified image correspond. The rectified image is then used as an input to a convolutional neural network.

To train the CNN, the training images need to be labeled. To label the image captured at time t_0, the corresponding surface friction $\phi(t_0)$, Eq. (11), is found. We define a fixed measurement distance d in the half of the scene rectangle for all of our experiments. Then the time t_0, when the vehicle travels on the surface over the measurement distance d, is calculated using the vehicle velocity as $t_0 = d/v(t)$, assuming the vehicle trajectory is straight.

The extent of the input image, which is used by the CNN to predict the surface friction, was found experimentally. A too small region around the measurement distance would not provide enough context, while a too large region would capture an irrelevant background.

In principle, image rectification is unnecessary. Nevertheless, this is an easy way to normalize images and possibly combine different cameras with different viewpoints without retraining. In all our experiments, the homography is estimated offline. Any deviations from the estimated mapping, which may occur due to camera tilting because of mounting on the vehicle body, are neglected.

B. Convolutional neural network

Using the process above, we collected a labeled dataset $\{(I_1, \phi_1), \ldots, (I_n, \phi_n)\}$. RESNET-50 [35] was used as the backbone, with the input of the rectified RGB image I resized to 224×224 px receptive field. The network has a single scalar output, after ReLU, the estimated surface friction $\text{CNN}(I; \Theta) = \hat{\phi}$. Vector Θ concatenates all network weights.

L_2-regression loss $L(\Theta) = \sum(\phi_i - \hat{\phi}_i)^2$ was used as the loss function penalizing differences between the ground

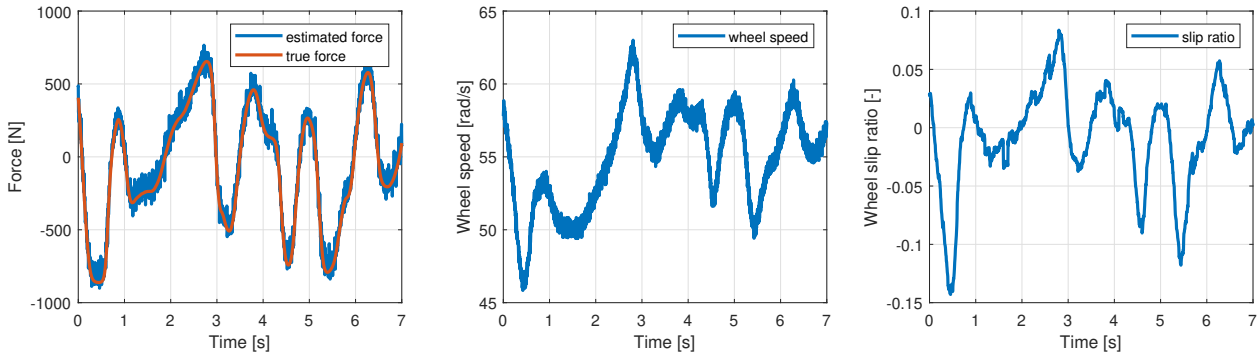


Fig. 4: Comparison between the estimated and the true traction force together with wheel speed and slip ratio. Simulations with nonlinear model described in Sec. II-A.

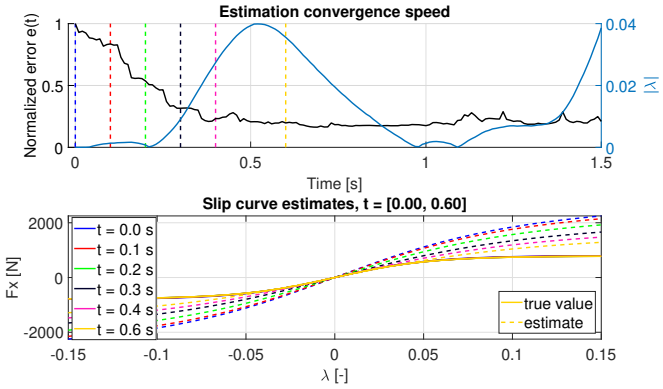


Fig. 5: Convergence of the estimator for bad initial value. Even for low slip ratio the estimator converges reasonably.

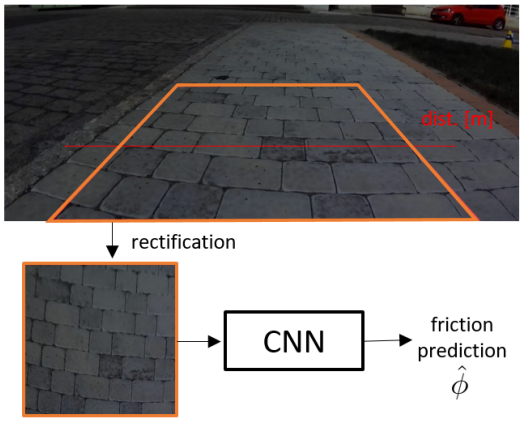


Fig. 6: Overview of the visual predictor. An input image from the vehicle camera is first rectified and cropped, so the rectified image were capturing a rectangular region of size 1.5m $\times$ 1.5m in front of the vehicle (in yellow). The image is than fed into a CNN to do the prediction of the friction at measurement distance of d, set as $d = 0.75$m (delineated as a red horizontal line).

truth and predicted labels of the surface friction. Batch-normalization layers were inserted and network weights Θ were found by ADAM optimizer [36]. As data augmentation, we used color jitter (contrast, brightness, hue), and random horizontal flipping. Only small $\pm 4^\circ$ image rotations were made, since direction of the image texture may be important for the perceived surface friction. The training converged after about 50 epochs.

IV. Experiments

A. Experimental subscale vehicle platform

The real-word data, both camera data and vehicle traction system data, were collected using an experimental subscale platform, shown in Fig. 7. The LOSI buggy Radio Controlled car of 1:5 scale was modified for this purpose. The overall dimensions are length: 844mm, width: 501mm, height: 308mm. The ground clearance is 69mm and the wheelbase is 552mm. Finally, the bare platform without data acquisition equipment is 13.8kg with traction batteries.

The platform was mechanically modified to accommodate the camera and vehicle data acquisition system. The traction system was modified from four-wheel drive to rear wheel drive to obtain a reliable and accurate vehicle traveling speed derived from the front wheel RPM sensors.

The subscale platform is equipped with four Electronic Control Units (ECU), providing vehicle motion control, camera and vehicle data acquisition and necessary safety of operation functionality. The high level ECU hardware is Intel *NUC7i7BNK* mini PC with Matlab & Simulink and Python software capabilities. The low-level traction control system, data acquisition, and safety features are performed by Raspberry Pi 4, Arduino Nano, and STM32 Nucleo (*STM32L432*) microcontrollers. The measurement system sampling time is set to 10ms. The vehicle measurements are:

- The vehicle velocity is estimated from the nondriven wheels (front wheels) as a mean value $v_x = \frac{r(\omega_1+\omega_2)}{2}$.
- Wheel angular speed is measured using Hall effect-based sensors and magnetic signs on the wheel.
- BLDC motor torque is estimated from two-phase current measurements. Ideal differential operation is assumed. Therefore, an equal split of the torque for both wheels is assumed.
- The drivable surface images are captured with a stereo ZED (StereoLabs). A single camera of the rig is used.

Fig. 7: Sub-scale vehicle platform used in our experiments.

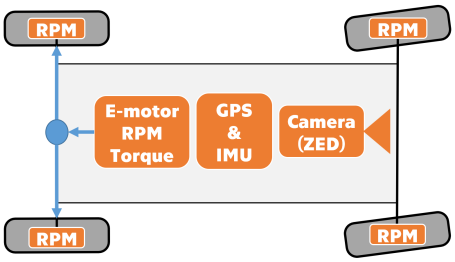


Fig. 8: Sub-scale vehicle platform sensors instrumentation.

B. Experimental platform estimation results

The proposed estimation algorithm was experimentally validated using the platform described in IV-A. Sampling time of the estimator is the same as the platform.

As the lateral velocity is not measured on this platform, only straight drive maneuvers are used in the measurement update phase of the UKF (no side slip angle of the wheel is considered). This situation also simplifies the wheel pivot point velocity measurement in x direction – it is the same as the velocity measured by the nondriven wheels (see Sec. IV-A). The friction torque was identified as a function of wheel angular speed $\tau_{drag} = \tau_{drag}(\omega)$. Parameters of the wheel dynamics equation were also identified. Wheel angular velocity is measured as described in IV-A. This way, all assumptions listed in Sec. II are satisfied.

The estimation algorithm was performed offline after the riding, but is ready for online execution (nature of the UKF filtering). Results of three experiments on three different surfaces are shown in Fig. 9. The surfaces are snow, gravel, and asphalt with the highest adhesion of the presented surfaces. For each surface, two estimated values of $\hat{\tilde{x}}_4$ (for two driven wheels) are shown in the Fig. 9. The $\hat{\tilde{x}}_4$ corresponds to the friction ϕ. The maximum of both wheel friction ϕ estimates was taken as the surface label that is later used for neural network training. Segmentation of the friction estimate for different surfaces can be clearly seen in the Fig. 9.

The estimation algorithm needs sufficient excitation of the dynamics to provide good estimates as described in the Sec. II-D. To provide the neural network friction estimates

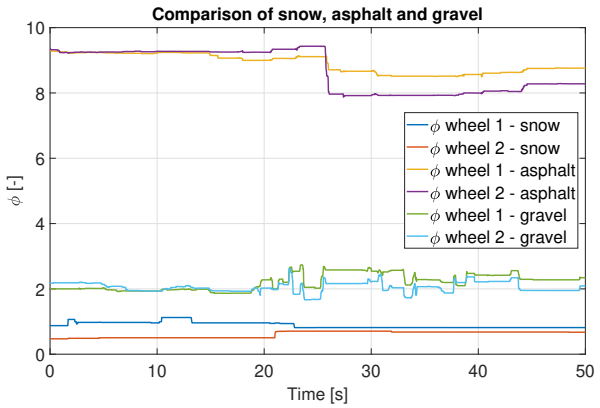


Fig. 9: Comparison of the estimated friction for different surfaces based on the subscale platform measured data.

that are not influenced by the convergence of the initial value and poor dynamics excitation, two runs of the UKF algorithm were executed. The second run of the estimation algorithm is initialized by the average value of the last 30 seconds of the first run. This approach improves the results for surfaces with piecewise constant parameters, while not compromising the algorithm performance for the surface with parameters changing even within a single run.

C. Dataset

The dataset was collected by driving the vehicle on various surfaces. The driving maneuvers consist of the vehicle acceleration and deceleration phases with the usual value of slip ratio in [-0.4, 0.4] range. During extreme maneuvers, the slip ratio was reaching values of 0.9. We acquired about 2.5 hours of synchronous raw recordings of all vehicle sensors and control signals, including the camera. Test rides were performed at 5 different locations, each containing a collection of different surfaces, e.g., dry/wet tarmac, snow, pavement, gravel, etc. Recordings were repeated at different days and day times to capture the effect of illumination.

As discussed above, both the friction estimation method implementation for the subscale platform and the image labeling assume that the vehicle goes approximately straight. Therefore, the data with significant deviations from the straight direction were discarded in the postprocessing.

We took images every 0.5s to keep the dataset of a reasonable size with enough diversity. The resulting dataset, automatically labeled, includes about 4 000 samples. The dataset was split into disjoint subsets: the training subset (70%), the validation subset (10%) to select the best training epoch, and the test subset (20%) used for evaluation. We made sure the test set does not overlap with the training data by using location and temporal metadata.

A sample of our dataset together with the distribution of the ground-truth friction labels of surface friction is shown in Fig. 10.

D. Evaluation and results

The testing was performed on the independent test split of our dataset. The results are summarized in Tab. I. The statis-

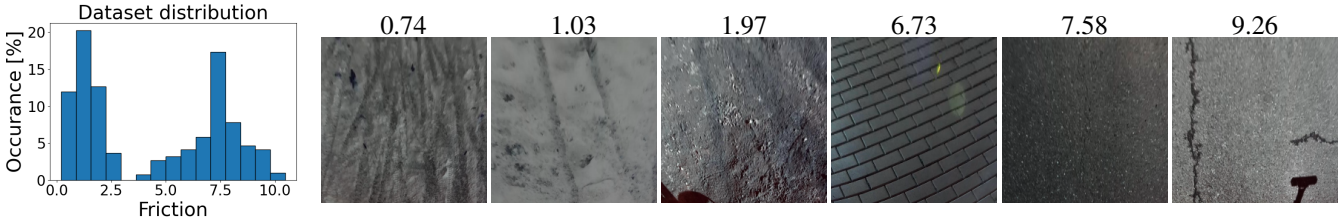


Fig. 10: Distribution of surface friction ϕ in our datasets. Samples sorted from the lowest to highest friction.

MAE	RMSE	Corr	P_{95}	$e_{0.5}$
0.36755	0.56746	0.9824	1.34325	78.17%

TABLE I: Error statistics

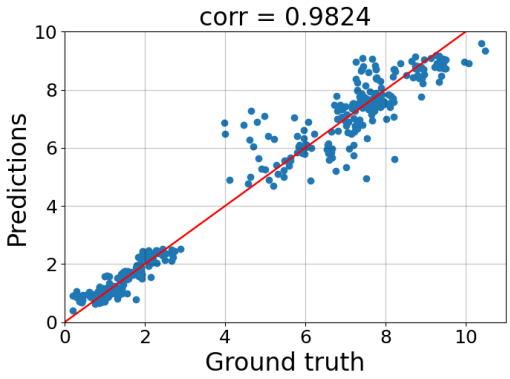


Fig. 11: Scatter plots for the proposed CNN visual prediction. Ideal predictions would lie on the red diagonal line. Correlation coefficients are shown in title of the plots.

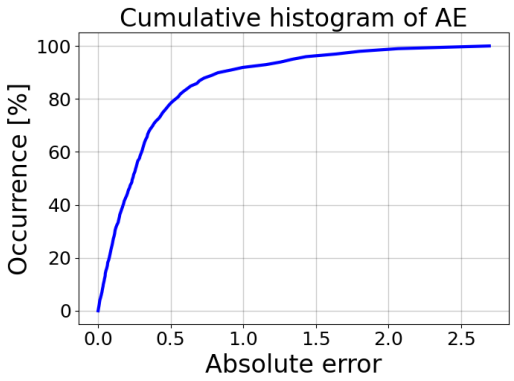


Fig. 12: Cumulative histograms of absolute error.

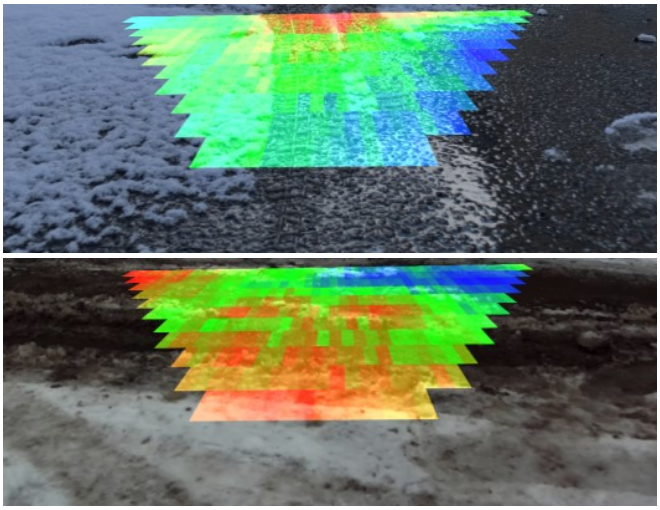


Fig. 13: Color-coded surface friction maps calculated by the trained CNN executed in a scanning window over the input image rectified to bird's-eye view, colorized by the estimated friction, and finally back-projected to the raw camera image.

tics are Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Correlation coefficient (Corr), Percentile-95 (P_{95}) meaning the prediction error lower or equal for 95% of test samples, and Error-0.5 ($e_{0.5}$) showing the percentage of samples having prediction error lower than or equal to 0.5.

Besides the statistics, we show a scatter plot in Fig. 11 visualizing the correlation between the ground truth and the predicted surface friction labels. In Fig. 12, we show the cumulative histograms of absolute errors to provide an insight on the error distributions. Statistics P_{95} and $e_{0.5}$ are easily seen in the plots. All statistics corroborate a high visual prediction accuracy.

E. Friction maps – qualitative results

The following experiment demonstrates that the trained model generalizes to a slightly different problem, predicting a coarse map of the surface friction.

Our CNN-model was trained to predict the local surface friction in the center of the given image, i.e., a single scalar for an image. The input to the CNN represents a square of $1.5\text{m} \times 1.5\text{m}$ in the scene. Therefore, we scan a larger area of the surface seen by the camera with the CNN evaluated at many locations. In particular, the scanning is done in the rectified bird's-eye view, and partially overlapping images of $1.5\text{m} \times 1.5\text{m}$ windows are one by one fed to the CNN. The outputs are stored, the center window locations are colorized based on the surface friction, and finally back-projected to the raw camera image.

Results are shown in Fig. 1 and Fig. 13. We can see that the boundary between the surfaces of different frictions is visible. This experiment is presented as a qualitative result, since we do not have the ground-truth data for other locations than the measurement point (0.75m in front of the vehicle for an image). The CNN was not trained for other locations, and

it is rather surprising that it generalizes well for far distant locations (up to 5 meters), where a loss of resolution degrades the image quality. The performed experiment demonstrates the accuracy and spatial consistency of the predictor.

V. Conclusions

We presented a method for surface friction estimation from vehicle responses and used the estimates from test rides to automatically label the images observed by the camera to train a visual predictor without any manually annotated data. An automatically labeled dataset of about 4 000 images was collected, and a convolutional neural network was trained. The method was evaluated on the independent test split of the dataset, achieving accurate predictions. As a qualitative result, we showed a coarse friction maps found by scanning the input image with the trained CNN predictor.

Acknowledgment

The research was supported by Toyota Motor Europe, and by CTU student grant under project SGS19/174/OHK3/3T/13.

References

[1] M. Bahnik, D. Filyo, D. Pekarek, M. Vlasimsky, J. Cech, T. Hanis, and M. Hromcik, "Visually assisted anti-lock braking system," in *Proc. IEEE Intelligent Vehicles Symposium (IV)*, 2020.

[2] L. Chu, W. Guo, and T. F. Fwa, "Theoretical and practical engineering significance of british pendulum test," *International Journal of Pavement Engineering*, 2020.

[3] W. Kleine-Beek, "Runway friction characteristics measurement and aircraft braking (rufab), volume 3: Functional friction," European Union Aviation Safety Agency, Tech. Rep., 2010.

[4] Lufft, "Road & runway sensors," 2021. [Online]. Available: https://www.lufft.com/products/road-runway-sensors-292/

[5] D. Grabowski and A. Czyzewski, "System for monitoring road slippery based on cctv cameras and convolutional neural networks," *Journal of Intelligent Information Systems*, vol. 55, pp. 521–534, 2020.

[6] S. Khaleghian, A. Emami, and S. Taheri, "A technical survey on tire-road friction estimation," *Friction*, vol. 5, no. 2, pp. 123–146, 2017.

[7] M. Aki, T. Rojanaarpa, K. Nakano, Y. Suda, N. Takasuka, T. Isogai, and T. Kawai, "Road surface recognition using laser radar for automatic platooning," *IEEE Trans. on Intelligent Transportation Systems*, vol. 17, no. 10, 2016.

[8] M. Kalliris, S. Kanarachos, R. Kotsakis, O. Haas, and M. Blundell, "Machine learning algorithms for wet road surface detection using acoustic measurements," in *Proc. IEEE International Conference on Mechatronics (ICM)*, 2019.

[9] A. Bystrov, E. Hoare, T.-Y. Tranb, N. Clarke, M. Gashinova, and M. Cherniakov, "Road surface classification using automotive ultrasonic sensor," *Procedia Engineering*, vol. 168, 2016.

[10] M. Jokela, M. Kutila, and L. Le, "Road condition monitoring system based on a stereo camera," in *International Conference on Intelligent Computer Communication and Processing*, 2009.

[11] J. Zhao, J. Zhang, and B. Zhu, "Development and verification of the tire/road friction estimation algorithm for antilock braking system," *Mathematical Problems in Engineering*, vol. 2014, 2014.

[12] K. Han, Y. Hwang, E. Lee, and S. Choi, "Robust estimation of maximum tire-road friction coefficient considering road surface irregularity," *International Journal of Automotive Technology*, vol. 17, no. 3, pp. 415–425, jun 2016. [Online]. Available: http://link.springer.com/10.1007/s12239-016-0043-8

[13] X. Zhang and D. Göhlich, "A hierarchical estimator development for estimation of tire-road friction coefficient," *PLoS ONE*, vol. 12, no. 2, pp. 1–21, 2017.

[14] X. Xia, L. Xiong, K. Sun, and Z. P. Yu, "Estimation of maximum road friction coefficient based on Lyapunov method," *International Journal of Automotive Technology*, vol. 17, no. 6, pp. 991–1002, dec 2016. [Online]. Available: http://link.springer.com/10.1007/s12239-016-0097-7

[15] M. Wielitzka, M. Dagen, and T. Ortmaier, "Sensitivity-based Road Friction Estimation in Vehicle Dynamics using the Unscented Kalman Filter," *Proceedings of the American Control Conference*, vol. 2018-June, pp. 2593–2598, 2018.

[16] Y. Wu, F. Liu, L. Guan, and X. Yang, "A survey of vision-based road parameter estimating methods," in *Proc. Intelligent Computing Methodologies ICIC*, 2020.

[17] E. J. Almazan, Y. Qian, , and J. H. Elder, "Road segmentation for classification of road weather conditions," in *Proc. ECCV Workshops*, 2016.

[18] J. Zhao, H. Wu, and L. Chen, "Road surface state recognition based on SVM optimization and image segmentation processing," *Journal of Advanced Transportation*, 2017.

[19] A. Jonnarth, "Camera-based friction estimation with deep convolutional neural networks," Master's thesis, Uppsala University, Uppsala, Sweden, 2018.

[20] M. Nolte, N. Kister, and M. Maurer, "Assessment of deep convolutional neural networks for road surface classification," in *Proc. International Conference on Intelligent Transportation Systems (ITSC)*, 2018.

[21] G. Pan, L. Fu, R. Yu, and M. I. Muresan, "Winter road surface condition recognition using a pre-trained deep convolutional neural network," in *Transportation Research Board 97th Annual Meeting*, 2018.

[22] S. Roychowdhury, M. Zhao, A. Wallin, N. Ohlsson, and M. Jonasson, "Machine learning models for road surface and friction estimation using front-camera images," in *International Joint Conference on Neural Networks (IJCNN)*, 2018.

[23] E. Sabanovic, V. Zuraulis, O. Prentkovskis, and V. Skrickij, "Identification of road-surface type using deep neural networks for friction coefficient estimation," *Sensors*, vol. 20, no. 3, 2020.

[24] D. Jin, B. Leng, X. Yang, L. Xiong, and Z. Yu, "Road friction estimation method based on fusion of machine vision and vehicle dynamics," in *Proc. IEEE Intelligent Vehicles Symposium (IV)*, 2020.

[25] E. Svensson, "Transfer learning for friction estimation using deep reduced features," Master's thesis, Department of Electrical Engineering, Linköping University, Linköping, Sweden, 2020.

[26] N. dynamics, "The grip indicator," 2021. [Online]. Available: https://niradynamics.se/tire-grip-indicator/

[27] J. Cech, T. Hanis, A. Konopisky, T. Rurtle, J. Svancar, and T. Twardzik, "Self-supervised learning of camera-based drivable surface roughness," in *Proc. IEEE Intelligent Vehicles Symposium (IV)*, 2021, in Review.

[28] H. Pacejka, *Tyre and Vehicle Dynamics*. Elsevier LTD, Oxford, 2002.

[29] D. Schramm, M. Hiller, and R. Bardini, *Vehicle Dynamics*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014. [Online]. Available: http://link.springer.com/10.1007/978-3-540-36045-2

[30] Smart Driving Solutions Research Center, "Twin Track vehicle model," 2020. [Online]. Available: https://github.com/SDS-RC-FEE-CTU-in-Prague/TwinTrack

[31] D. Efremov, T. Hanis, and M. Hromcik, "Introduction of Driving Envelope and Full-Time-Full-Authority Control for Vehicle Stabilization Systems," in *2019 22nd International Conference on Process Control (PC19)*. IEEE, jun 2019, pp. 173–178.

[32] Adams/Tire. Using the PAC2002Tire Model. [Online]. Available: https://docplayer.net/54206752-Using-the-pac2002tire-model.html

[33] K. Maeda, H. Fujimoto, and Y. Hori, "Four-wheel driving-force distribution method based on driving stiffness and slip ratio estimation for electric vehicle with in-wheel motors," in *2012 IEEE Vehicle Power and Propulsion Conference*. IEEE, oct 2012, pp. 1286–1291. [Online]. Available: http://ieeexplore.ieee.org/document/6422490/

[34] A. H. Jazwinski, *Stochastic processes and filtering theory*. Courier Corporation, 2007.

[35] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. CVPR*, 2016.

[36] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization." in *Proc. ICLR*, 2015.