CHALMERS

Vehicle Road Grip Estimation using a Front-facing Camera and Temperature Readings

Master's thesis in Electrical engineering

ANDREAS GELOTTE

Department of Electrical Engineering
Division of Systems and Control
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2023

MASTER'S THESIS IN ELECTRICAL ENGINEERING

Vehicle Road Grip Estimation using a Front-facing Camera and Temperature Readings

ANDREAS GELOTTE

Department of Electrical Engineering
Division of Systems and Control
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2023

Vehicle Road Grip Estimation using a Front-facing Camera and Temperature Readings
ANDREAS GELOTTE

Supervisor: Magnus Gustafsson, Attribute Leader Brake, China Euro Vehicle Technology AB
Hasith Karunasekera, Postdoc, Chalmers University of Technology
Examiner: Jonas Sjöberg, Full Professor, Chalmers University of Technology

Master's thesis
Department of Electrical Engineering
Division of Systems and Control
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden
Telephone: +46 (0)31-772 1000

Cover:
Car wheel driving on snow-covered asphalt, as envisioned by the
Stable Diffusion deep learning text-to-image model.

Chalmers Reproservice
Gothenburg, Sweden 2023

Vehicle Road Grip Estimation using a Front-facing Camera and Temperature Readings
Master's thesis in Electrical engineering
ANDREAS GELOTTE
Department of Electrical Engineering
Division of Systems and Control
Chalmers University of Technology

Abstract

An approach to estimating road friction and corresponding passenger vehicle braking and cornering performance is presented. In particular, it classifies dry and wet asphalt, gravel, snow- and ice-covered roads using an onboard front-facing camera. To estimate road surface condition, the convolutional neural network SqueezeNet is retrained using 4100 images of the mentioned road conditions. All five classes are associated with individual tyre slip-friction relations as per Pacejka's Magic Formula. A final estimation of surface condition is given by an arbiter, taking input from convolutional neural networks, ambient temperature and road surface type databases for the current GPS coordinates. Models for minimum braking distance and safe cornering speed are proposed using point mass modelling for the current road friction estimate.

The model is verified in vehicle on automotive proving grounds and public roads, evaluating road surface condition classification performance in steady-state driving and road friction estimation during dynamic manoeuvres. Predicted braking distances are within 22% of experimental data, and safe cornering speeds within 3%. Surface type is correctly identified in most live vehicle tests, while classification performance reach 78% for pre-recorded images of roads.

Keywords: Road grip, Road Friction Estimation, RFE, Tyre, Computer vision

Acknowledgements

I would like to express my thanks to supervisor Magnus Gustafsson at CEVT (China Euro Vehicle Technology AB) for his support and for giving me the opportunity to write my thesis at such an exciting company. He showed the ins and outs of vehicle dynamics development, proving with the daily work how diverse and rewarding automotive testing can be.

A special thanks go to academic supervisor Hasith Karunasekera for his guidance and useful feedback on the subject of computer vision, complementing my knowledge gaps as a mechanical engineering student. I also wish to thank my examiner Jonas Sjöberg for his input and great patience with the thesis writing process.

Furthermore, I thank Martin Pedersen, Erik Nordin and Felix Collin for being excellent deskmates and inspiring colleagues, providing both technical insights and words of encouragement.

Contents

List of Figures

List of Tables

Nomenclature

Symbol	Unit	Description
$\boldsymbol{A}$	$-$	State matrix for the point mass model
g	m/s^2	Acceleration due to gravity
F	N	Resultant force magnitude for the point mass model
F_x, F_y	N	Longitudinal and lateral force at tyre
F_z	N	Normal force at tyre
m	kg	Vehicle mass
$\boldsymbol{q}$	$-$	Input function for the point mass model
s_x	$-$	Longitudinal slip
t	s	Time
v	m/s	Absolute velocity
v_x, v_y	m/s	Longitudinal and lateral velocity
w	m	Track width
α	$degree$	Slip angle of tyre
$\boldsymbol{\gamma}$	$-$	Vehicle control input for the point mass model
δ	$degree$	Steering wheel angle
μ	$-$	Coefficient of friction
ϕ	rad	Angle of the resultant force for the point mass model
ω_w	m/s	Wheel circumferential velocity

Acronyms

ABS	Anti-lock braking system
ADAS	Advanced driver-assistance systems
ANN	Artificial neural network
API	Application programming interface
CAN	Controller area network
CEVT	China Euro Vehicle Technology AB
CNN	Convolutional neural network
COG	Centre of gravity
DNN	Deep neural network
DOF	Degrees of freedom
ECU	Electronic control unit
ESC	Electronic Stability Control
GIS	Geographic information system
GPS	Global positioning system
ML	Machine learning
NN	Neural network
OBD	On-board diagnostics
RFE	Road friction estimation
ROI	Region of interest
VTI	The Swedish National Road and Transport Research Institute

1 Introduction

1.1 Background

A key component of vehicle safety is road friction, which dictate the available friction for the tyres and thus vehicle stability. While driving, a lack of friction can manifest as a feeling of imprecise steering and unexpected sliding during cornering and braking, leading to low driver confidence and safety margins.

Wallman and Åström [1] of The Swedish National Road and Transport Research Institute concluded that while road friction is important for traffic safety, it is not the only contributing factor to accident risk. Human drivers adjust their behaviour depending on the apparent road environment, weather and tyre noise. These attributes are usually not measured and processed by the vehicle on-board computers, leading to a gap in available information between human drivers and autonomous driving solutions.

Adverse road conditions due to heavy rain, snow or ice greatly affects vehicle handling, and are currently combated with reactive solutions like anti-lock braking system (ABS) and electronic stability control (ESC) that intervene once the wheels slip.

By combining a front-facing camera and image classification algorithms, it is possible to estimate road condition ahead of the vehicle and by extension the minimum braking distance, advanced driver-assistance systems such as lane keeping assist and automatic emergency braking will be able to adjust the cruising speed and brake actuation aggressiveness to preemptively avoid a loss of vehicle control.

This study validates road condition classification performance for a camera-based system in real-world scenarios. Classes of road condition include dry and wet asphalt, gravel, snow and ice.

1.2 Related work

In 2011, Tang and Breckon [2] performed road environment classification with a *k-nearest neighbors algorithm* (k-NN) applied to color and texture features extracted from vehicle dashcam images using the GLCM method and Gabor filters. The study managed to differentiate environments such as off-road, highway and urban when posed as binary classification problems. Additionally, they employed a simple artificial neural network (ANN) with two layers and managed a 80-85% success rate in a four-class determination problem, showing a substantial improvement over the k-NN method. Tang and Breckon discussed potential driver assistance applications and autonomous driving. While not related to road grip or used a neural network for classification, it is of interest as it is one of the earliest studies in the field.

Taylor et al. [3] classified road type in 2012 by analysing vehicle on-board CAN network data with decision tree methods, achieving up to 89% accuracy for classification of road type, signage and carriageway type. Investigated signals included wheel speed, gear position, and suspension movement. Further, Tumen et al [4] determined road type and quality by differentiating pothole-riddled roads from smooth roads using camera images and a novel *convolutional neural network* (CNN) in 2018. They classified six different types of dry roads with an accuracy of $91,41\%$, and binary classification of pothole-defect roads to $91,07\%$ accuracy. In the same year, Šabanovič et al. [5] developed a *deep neural network* (DNN) classification algorithm to improve anti-lock braking system (ABS) performance during hard braking on various road surfaces, with a road-surface type classification accuracy of 96% and road condition classification of 92%. Stopping distance improved a remarkable 18%, although it should be noted that current ABS technology has improved over the second generation Toyota Prius (2003-2009) used in the study. Šabanovič retrained the SqueezeNet [6] neural network.

Volvo Cars [7] researched and patented a method for semi-autonomous image segmentation and annotation, in which the ego lane in the dashcam image is transformed into bird's-eye view and divided into a grid of sub-images. Yet again, the SqueezeNet neural network was employed to identify road surface condition from the full dashcam image, subsequently the sub-images in the bird's-eye view were subject to a rule-based road friction estimation.

Student thesis work has been performed by Jonnarth [8] at Uppsala University with the DenseNet [9] densely connected convolutional network analysing 37,000 images collected on-road, differentiating between dry and snow-covered roads. Jonnarth's thesis explored the feasibility of extending the training dataset with simulated images from a racing game, seeing only a marginally increase in prediction accuracy by 0.4% over the original datasets' 84.14%.

RoadEye, Sweden [10] developed in 2007 an optical system that illuminates the road surface with two separate wave lengths of light and measures the difference in absorption for each reflection. Ice absorbs the most light at 1550 nm while water absorbs the most at 1450 nm, allowing RoadEye to calculate the composition of ice and water on the measured surface. Subsequent derivatives include OmniKlima, Sweden [11] and RoadCloud Oy, Finland [12]. Similar research has been done by VTT Technical Research Centre of Finland [13].

Minges student thesis [14] investigated the possibility of estimating road friction with a neural network using weather information from SMHI, the Swedish Meteorological and Hydrological Institute, and Trafikverket. Although the neural network performed poorly in the classification tasks, the statistical analysis is of interest to this study. Minges found Pearson correlation coefficients, ranging from moderate (0.26 - 0.2 for peak wind speed, dew point and temperature), to weak negative (-0.021 for air humidity) correlation with friction.

There have been several projects with the aim of mapping road friction on public roads using a network of connected vehicles, with the largest national study of its kind being Digital vinter [15] by Trafikverket, the Swedish Transport Administration. Private initiatives include industry suppliers such as NIRA Dynamics, Sweden [16], RoadCloud Oy, Finland. [12], Continental AG Preview ESC, Germany. [17], and Tactile Mobility, Israel. [18]. Most solutions use virtual sensors based on on-board diagnostic signals (OBD) such as wheel torque, rotational speed, gyroscopic data and GPS speed. These systems are reactive rather than predictive, a clear disadvantage to the camera-based systems that estimate the road surface ahead of the vehicle.

1.3 Contributions

This study provides an estimation of the utilised road friction coefficient, maximum available friction coefficient and minimum braking distance in real-time for a passenger vehicle driven on an automotive proving ground. The estimator gather images of the road from a front-facing camera, ambient temperature and wheel speed from the vehicle's CAN network, and road pavement data from two national databases based on GPS coordinates. Next, retrained SqueezeNet neural networks classifies the image feed for five types of road surface condition ("Gravel", "Dry", "Wet", "Snow" and "Ice") and the presence of snow in the road shoulder and surrounding areas.

Road condition is arbitrated based on these input in a switch statement control mechanism to mitigate misclassifications by the neural networks. Road friction coefficient is given by Pacejka's Magic Formula for the wheel slip and current surface type, with distinct friction-wheel slip relations for each road surface condition.

This work is limited in scope by excluding radar and lidar data, along with online live data such as weather forecast and road accident information. Low light conditions are not considered, and neither are split friction scenarios, i.e. road grip difference between the left and right wheel.

1.4 Thesis outline

In Chapter 1, the context of the study is established through literature review of previous work and contributions are given. Chapter 2 provides theoretical background on vehicle road grip and image analysis. Chapter 3 outlines the tools, procedures and materials used in the study.

Chapter 4 - 5 presents the results from public road and automotive proving ground tests. Chapter 6 reflects on the methodology and suggests potential areas for future work. Lastly, Chapter 7 concludes the thesis.

2 Theory of visual road grip estimation

This chapter gives the foundation for visually estimating road grip, with a description of the Computer vision field and methods of vehicle modelling, explaining why road surface conditions affect a vehicle's handling ability and how vehicle performance is calculated.

2.1 Vehicle road grip and tyre modelling

Road grip is the traction force between a vehicle's tyres and the road surface, and is what allows a vehicle to accelerate, turn and brake. It is composed of dry friction, the force that resist the relative motion between two surfaces that are in contact:

$$F = \mu F_z, \tag{2.1.1}$$

where F_z is the normal force acting perpendicular to the surfaces and it is a direct result of the weight acting on the tyre. On level ground, normal force F_z equals the weight of an object of mass g:

$$F_z = mg. \tag{2.1.2}$$

While the coefficient of friction is often expressed as the ratio between the previously mentioned forces, the numerical values of the forces are not easily measured in practical applications.

In real-time vehicle modelling, this coefficient is often derived from empirical tyre models such as the Pacejka Magic Formula [19], consisting of a curve of tyre slip versus coefficient of friction fitted to experimental data. Longitudinal tyre slip s_x, as defined by the SAE670e standard, is the relative difference between a driven wheel's circumferential velocity ω with tyre radius r_e, and the vehicle's absolute velocity v:

$$s_x = \frac{\omega r_e - v}{\max(v, \omega r_e)}. \tag{2.1.3}$$

Tyres provide peak traction with a certain amount of slip occurring, as illustrated in Figure 2.1.1 for a few road surfaces identified in [20].

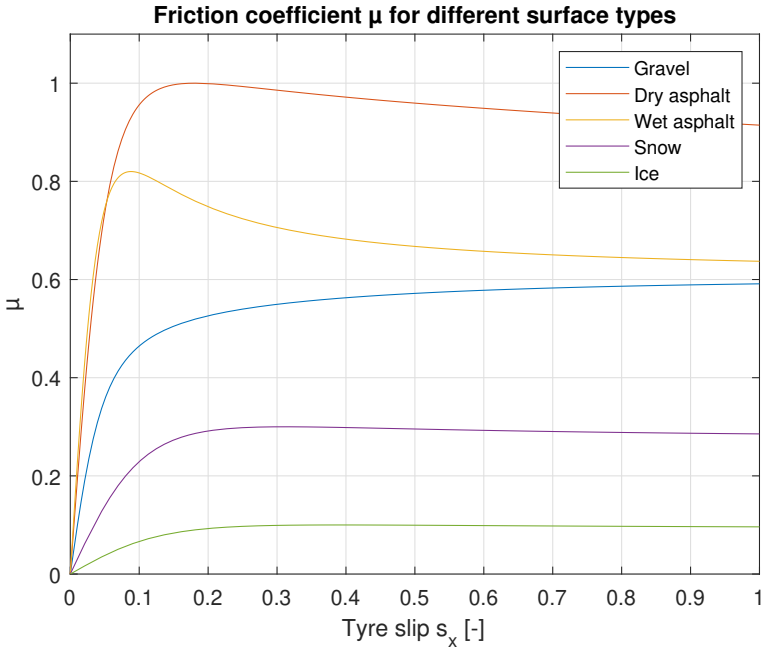


Figure 2.1.1: *Road friction as a function of tyre slip.*

The difference in curve shapes found in the figure above, i.e. an initial peak in friction for paved surfaces and the gradual increase for looser surfaces, have their cause in material properties and texture of the road

surface. On a micro scale, the surface materials adhere to each other through molecular adhesion acting in the opposite direction of travel at the contact patch. On a macro scale, the surface roughness of the road causes indentations in the rubber that accounts for 80% of tyre grip [21]. Due to the elastic hysteresis of rubber, more force is required during the compression of material than the rebound, causing an asymmetrical reactive force contributing to tyre grip. This energy loss is responsible for the rolling resistance and constitutes approximately 5% of the energy required to propel a passenger vehicle.

The aforementioned peak in friction coefficient for paved surfaces, as shown by yellow and red lines in Figure 2.1.1, is explained by the transition from a large static friction to a weaker sliding friction, both of which are subdivisions of the dry friction in Eq. 2.1.1. Looser road surfaces such as dirt, gravel and snow exhibit less static friction than paved surfaces due to shearing deformation of the surface, while also being subject to material build-up due to the plowing effect of the wheel.

While there is no international standard for road friction measurements and threshold values, The Swedish National Road and Transport Research Institute (VTI) have performed extensive research on road friction coefficients on public roads using the systems by SARSYS-ASFT [22], which spray a small amount of water on the road surface in front of a separate measurement wheel with 15% slip.

Upon surveying 102 km of public roads, VTI [23] found a mean friction of 0.8 which could be considered high friction, with 13% of the roads having 0.7 or worse friction. An upper threshold for poor road friction could be 0.25, as Trafikverket has established this as the minimum legal requirement on unplowed, snow-covered roads [24].

2.2 Vehicle dynamics

To accurately predict the outcome of vehicle manoeuvres, a mathematical model is needed to describe the current and future state in terms of position, velocity and acceleration. This section presents common approaches to vehicle modelling, including a low-fidelity point mass model and the single track model, alias bicycle model, for lateral and longitudinal movement. Following that, vertical movement and load transfer are outlined. All expressions are given in local XY-coordinate systems as defined in [25], with a right-handed axis system with origin commonly located at the sprung mass centre of gravity (COG).

2.2.1 Point mass model

The simplest model of vehicle dynamics offers two degrees of freedom (DOF), representing the vehicle as a single point mass moving in a plane. It is commonly used in studies of vehicle rigid body kinematics and energy efficiency, but is not sufficient for transient behaviours found in dynamic driving. If restricted to a singular tractive force $-\mu mg \leq F_x \leq \mu mg$, the state-space representation of a point mass model is expressed as

$$f\left(\mathbf{x}, \boldsymbol{\gamma}, t\right) = \mathbf{A}\mathbf{x} + \mathbf{q}(\boldsymbol{\gamma}) = \begin{bmatrix} v_x & v_y & a_x & a_y \end{bmatrix}^{\mathsf{T}} \tag{2.2.1}$$

evaluated for vehicle state vector $\boldsymbol{x}$ in global coordinates, vehicle control input vector $\boldsymbol{\gamma}$ comprising tractive force F and steering angle ϕ, and time t:

$$\mathbf{x} = \begin{bmatrix} X \\ Y \\ \dot{X} \\ \dot{Y} \end{bmatrix}, \tag{2.2.2}$$

$$\boldsymbol{\gamma} = \begin{bmatrix} F \\ \phi \end{bmatrix}, \tag{2.2.3}$$

with state matrix $\boldsymbol{A}$ and input function $\boldsymbol{q}$

$$A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \tag{2.2.4}$$

$$\mathbf{q}(\boldsymbol{\gamma}) = \frac{F}{m} \begin{bmatrix} 0 \\ 0 \\ \cos\phi \\ \sin\phi \end{bmatrix}, \tag{2.2.5}$$

where F is the resultant force magnitude with components F_x and F_y in local coordinates:

$$\begin{bmatrix} F_x \\ F_y \end{bmatrix} = \begin{bmatrix} 0 & 0 & \cos v & \sin v \\ 0 & 0 & -\sin v & \cos v \end{bmatrix} \mathbf{q}(\boldsymbol{\gamma}) \ , \ \tan v = \frac{\dot{Y}}{\dot{X}}. \tag{2.2.6}$$

2.2.2 Single track model

The single track model shown in Figure 2.2.1 introduce two wheels mounted longitudinally with a known wheelbase L, allowing Z-axis yaw rotation around the centre of gravity and thus 3 DOFs. Slip angles α_f and α_r are individual for each wheel, enabling the under and oversteer phenomena commonly found in front-wheel drive and rear-wheel drive vehicles respectively.

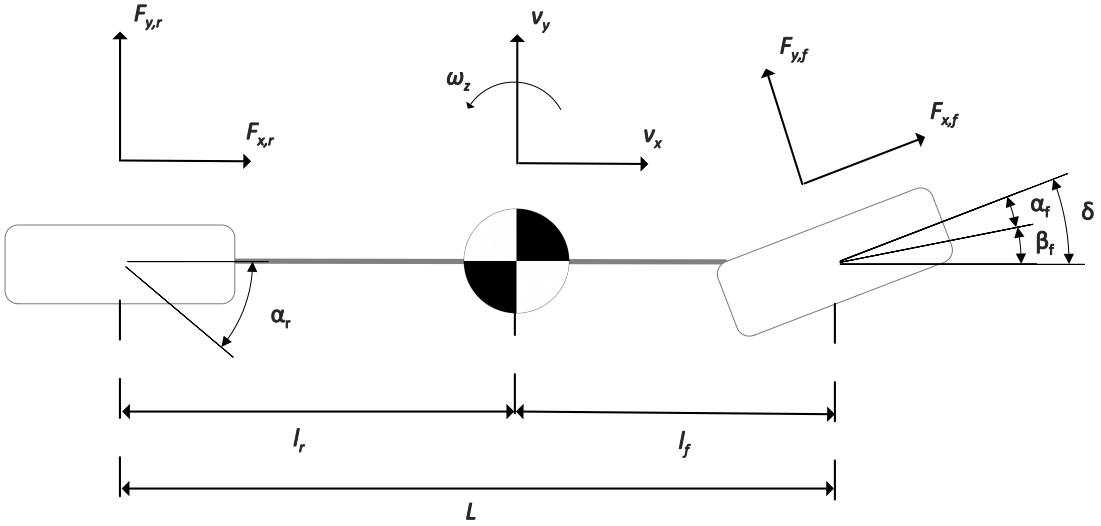


Figure 2.2.1: *Single track model.*

A state-space representation of the single track model is not presented in this chapter, but can be constructed from force and moment equilibria, rigid body compatibility between the COG and axles, and an appropriate tyre model. In order of complexity, tyre modelling range from linear with constant slip stiffness to Pacejka's model as previously outlined, and advanced brush models. Vehicle control input for the state-space model commonly consist of lateral speed and steering angle.

2.2.3 Lateral and longitudinal load transfer

Phenomena like pitch, roll and heave require more comprehensive vehicle dynamics models. Transient longitudinal driving such as braking and acceleration may be analysed with the model with axle pivot points defined in [26], illustrated in Figure 2.2.2. Longitudinal wheel forces transfer through trivial front and rear suspension linkages with virtual pivot points, while suspension damping help capture transient dynamics in short time frames.

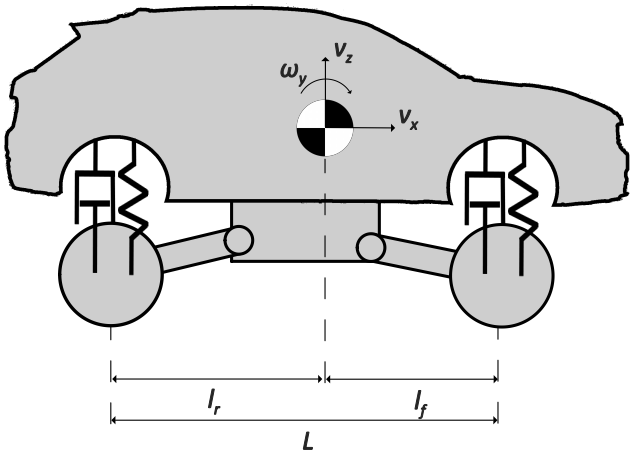


Figure 2.2.2: *Model with axle pivot points for pitch and heave analysis.*

Load transfer during cornering is commonly modelled with a two-track model, consisting of two single track models constrained geometrically with a fixed track width w, shown in Figure 2.2.3 below. The two-track model enables roll by introducing a height for the centre of gravity, roll angle and roll moment of inertia with a known roll damping and stiffness, therefore resulting in 4 DOFs. Vehicle yaw due to braking and acceleration is implemented similarly.

At lower speeds, accurate modelling demand Ackermann steering geometry, where the left and right front wheel steer with different angles so that the respective axial centreline align at a common centre point M. This accommodates the difference in radii for the left and right wheels during cornering.

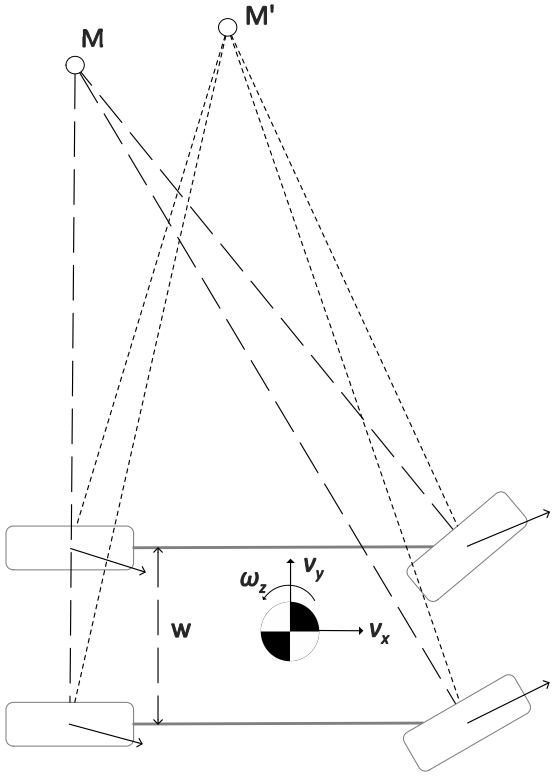


Figure 2.2.3: *100% Ackermann steering geometry and the instantanious turn center for a two-track model.*

As can be seen in Figure 2.2.3, the axial centreline of all wheel intersect at M. It should be noted that this is a geometrically sound solution for a wheel that experience zero slip angle and is not representative of real-life applications. When slip angle is considered, the vehicle will turn around the instantaneous turn centre M' positioned ahead of M due to the slip angles in the rear tyres. Tyre rubber properties, driving style and cornering radius all play an important role when deciding on a steering geometry. In certain scenarios, race cars utilise anti-Ackermann steering to optimise slip angle for the outermost tyre by increasing toe-in [27], as this might allow for increased lateral forces.

2.3 Road surface condition recognition using computer vision

Computer vision is a field within artificial intelligence allowing computers to perceive objects in images and video, tracing its roots to the 1960s and neuropsychological experiment on object recognition [28] and the emerging academic discipline of artificial intelligence. By detecting edges, features and objects, computer vision is a good candidate for visually determining weather condition in an automotive setting.

2.3.1 Artificial neural networks

The computer equivalent of biological nerve cells, with axons and synapses, are *nodes*, also called units, connected by links. Artificial neural networks (ANNs) represent these nodes with simple mathematical operations called activation functions, where the output of a node propagate through the links to the input of another node. Input function in_j for input a_i and associated weights $w_{i,j}$ is defined as

$$in_j = \sum_{i=0}^{n} w_{i,j} a_i. \tag{2.3.1}$$

Output a_j is derived from the activation function g applied to the sum of the inputs:

$$a_j = g(in_j) = g\left(\sum_{i=0}^{n} w_{i,j} a_i\right). \tag{2.3.2}$$

By connecting several nodes to each other, a neural network is formed. One example is the multi-layer neural network in Figure 2.3.1 below, where each unit coloured in red, blue and green represent a neuron. The arrows represent input- and output links between layers. In the case of image classification networks, the input layer has an input size equal to the number of pixels times colour channels in the image.

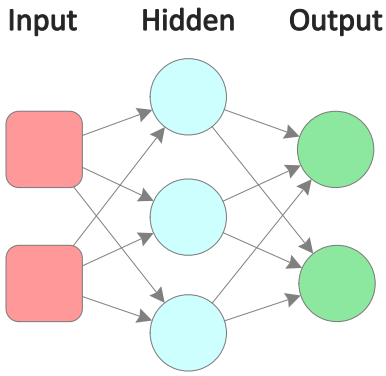


Figure 2.3.1: *An artificial neural network with two inputs, one hidden layer with three neurons, and two outputs.*

The arrangement of these connections decide the type of network and affects functionality greatly; as either a feed-forward or a recurrent network [29]. The former connects all neurons in a single direction, letting information flow from the input image to the network output, while the latter feeds its output to the nodes in earlier layers. In short, this allows recurrent networks to hold an internal state that can be subject to oscillations and chaotic behaviour, and can be expressed as Long short-term memory (LSTM). This is particularly useful when analysing sequences of data in fields such as speech recognition, anomaly detection and stock market forecasts. Feed-forward networks can be trained with backpropagation, where the gradient of a loss function is computed efficiently enough to use gradient optimisation methods for adjusting the weights of each node, allowing a learning process to take place.

Activation functions introduce complexity as they are often either binary or logistic and thus non-linear [30]. This lets a network of connected neurons represent nonlinear functions, as a linear network could just as well be represented by a single layer. Linear layers do not allow backpropagation, as the derivative of the activation function is constant and unaffected by the input. The lack of backpropagation prohibits learning in the network, making linear activation functions unsuitable for neural networks.

An important aspect of non-linear activation functions is the derivative and its influence on the learning process. Small derivatives, or gradients, with a distribution below one tend to decrease as they get multiplied through the layers of a neural network, resulting in slow backpropagation during learning. This is known as the vanishing gradient problem. On a similar note, aggressive derivatives and large weights may result in an exploding gradient problem, where the network is unstable during training and oscillates between weights, never converging on a solution.

Some of the most commonly used activation functions are described in the following section, with the sigmoid function on the interval (0,1) used especially for probability output:

$$g(z) = \frac{1}{1 + e^{-z}}. \tag{2.3.3}$$

As the sigmoid function is smooth and differentiable, derivatives are non-negative. One drawback is that large absolute input values results in small derivatives, namely the aforementioned vanishing gradients leading to slow or no progress during training. For an input with mean distribution 0, sigmoid function have a mean output value of 0.5. The output magnitude of a layer is therefore dependant on the number of neurons. The linear and sigmoid activation functions are compared in Figure 2.3.2.

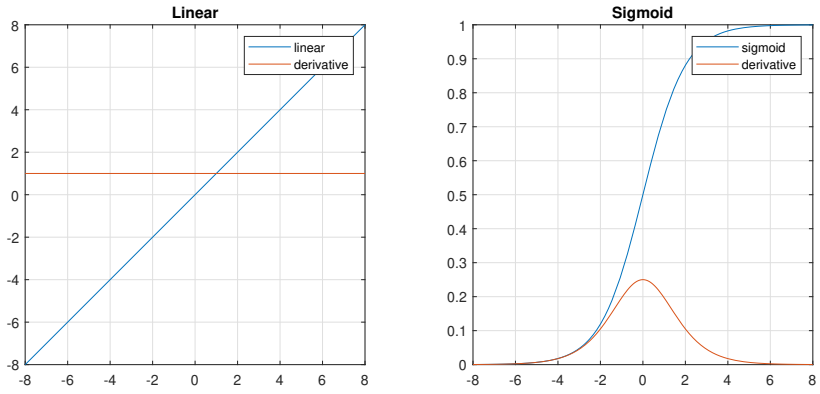


Figure 2.3.2: *Linear and sigmoid activation functions with respective derivative.*

Similarly, the softmax function is commonly used for the output layer and normalise input into a probability distribution:

$$g_i(\vec{z}) = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}. \tag{2.3.4}$$

One clear advantage of the softmax function is that the sum of the results from a layer always equals one, avoiding the problematic magnitude of the summation found with sigmoid functions.

The hyperbolic tangent activation function (tanh) is a scaled and shifted version of the sigmoid function, ranging (-1,1) with negative inputs mapping to negative output:

$$g(z) = \frac{e^x - e^{-x}}{e^x + e^{-x}}. \tag{2.3.5}$$

A computationally efficient activation function is the rectified Linear Unit (ReLU), and functions the same way as an electrical rectifier that limits output to zero for negative input values. As the derivative is constant for positive input values, ReLU avoids saturation in these region.

$$g(z) = \max(0, z). \tag{2.3.6}$$

Leaky ReLU is a variant that addresses the zero gradient value for negative input, improving the backpropagation process and preventing the creation of dead neurons which never activate:

$$g(z) = \max(az, z). \tag{2.3.7}$$

A comparison between the hyperbolic tangent and ReLU activation functions is made in Figure 2.3.3 below. Notable is the gradient for input values of zero.

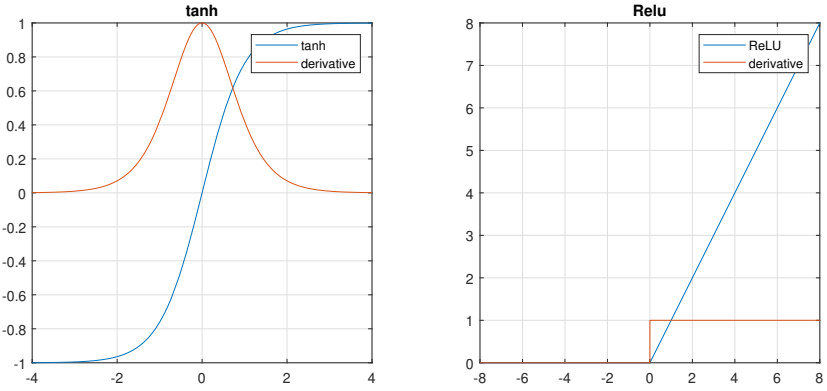


Figure 2.3.3: *Tanh and ReLU activation functions with respective derivative.*

2.3.2 Convolutional neural networks

As Section 1.2 suggest, convolutional neural networks are a highly topical subject as they provide great runtime performance and generalisation, meaning that they are suitable for the ever-changing environments a vehicle passes through. In mathematics, convolution is the operation on two functions f and g, where the output integral h express the amount of overlap as g is shifted over f. In two-dimensional applications such as image processing with pixels x and y, the convolution h is expressed as

$$h(x,y) = (f * g)(x,y) = \sum_{u=-\infty}^{+\infty} \sum_{v=-\infty}^{+\infty} f(u,v)g(x-u, y-v). \tag{2.3.8}$$

Convolutional neural networks (CNN) applies this operation with filters in hidden layers, where the filters have learnable parameters. These parameters, or function weights as described in Chapter 2.3.1, are shared across all filters within a layer. That way, features learned by neurons in one section of the image can be recognised in another. Further, CNN:s employ local connectivity between neurons in adjoining layers, resulting in a strong response to local features within a region of the image. This allows stacked layers to recognise texture and objects from several smaller features.

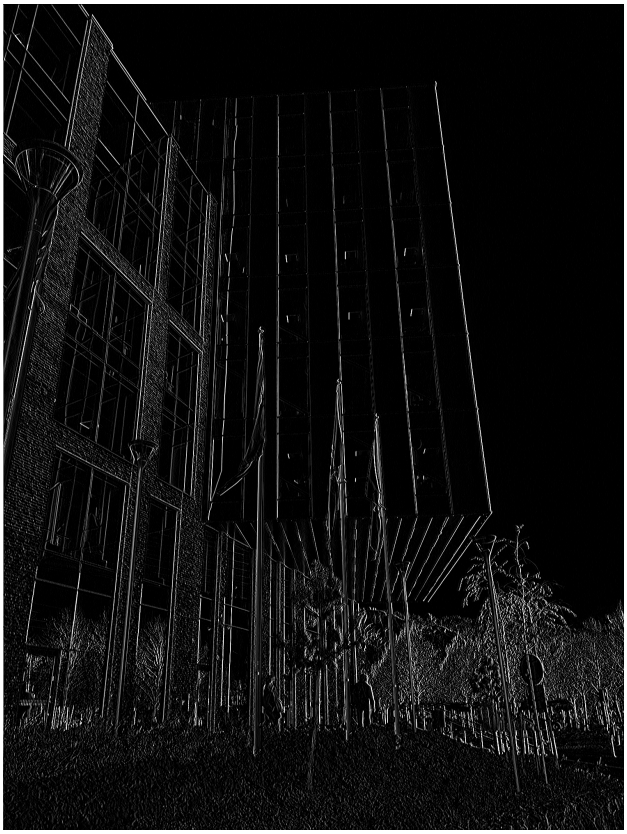
In computer vision, convolution is the operation of iteratively applying a kernel to a patch of pixels sliding over the image, where image kernels are small matrices that operate on the pixel values to modify the output image. Notably, the derivative $(f * g)'$ equals the convolution with the derivative $f * (g')$. This saves computational time by reducing two operations into one, for example when combining edge detection with smoothing to reduce the number of visible edges. A few image kernels are exemplified in Table 2.3.1.

Operation	**Kernel**
Sharpen	$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$
Box blur	$\frac{1}{9}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$
Outline	$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$

Table 2.3.1: Matrices of common image kernels.

Convolution is illustrated in Figure 2.3.4 with a kernel that identifies vertical edges in a black and white image, indicating vertical edges with bright pixels.

(a) *Black and white*



(b) *Filtered*

Figure 2.3.4: *Image filtered with a vertical edge detection kernel.*

Digital images are generally represented by a pixel grid where the colour of each pixel is defined by the additive primary colours red, green and blue. 24-bit true colour images use 8 bits for each of the primary colours, allowing $0 - 255$ distinct values for a combined colour space of $2^{24} \approx 16.8$ million colours. Grayscale pixel values range from 0 for black to 255 for white, exemplified in Figure 2.3.5 portraying a colour transition along a vertical edge:

128	128	192	224	224
128	128	192	224	224
128	128	192	224	224
128	128	192	224	224
128	128	192	224	224

Figure 2.3.5: *Pixel values of a vertical edge image.*

The 3×3 kernel, marked in green in the figure below, impose a pixel-wise multiplication of which the sum is given to the centre pixel:

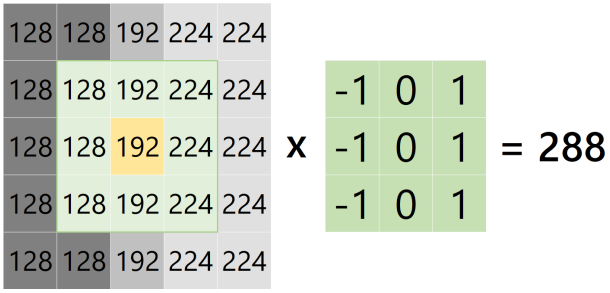


Figure 2.3.6: *Convolutional operation with the kernel in green. Output value applies to the cell marked in yellow, fully saturating a 255 pixel value.*

By iterating the operation for every pixel in the image, all vertical edges are identified and displayed as a filtered image:

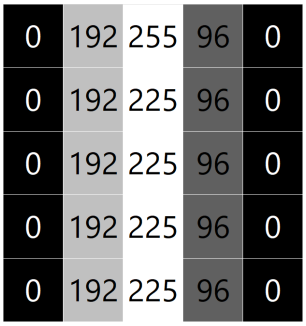


Figure 2.3.7: *Pixel values after filtering. Higher values indicate presence of an edge.*

In the above figures we assume all edge values extend to the surrounding pixels, facilitating the kernel operation. Other methods of edge handling include cropping, wrapping and mirroring edge pixels.

2.3.3 Network training and data set diversity

In an effort to increase transferability of models and save training time, transfer learning was introduced in the late 1970s. It is a method of applying previous knowledge on new problems, and is best exemplified in computer vision with neural networks that have been retrained to identify objects other than what was part of the original training data. This can be done by replacing the later network layers that combine known features such as patterns and edges into objects. In doing so, a well-known neural network trained on thousands of objects can form the basis for a network with the purpose to identify weather and road condition visually.

As touched upon in Chapter 2.3.1, the output of a loss function is minimised during training in machine learning. The function is a measure of how well the prediction model does in terms of predicting an expected outcome. Optimisation methods such as stochastic gradient descent with momentum (SGD) and root mean square propagation (RMSprop) are commonly used. To facilitate this, image datasets are split for training, test and validation with no overlapping images. This is an example of *supervised learning,* where the training data is labelled or sorted into categories corresponding to each output class. Other learning types are unsupervised learning and reinforcement learning. The former use unlabelled data that the neural network has to identify patterns independently, while the latter is feedback-based by rewarding correct predictions made during training.

During transfer learning, the model runs the risk of under- or overfitting to the training data. An underfit neural network does not account for enough points in the data set and lacks the complexity needed for the task, either by not learning fast enough, having too few layers or to little training data. The opposite problem, overfitting, occurs when the neural network models the training data too well and is characterised by substantially higher validation loss than training loss from the optimisation function. The model has learned to recognise the training data, rather than the common patterns for each class.

Dataset variation can be increased with image augmentation methods such as randomised geometric transformation like rotation, reflection and translation. To combat overfitting, regularisation methods such

as dropout layers are employed during training. These layers remove a portion of random neurons and links as shown in Figure 2.3.8, forcing the neural network to improve generalisation and consider additional input features.

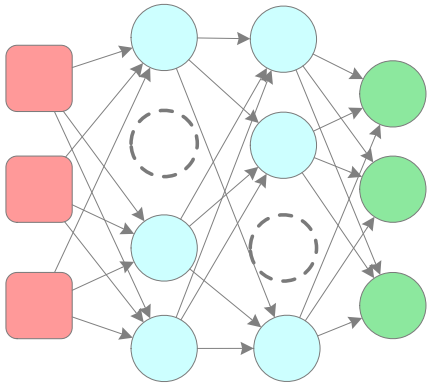


Figure 2.3.8: *Neural network with dropped neurons in dotted lines.*

The explicit parameters controlling the learning process are called hyperparameters and constitute an important aspect of neural network training. These parameters can be tuned with optimisation methods like parameter sweep, random selection, Bayesian and gradient-based optimisation. Examples of hyperparameters include learning rates, dropout layers, number of epochs for training duration and batch size before updating model parameters. Another design choice is the weight initialisation for the newly replaced layers, where the type of underlying activation function and initialisation method play a significant role in the time to convergence [31].

3 Grip estimator implementation and data collection

In this chapter, methods, tools and materials are presented. The study constitutes applied exploratory research as it seeks to develop a procedure in an under-researched field with potential application as a product in the automotive industry. Research data include primary and secondary sources as images were captured on-road and collected from an online database.

3.1 Tools and procedures

The grip estimator is developed in MathWorks Simulink, using sensor data gathered from a Lynk & Co 01 passenger vehicle driven on public roads and at automotive proving grounds. Image processing is performed with the Computer Vision, Image Acquisition, Image Processing and Deep learning toolboxes. Braking distances are measured with a Racelogic VBOX 3iSL data logger and a Vector VN7610 CAN FD interface reads messages from the chassis ECU in the vehicle, using the Vehicle Network Toolbox. Offline road surface data is read with the Mapping toolbox.

3.2 Materials and image sampling method

Approximately 2100 images were collected across a dozen occasions on public roads using a windshield mounted mobile phone. An additional 2000 images from the University of California, Berkeley BDD100K dataset [32] extended the image dataset. Additionally, offline geographic data of the wearing course, i.e. the upper layer of the roadway, was fetched from the National Road Database [33].

The collective 4100 images create two datasets: one with the driven lane transformed into bird's-eye view, and a separate for the full image including the surrounding lanes and road shoulder. This distinction is motivated in Section 3.3.2. The former dataset comprise five surface conditions and is used to train a multi-label classification network, and its image composition is shown in Table 3.2.1:

Source	**Dry asphalt**	**Wet asphalt**	**Gravel**	**Snow**	**Ice**
BDD100K [32]	537	589	0	96	0
Own driving tests	904	211	760	31	59
Total	1441	800	760	130	59

Table 3.2.1: Composition of the multi-label classification image dataset.

Images were labelled manually, with each category of surface condition exemplified in Figure 3.2.1. Category *(a) Dry* contains only images of dry tarmac with no loose gravel, with *(b) Wet* as the equivalent for roads ranging from slightly misty to puddle-covered. *(c) Gravel* include unpaved roads and tarmac with loose gravel, both wet and dry. *(d) Ice* images were captured on a handful occasions, on both an ice-covered lake and slushy stretches of road where braking tests were performed to confirm the layer of ice beneath. *(e) Snow* depicts completely or partially snow-covered roads with no ice underneath.

(a) *Dry* (b) *Wet* (c) *Gravel* (d) *Ice* (e) *Snow*

Figure 3.2.1: *Bird's-eye view of the five road surface condition categories.*

The dataset with full images, which includes surrounding lanes and road shoulders, enables training of a binary classification network with the image composition shown in Table 3.2.2:

Source	**Snow**	**Snowless**
BDD100K [32]	691	1372
Own driving tests	244	874
Total	935	2246

Table 3.2.2: Composition of the binary classification image dataset.

Images from the dozen self-collected occasions and numerous occasions in the BDD100K dataset were not separated for training and validation purposes. Instead, all datasets were split randomly into training and validation with a 70/30 distribution. Images were captured with 15 second intervals from a moving vehicle, eliminating the possibility of any identical images. Duplicate standstill images, i.e. from stops in intersections were removed. To prevent network bias towards more prevalent classes in the datasets, the imbalanced number of images per category are accounted for using inverse-frequency class weights:

$$w_i = \frac{N_{samples,tot}}{N_{classes} N_{samples,i}}. \tag{3.2.1}$$

3.3 Simulink implementation

This section details the grip estimator. First, the Simulink implementation is outlined. Second, the image pipeline and image classifiers are presented. Third, road property databases and vehicle sensors are discussed. Lastly, the arbiter that aggregate the data and decide on a final surface condition classification is presented.

3.3.1 Overview

An overview of the road grip estimator's working method is presented below in Figure 3.3.1 and is detailed in Appendix A. Video is captured by a front-facing camera onboard the vehicle and fed to neural networks, producing two types of classification: five classes of road surface condition and a binary indicator for presence of snow, further explained in Section 3.3.2. An arbiter gather these classifications and additional information from a road database and vehicle sensors, producing a final surface condition classification of dry or wet asphalt,

ice, snow or gravel. Finally, the surface condition dictate the friction coefficient between the road and tyre, resulting in numerical values for the minimum braking distance and maximum cornering speed for the current conditions.

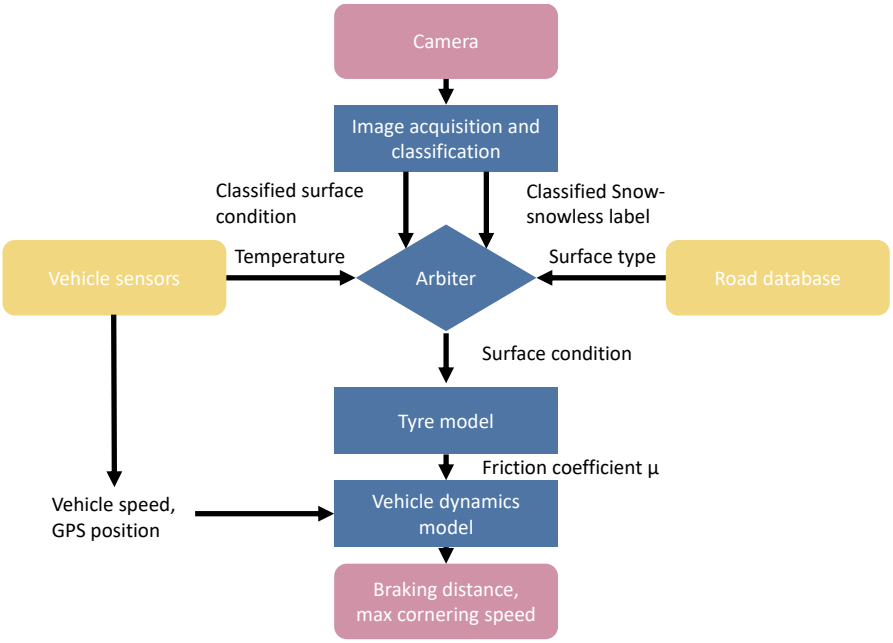


Figure 3.3.1: *Functional overview of the Simulink model, depicting the image pipeline and classification, database and vehicle sensor data, arbitration and lastly model output as braking distance and maximum cornering speed.*

3.3.2 Image acquisition and analysis

Image analysis forms the basis for the Simulink model, where a windshield-mounted camera constitutes the main source of information for classifying road condition. The image pipeline is visualised in Figure 3.3.2 and consists of a transformation of the camera image into bird's-eye view perspective, masking the predefined driving lane, resizing the image and classifying road condition with a convolutional neural network.

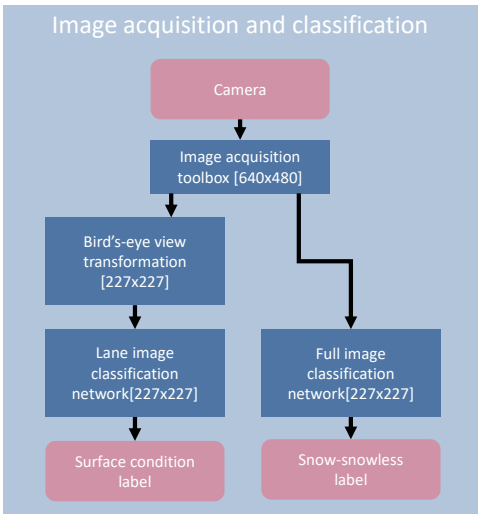


Figure 3.3.2: *Image pipeline and resolutions for the Simulink model.*

In parallel, the full image is fed to a second neural network that identifies the presence of snow indiscriminate of location; i.e. in the driven lane or road shoulder. This emulates one of the contextual clues human drivers identify when determining road condition. Both classification networks use retrained SqueezeNet convolutional

neural networks with the datasets presented in Chapter 3.2.

(a) *Full image*

(b) *Bird's-eye view of driven lane*

Figure 3.3.3: *Full image and bird's-eye view transformation in ambiguous conditions.*

The driven lane can be dynamically identified in both straight and curved sections of road by detecting lane markings and transforming the corresponding area into bird's-eye view perspective. This functionality was cancelled as camera type and placement in the BDD100K images are inconsistent, preventing automation of the neural network training procedure. An approximate region of interest is chosen in front of the vehicle, narrow enough to prevent interference from the vehicle bonnet, roof and A-pillars in the images from the BDD100K data set.

An accurate bird's-eye view transformation is achieved for images captured during the own driving tests. Here, a static region of interest (ROI) is chosen 5-15 meters ahead of the vehicle, four meters wide. Camera parameters are identified by capturing a checkerboard pattern from various angles and distances as demonstrated in Figure 3.3.4. The Matlab Computer Vision toolbox calculates intrinsic camera parameters such as focal length and field of view.

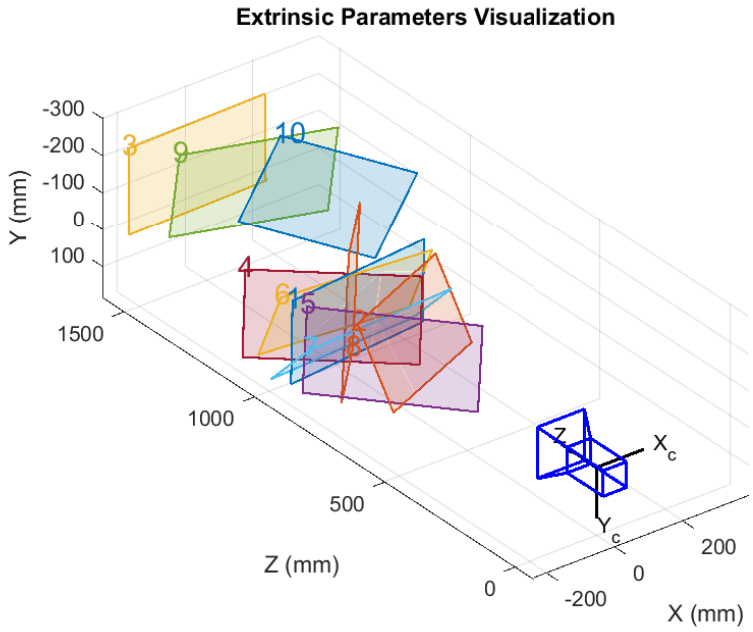


Figure 3.3.4: *Visualisation of the camera calibration process.*

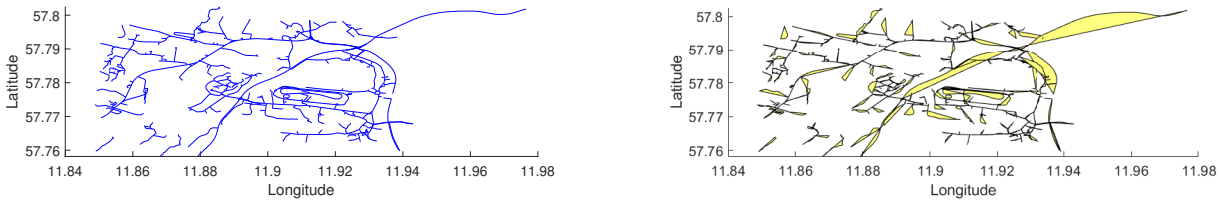
This process allows world coordinates to be translated to image pixel coordinates and vice versa, enabling precise measurements of distances to road features in the image. While not implemented in this study, these precise measurements could be used in conjunction with a grid of road regions extracted from the bird's-eye view projection. Each region is then classified separately, allowing road condition to be estimated granularly along the path of each wheel of the vehicle. As the distance to each region is known, safety functions such as ABS and ESC could adjust preemptively with great detail.

3.3.3 Databases and vehicle sensors

The Simulink model input is limited to a front-facing camera, on-board vehicle sensors and road pavement data. Two key databases were identified: the Swedish Transport Administration (Trafikverket) [34] and the National Road Database (NVDB) [33]. Both databases are implemented in the Simulink model, with Trafikverket cloud API calls performed with an extrinsic Python script, and NVDB data accessed locally.

Trafikverket's cloud API only provide information on wearing course for national roads, resulting in limited coverage. It does not provide any information on roads that lack dedicated numbering such as municipal side roads and city roads, contrary to NVDB which include all registered roads. While this information is accessible through the ADASIS map-enhanced driver assistance systems protocol [35] and can be communicated over CAN and TCP/IP, it was not implemented in the chosen vehicle.

The geospatial data is transformed from line to polygon geometry as portrayed in Figure 3.3.5 through the geographic information system software QGIS [36]. Implementation is suboptimal as it causes polygons from longer stretches of road to overlap with smaller side roads, resulting in ambiguous data in some cases.



(a) *Line geometry*

(b) *Polygon geometry*

Figure 3.3.5: *Comparison of line and polygon geospatial data depicting Säve locality, Gothenburg.*

The transformation accommodates the use of a Matlab File Exchange submission for determining corresponding values for positions in shape files [37]. It receives the current GPS coordinates and returns the appropriate road surface type if the position is located within a polygon.

Testing showed a significant performance difference between the two road data providers, depending on the chosen coverage as seen in Table 3.3.1. Response times scale with polygon count, concluding that the most feasible way of optimising performance is by limiting coverage.

Data provider	**Coverage**	**Response time [s]**	**File size [MB]**	**Polygon count**
Trafikverket	All national roads	0.16	online	-
NVDB	Västra Götaland county [1]	50	93	283,573
NVDB	Gothenburg municipality [1]	5	8.2	33,292
NVDB	Säve locality [1]	0.12	0.1	486

Table 3.3.1: Database response times and geographical coverage of road wearing course.

[1] Includes both national and municipal roads

3.3.4 Vehicle model

Initial tests revealed substantial computing power requirements for real-time application, as estimating stopping distances require stepping through the state-space model with potentially hundreds of calculations from the current vehicle speed to standstill. When performed at regular intervals, e.g. every second, heavy stutters are introduced to the image classification network. While the single-threaded nature of Simulink can be overcome with co-simulation components for parallel computing operations, this was not implemented.

The problem is especially apparent during cornering, as the single track model is only valid for the current steering angle and velocity. Increased velocity requires different steering input due to new slip angles, therefore requiring additional simulations for multiple steering angles in order to determine the optimal angle and peak velocity for a given corner radius. This, paired with the general high computing time, led to the use of a simple point mass vehicle dynamics model.

For a point mass model, stopping distance is given by the kinetic energy at velocity v and the work W_b performed during braking to standstill. Kinetic energy of the moving vehicle is

$$E_k = \frac{1}{2}mv^2, \tag{3.3.1}$$

and braking work by a constant force along distance d, with friction force F_f as given in Equation 2.1.1 is

$$W_b = F_f d = \mu mgd. \tag{3.3.2}$$

Braking distance d is found by substitution into Equation 3.3.1 and 3.3.2:

$$d = \frac{v^2}{2\mu g}, \tag{3.3.3}$$

where v is the vehicle speed in m/s, μ is the coefficient of friction ranging approximately 0.1 - 1.2 and g is the gravitational acceleration in m/s^2. Braking distance thus increase exponentially with vehicle speed, and decrease with road friction. The coefficient of friction is estimated with Pacejka's Magic Formula, where friction is a function of wheel slip and surface type as described in Section 2.1. Wheel slip is given by CAN signals for wheel speed relative to vehicle speed.

Maximum cornering speed for an unbanked curve of radius R is given by the centripetal acceleration, friction coefficient μ and gravitational acceleration g. Force along the local Y-axis due to centripetal acceleration is

$$F_c = ma_c = m\left(\frac{v^2}{R}\right) \tag{3.3.4}$$

with friction force previously given as

$$F_f = \mu mg. \tag{3.3.5}$$

Maximum speed is achieved when centripetal force F_c equals friction F_f

$$v^2 = \frac{\mu mgR}{m} : \tag{3.3.6}$$

resulting in the following expression for maximum cornering speed:

$$v = \sqrt{\mu Rg}. \tag{3.3.7}$$

A simple method of identifying corner radius is to fit a circle to three points in space, as Figure 3.3.6 illustrates.

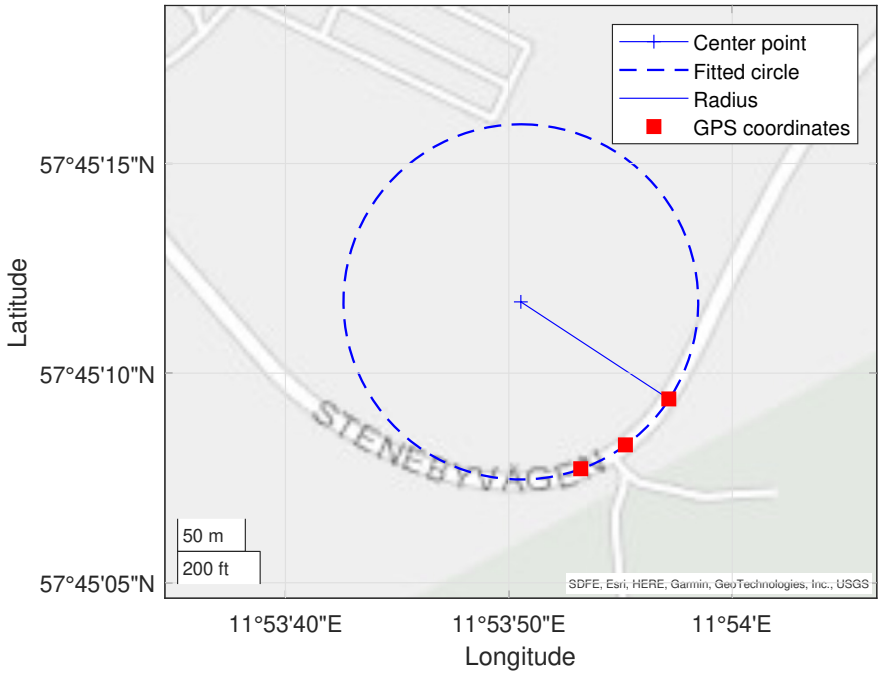


Figure 3.3.6: *Radius of a circle fitted to three GPS coordinates.*

This is implemented in an automotive setting by capturing the GPS position with a sampling interval of 1 Hz, providing a balance between responsiveness and resistance to signal noise. The accuracy of the radius estimation is evaluated in Chapter 5.2.

3.3.5 Addressing misclassifications with an arbiter

To combat misclassifications, a rule-based arbiter gathers the output labels of the image classifiers, time of year, CAN data and API responses and presents an enumerated label for the final road surface estimation. This is done in an effort to minimise the misclassifications identified in previous research and during neural network training. It is presented by a flowchart in Appendix B, where switches for input signals are indicated in blue, cases for the respective signal values are yellow and output labels are green.

Two outputs are given: road surface condition (”Gravel”, ”Dry”, ”Wet”, ”Snow”, ”Ice” and ”Unknown”) and a boolean signal for consensus, acting as a rudimentary quality flag.

3.4 Verification and validation

In this section, the two testing environments in this work are presented.

3.4.1 Simulink

The Simulink model is verified with pre-recorded video sequences and matching time series data of CAN signals with temperature, GPS position and velocity. Driving routes are replayed in Simulink and output data (classifications, bird’s-eye view transformed images) is monitored.

3.4.2 Automotive proving ground

The Simulink model validity and the fulfilment of problem statements are verified on public roads and proving grounds. Real-time CAN signals for temperature, vehicle and wheel speeds are used. GPS position is fed through the MathWorks cloud using the Matlab Android app.

Test scenarios include straight-line braking and steady-state cornering on various surface conditions as detailed in Chapter 1.3 Contributions and is summarised in Table 3.4.1 below:

Scenario	Surface condition	Vehicle speed [km/h]	Radius [m]
Straight-line braking	Dry, wet asphalt	40, 100	-
Straight-line braking	Gravel	40, 80	-
Straight-line braking	Low μ basalt tiles	40, 80	-
Steady-state cornering	Dry, wet asphalt	-	22.5, 37.5

Table 3.4.1: Test scenarios on the automotive proving ground.

The grip estimator is evaluated by measuring stopping distances in the braking scenarios, and by experimentally identifying maximum cornering speed on skidpad testing for the cornering scenarios. These values are compared to the distances and speeds predicted by the grip estimator.

4 Evaluation of road surface condition classification

This chapter evaluates the classification performance, with sections dedicated to neural network training performance, proving ground and public road testing.

4.1 Neural network training performance

Training results of the SqueezeNet convolutional neural networks are presented in this section. All possible combinations of predicted and true class labels are shown as confusion matrices, with additional columns indicating the per class accuracy. As the image classification task is divided into two sections — one for the bird's-eye view projection of the driven lane and one for the full image — the results are presented seperately for each neural network.

The main network tasked with analysing the driven lane performed well in all categories but snow, as shown in Figure 4.1.1. A noteworthy aspect of the matrix is the performance discrepancy between the ice and snow classes, where the latter was often misclassified as wet asphalt.

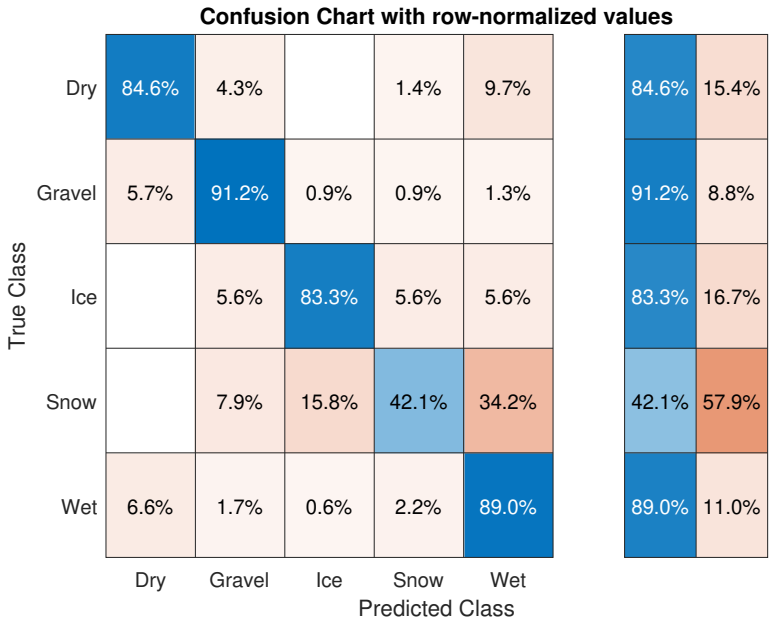


Figure 4.1.1: *Confusion matrix for the multi-label classification network trained on the complete data set combining BDD100K and own images.*

The second network, trained for binary classification of snow–snowless conditions in the full image, performed with greater accuracy and is presented in Figure 4.1.2.

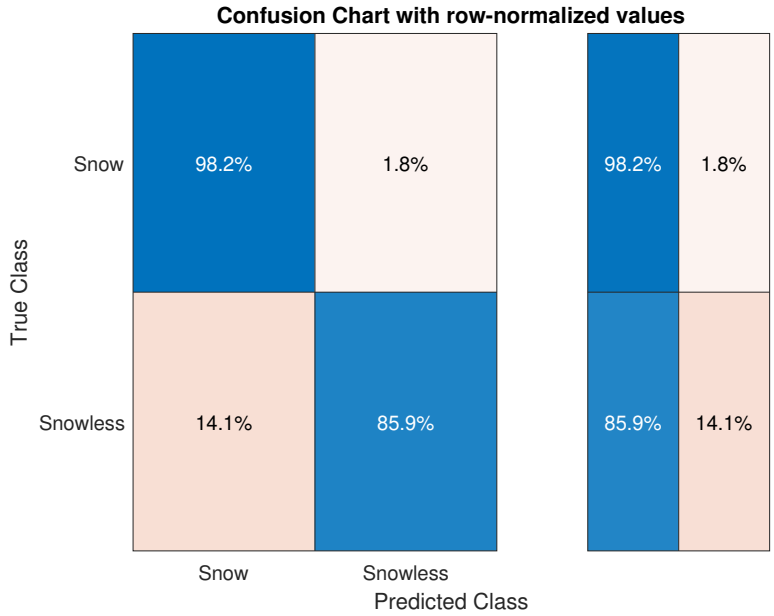


Figure 4.1.2: *Confusion matrix for the binary classification network.*

Lastly, GPU runtime performance was evaluated for the image pipeline and the outcome is presented in Table 4.1.1 below. Other computational operations, besides road database lookup, had negligible performance impact.

GPU	**Form factor**	**CUDA cores**	**Compute capability**	**Runtime [ms]**	**Rate [fps]**
Nvidia MX150	Laptop	384	6.1	69	14
Nvidia GTX 1070	Desktop	1920	6.1	16	63

Table 4.1.1: GPU runtime performance for the image pipeline.

4.2 Public road and proving ground testing

Qualitative testing was performed on Hällered proving ground on June 2, 2022 in mainly dry and partially light rain conditions, with air and track temperatures of 14 and 12 degrees Celsius respectively. The proving ground is depicted in Figure 4.2.1 with its layout described in the subsequent paragraph. The classified road surface conditions are displayed as coloured markers.

Figure 4.2.1: *Classified road surfaces on Hällered proving ground.*

The main oval, marked in blue, is a paved high-speed circuit. For this section of road, the image classifier reached 91.6% accuracy.

North-west of the oval lies a gravel track with a paved northernmost section. Here, an accuracy of 88.8%. A skidpad is located in the middle of the oval track and is further detailed in Figure 4.2.2 below. The dedicated track for Noise, vibration, and harshness (NVH) testing forms a loop along the western part of the oval, with the low μ basalt tile track routed along the straight section.

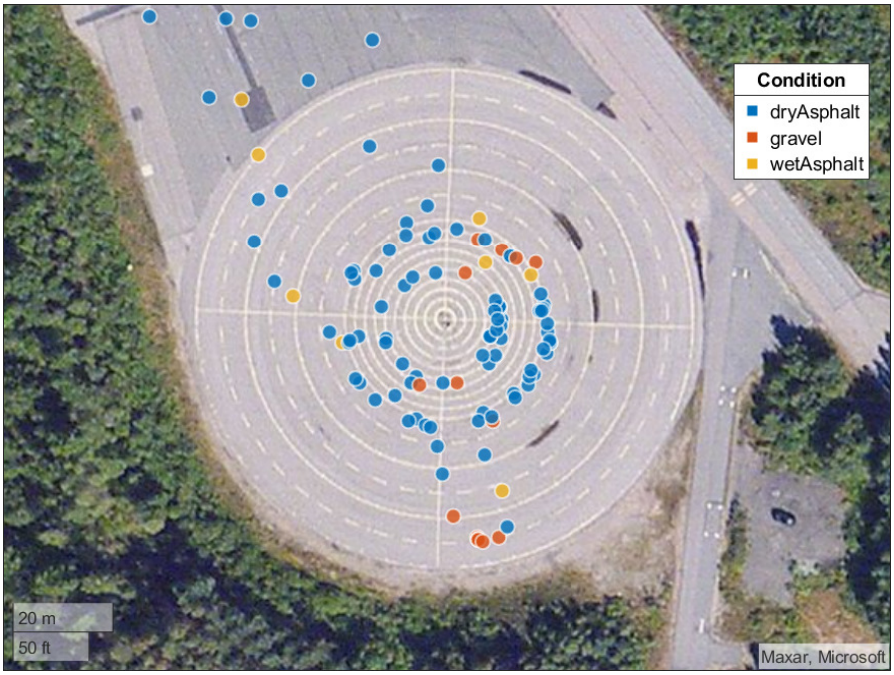


Figure 4.2.2: *Classified road surfaces from skidpad testing.*

As indicated by the "gravel" data points in the above figure, the image classifier performed slightly worse on the coarse paved surface of the skidpad than the main oval, with 81.2% accuracy on the skidpad for an aggregated dry pavement-accuracy of 86%. The scattered position of the data points is due to the low logging frequency of 1 Hz, hinting that three to four laps were carried out at each circular driving route of 12 and 20 meter radius respectively.

The dataset include little to no gravel images captured at higher vehicle speeds, neglecting to account for motion blur found in testing. This phenomenon became apparent during the virtual verification phase, where a movie and synthetic CAN-data of a pre-recorded gravel route were supplied to the Simulink model. The classifier output transitioned from gravel to dry asphalt as the vehicle speed increased, marked in blue in Figure 4.2.3 as the vehicle travelled north. However, none of this phenomenon was noticed on the automotive proving ground.

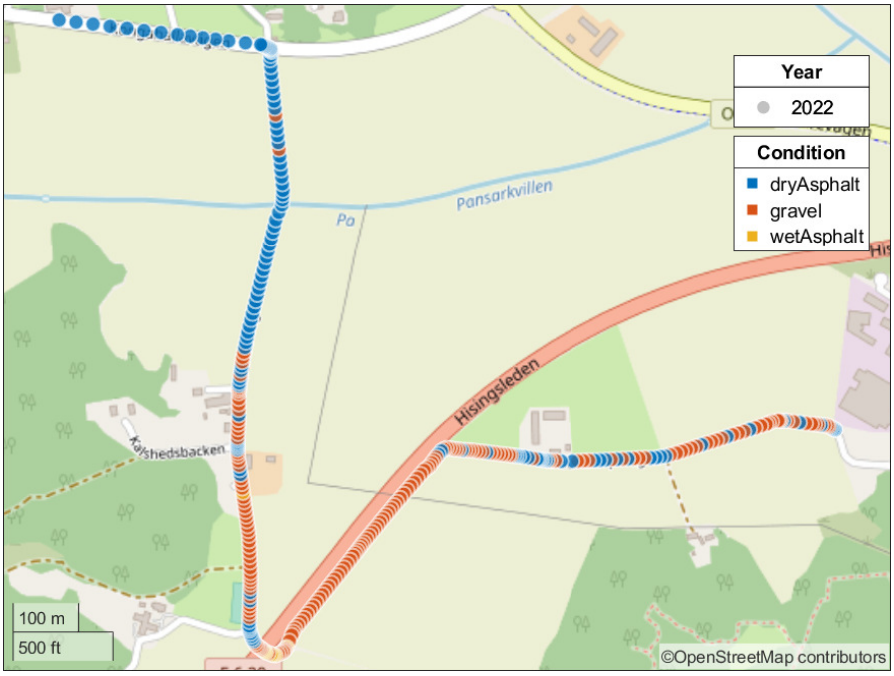


Figure 4.2.3: *Road surface classifications on a gravel route along the main road Hisingsleden. The north-south road stretch is incorrectly classified as dry asphalt (blue).*

Lastly, Figure 4.2.4 exhibit classification performance in medium rain on public roads with occasional breaks in rainfall. Condition was correctly identified a majority of the time.

Figure 4.2.4: *Classified road surfaces on wet asphalt.*

5 Evaluation of vehicle handling estimation

To evaluate the limits of road friction and minimise the influence of other factors, straight-line braking and steady-state cornering performance was investigated. Due to limited track time and little variation in weather, testing was not performed for wet asphalt or snow conditions. First, minimum braking distance estimations were evaluated on gravel, damp asphalt and an emulated ice surface on level ground. Second, maximum cornering speeds were identified through skidpad testing for two circular paths.

5.1 Straight-line braking

Table 5.1.1 provides experimental data on braking performance for the Lynk & Co 01 passenger vehicle, with estimated braking distances and estimation errors. Notable is the large estimation error on gravel surface and braking distance variance at 80 km/h.

Condition	**Target velocity [km/h]**	**Estimated braking distance [m]**	**Braking distance [m]**	**Sample Size N [-]**	**Distance estimation error [%]**
gravel	40	10.5	14.9 ±1.0	5	-26
	60	23.6	34.6 ±0.8	4	-31
	80	41.9	62.5 ±8.8	4	-33
asphalt [2]	100	39.3	43.7 ±0.2	4	-11
ice [3]	85	284 [4]	225.7	1	+26

Table 5.1.1: Straight-line braking tests given as mean and standard deviation.

5.2 Steady-state cornering

Cornering performance is displayed in Table 5.2.1 below. This test served two purposes: validating both the estimation of curvature radius, and the maximum velocity estimator.

Condition	**Corner radius [m]**	**Estimated radius [m]**	**Maximum velocity [km/h]**	**Estimated velocity [km/h]**	**Velocity estimation error [%]**
dry asphalt	12	12.3	39	40	+2.6
	20	20.2	49	51	-4.1

Table 5.2.1: Skidpad cornering tests.

It is apparent from this table that estimations of radius and safe cornering speed agree well with measured data. Testing was only performed in one type of surface condition due to limited track time.

[2] Slightly damp paved surface.
[3] Low μ basalt tile surface that emulate slippery winter conditions.
[4] Value calculated in post for icy condition, $\mu = 0.1$.

6 Discussion and future work

This chapter discusses potential applications within the automotive environment. Furthermore, testing procedure, dataset and neural network are evaluated. Finally, future work is suggested.

6.1 Benefits and ethical aspects of a practical application

As detailed in Chapter 1.1, preemptive grip estimation has numerous applications within vehicle safety systems. Currently grip is estimated a few meters ahead of the vehicle which limits usability to some extent. It has great value for steady surface conditions, however transient conditions might lead to insufficient notice for some safety systems. Ideally the prediction should hold true for a patch of road further away than the stopping distance required under poor conditions.

An ethical survey presented only a few negative ethical aspects: an increased prevalence of camera recordings and image recognition, posing a potential privacy invasion. Car manufacturers offer little to no insight of code implementation to the public, and this could prove problematic if there is no regulatory oversight.

On a more technical level, ethical impact depends on how the grip estimator is implemented in a vehicle. Driver reliance on the system is low if the system only acts in an informing manner with auditory and visual warnings, but if implemented as an intervening system that dictate break levels and vehicle speed consequences increase. The accommodation of machine learning in a regulatory setting is still controversial, as identified in [38]. One example is ISO 26262, an international standard that seeks to improve the safety of electronic systems in serial production road vehicles. In this, the SqueezeNet neural networks are considered non-transparent as it does not provide the a reliable way of tracing the results of each layer and would therefore require great effort to assess error rates and residual risks in order to comply with the ISO standard. The use of an arbiter could potentially mitigate these problems by shifting some of the dangers of misclassifications to a controlled, rule-based environment.

This study will further contribute to research within road traffic safety, an area of work in *Sustainable transport* that is part of the United Nations Sustainable Development Goal 11. Further, the potential application within adaptive cruise control has the potential of reducing particulate matter emissions by lessening the usage of friction brakes and thus the brake pad wear.

6.2 Test results

Experimental uncertainty is significant for certain aspects of the tests, by reason of the arbiter that potentially masks image misclassifications and outputs a more suitable classification. One such example is the nullification of "snow" classifications when the ambient temperature is above a threshold. Since only a few Simulink signals were logged, it was not possible to delve into the arbiter's impact in post.

Road condition was correctly classified in all scenarios, showing occasional deviating data points at each track on the automotive proving ground. In some parts of testing, track was slightly damp from light rain. Most of the time, this was not identified by the image classifier and the condition did not seem to impact vehicle performance. Notable is that the rumble strips and cobblestone in the NVH track was classified as gravel. Cobblestone in particular was considered for inclusion in the neural network training, but was decided against as no uniform friction value or braking performance could be found in previous research. The geometric pattern of cobblestone lends itself to image classifiers and would likely be easily identifiable.

For braking distances, a combined absolute estimation error of 22% was found when weighing all three tested surface conditions equally. This error could arguably be attributed to mismatched slip-friction curves, and by extension peak μ values, for the type of gravel road used at the proving ground. The vehicle experienced noticeable skidding due to loose gravel. Asphalt testing provided more accurate estimations, and due to the time of year, snow- and ice testing were substituted with a low μ basalt tile track watered by sprinklers. The arbiter classified the condition as "wet asphalt", understandably resulting in too short braking distances. The braking distance presented in Table 5.1.1 in Chapter 5 for "ice" was therefore calculated in post, emulating winter conditions.

The estimation of curvature radius and safe cornering speeds performed well beyond expectations and, in hindsight, would have made for an interesting subject if more surface types were included. The radius estimation seemed both accurate and precise, as the numerical value deviated very little and was close to the

radius of the circular paths travelled by the vehicle. Here, GPS antenna placement and its exact distance relative to the circular road markings play a significant role but was not measured during testing.

6.3 Dataset and neural network analysis

Image classification performance was weaker than expected, reaching only 78.0% accuracy due to poor snow performance, with 34% of snow pictures misclassified as wet asphalt. This contrasts to 92.2 % in [39] with a similar ROI transformation of the driving lane and BDD100K, Oxford Robocar and Cityscapes datasets sorted into five classes of road condition: wet and dry asphalt, wet and dry cobblestone and snow. [5] reached 88.8 % with a similar set of classes, replacing snow with dry and wet gravel. [7] delved deeper into winter condition classification by using classes for snow, slush, wet and dry asphalt, reaching 97.3 % accuracy when performing full image classification. Noteworthy is that these studies classified a similar number of road conditions, but with datasets two to eight times as large. As previously shown in Chapter 2.3, greater training accuracy can be achieved by training the network for more epochs, though it risks overfitting the network to the training data. Validation accuracy would not improve and validation loss would remain high.

The generality of the results is limited by the image dataset variety and road database coverage, but these issues are easily addressed by retraining the network to images of roads local to the intended area. A greater variety of images for each surface condition would for instance better capture the fine differences found in gravel roads due to varying shade and stone size. The increased variety of images could be utilised for additional classifications by differentiating loose and tightly packed gravel.

The dataset made an perhaps unnecessary distinction between snow- and ice-covered roads, leading to poor classification performance for those conditions. This includes an at times inconsistent labelling of similar images, exemplified by the subjective task of sorting slushy roads as ice or snow. Figure 6.3.1 illustrates such a scenario. A potential solution to more accurately distinct between snow and ice is to implement the light absorption method developed by RoadEye, as described in Section 1.2.

(a) *Snow* (b) *Ice*

Figure 6.3.1: *Illustration of the subjective distinction between snow- and ice-covered roads.*

6.4 Future work

Future studies should address generality, both in terms of geographical coverage and operating conditions. Low light scenarios can be addressed by extending the datasets with dusk images or by specialising on a smaller patch of road illuminated by the headlights, providing better control over light conditions. Classification validity can be improved with cross validation methods such as k-fold or Monte Carlo.

Pavement information from road databases should be handled by protocols in the likes of ADASIS. Besides providing better coverage and flexibility in road data, it would further facilitate implementation in a vehicle ECU. Information on upcoming roads can be given from the planned route or most probable path.

Crucially, the arbiter need further work in terms of robustness and verification, as very little data of the internal signals were saved during testing. The signal for consensus should be adapted into a quality flag with multiple values, as is customary in vehicle networking applications. This would allow potential dependant functions like cruise control to enter safe modes when uncertain data is received.

7 Conclusion

It is shown that road grip can be estimated from the image feed of a forward-facing camera onboard a passenger vehicle. By applying a specially trained convolutional neural network, prediction accuracy reached 78% for pre-recorded images of five different surface types. When paired with an arbiter that consider ambient temperature and online road pavement data, road condition was correctly classified 86% of the time on dry asphalt, and 89% of the time on gravel. Snow and ice classification accuracy was not measured in vehicle due to the time of year.

Road grip for each surface type was determined from distinct tyre characteristics as modelled by the Pacejka Magic Formula. Minimum braking distance and safe cornering speed were determined with a point mass model and always presented to the driver. Braking performance was estimated with moderate success, while curvature radius and safe cornering speeds were estimated within 1.8% and 3.4% respectively.

This study adds to the research field by investigating ways of confirming or rejecting classification outputs by neural networks, and by validating image classification performance in actual vehicles on automotive proving grounds.

Bibliography

[1] C-G. Wallman et al, "Friction measurement methods and the correlation between road friction and traffic safety," Statens väg- och transportforskningsinstitut (VTI), Tech. Rep. 911A, 2001.

[2] I. Tang and T. Breckon, "Automatic road environment classification.," *IEEE Transactions on Intelligent Transportation Systems, Intelligent Transportation Systems, IEEE Transactions on, IEEE Trans. Intell. Transport. Syst*, vol. 12, no. 2, pp. 476–484, 2011, ISSN: 1558-0016.

[3] P. Taylor, S. S. Anand, N. Griffiths, F. Adamu-Fika, A. Dunoyer, and T. Popham, "Road type classification through data mining," in *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, ser. AutomotiveUI '12, Portsmouth, New Hampshire: Association for Computing Machinery, 2012, pp. 233–240, ISBN: 9781450317511. DOI: 10.1145/2390256.2390295. [Online]. Available: https://doi.org/10.1145/2390256.2390295.

[4] V. Tumen, O. Yildirim, and B. Ergen, "Recognition of road type and quality for advanced driver assistance systems with deep learning," *Elektronika ir Elektrotechnika*, vol. 24, no. 6, pp. 67–74, Dec. 2018. DOI: 10.5755/j01.eie.24.6.22293. [Online]. Available: https://eejournal.ktu.lt/index.php/elt/article/view/22293.

[5] E. Šabanovič, V. Žuraulis, O. Prentkovskis, and V. Skrickij, "Identification of road-surface type using deep neural networks for friction coefficient estimation," *Sensors*, vol. 20, no. 3, 2020, ISSN: 1424-8220. DOI: 10.3390/s20030612. [Online]. Available: https://www.mdpi.com/1424-8220/20/3/612.

[6] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, *Squeezenet: Alexnet-level accuracy with 50x fewer parameters and ¡0.5mb model size*, 2016. arXiv: 1602.07360 [cs.CV].

[7] S. Roy Chowdhury, M. Zhao, A. Wallin, N. Ohlsson, and M. Jonasson, "Machine learning models for road surface and friction estimation using front-camera images," in *2018 International Joint Conference on Neural Networks (IJCNN)*, 2018, pp. 1–8. DOI: 10.1109/IJCNN.2018.8489188.

[8] A. Jonnarth, "Camera-based friction estimation with deep convolutional neural networks," M.S. thesis, Uppsala University, Division of Systems and Control, 2018.

[9] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, *Densely connected convolutional networks*, 2018. arXiv: 1608.06993 [cs.CV].

[10] S. Löfving, "Optisk metod att uppskatta egenskaper för ett is-eller vattenbelagt mätobjekt," Swedish patent SE 531949, May 2007.

[11] OmniKlima, *Solutions for detection of road condition*, Accessed: 2022-02-18. [Online]. Available: http://omniklima.com/.

[12] RoadCloud Oy. Accessed: 2022-02-18. (), [Online]. Available: https://roadcloud.com/sv/.

[13] M. Kutila, P. Pyykönen, K. Kauvo, and P. Eloranta, "In-vehicle sensor data fusion for road friction monitoring," in *2011 IEEE 7th International Conference on Intelligent Computer Communication and Processing*, 2011, pp. 349–352. DOI: 10.1109/ICCP.2011.6047895.

[14] F. Minges, "Neural network-based road friction estimation using road weather information," M.S. thesis, Chalmers University of Technology, Division of Vehicle Safety, 2020. [Online]. Available: https://hdl.handle.net/20.500.12380/300762.

[15] A. Asp, J. Casselgren, D. Eriksson, *et al.*, "Slutrapport införande av digital vinterväglagsinformation – "digital vinter"," Trafikverket, Tech. Rep. 2021:114, 2021, p. 11.

[16] Nira Dynamics, *Road perception*, Accessed: 2022-02-18. [Online]. Available: https://niradynamics.se/road-perception/.

[17] Continental AG. "Preview esc." Accessed: 2022-02-23. (), [Online]. Available: https://www.continental.com/en/press/fairs-events/techshow-2019/previewesc/.

[18] Tactile Mobility, *Oem car makers*, Accessed: 2022-02-18. [Online]. Available: https://tactilemobility.com/solution/car-makers/.

[19] H. Pacejka, *Tyre and Vehicle Dynamics*. Oxford: Butterworth-Heinemann, 2002, pp. 156–215, ISBN: 0750651415.

[20] D. Tavernini, M. Massaro, E. Velenis, D. I. Katzourakis, and R. Lot, "Minimum time cornering: The effect of road surface and car transmission layout," *Vehicle System Dynamics*, vol. 51, no. 10, pp. 1533–1547, 2013. DOI: 10.1080/00423114.2013.813557. [Online]. Available: https://doi.org/10.1080/00423114.2013.813557.

[21] Michelin. "Adhesion and indentation." Accessed: 2022-03-19. (), [Online]. Available: https://thetiredigest.michelin.com/performance-safety.

[22] SARSYS-ASFT, *Sarsys-asft friction measuring system*, Accessed: 2022-02-18. [Online]. Available: https://www.sarsys-asft.com/measuring-system.
[23] L. Sjögren et al, "Krav på belagda väg-, cykel- och gångbanors friktionsegenskaper vid barmarksförhållanden," Statens väg- och transportforskningsinstitut (VTI), Tech. Rep. 2015/0280-9.1, 2020.
[24] Trafikverket, *Standardbeskrivning för basunderhåll väg (sbv)*, 2021. [Online]. Available: https://www.trafikverket.se/for-dig-i-branschen/vag/underhall-vag/Entreprenorsdokument/Kontraktsdokument/.
[25] "Road vehicles — Multidimensional measurement and coordinate systems definition," The International Organization for Standardization (ISO), Tech. Rep. 21002:2021(en), 2021.
[26] B. Jacobson et al, *Compendium in Vehicle Motion Engineering*. Chalmers University of Technology, 2021. [Online]. Available: https://research.chalmers.se/en/person/bengtja#publications.
[27] R. Pütz and T. Serné, *Race Car Handling Optimization*. Wiesbaden: Springer Wiesbaden, 2022, pp. 148–155, ISBN: 978-3-658-35200-4. DOI: 10.1007/978-3-658-35200-4. [Online]. Available: https://doi.org/10.1007/978-3-658-35200-4.
[28] IBM. "What is computer vision?" Accessed: 2022-09-05. (), [Online]. Available: https://www.ibm.com/se-en/topics/computer-vision.
[29] S. J. Russell and P. Norvig, *Artificial intelligence : a modern approach*. Ser. Prentice Hall Series in Artificial Intelligence. Pearson Education Limited, 2016, pp. 928–947, ISBN: 9781292153964.
[30] A. Zaras, N. Passalis, and A. Tefas, "Chapter 2 - neural networks and backpropagation," in *Deep Learning for Robot Perception and Cognition*, A. Iosifidis and A. Tefas, Eds., Academic Press, 2022, pp. 17–34, ISBN: 978-0-323-85787-1. DOI: https://doi.org/10.1016/B978-0-32-385787-1.00007-5. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780323857871000075.
[31] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks.," in *AISTATS*, Y. W. Teh and D. M. Titterington, Eds., ser. JMLR Proceedings, vol. 9, JMLR.org, 2010, pp. 249–256.
[32] F. Yu, H. Chen, X. Wang, *et al.*, "Bdd100k: A diverse driving dataset for heterogeneous multitask learning," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2020.
[33] Trafikverket. "Nationell vägdatabas - lastkajen." Accessed: 2022-05-11. (), [Online]. Available: https://lastkajen.trafikverket.se/.
[34] ——, "Trafikverkets öppna api för trafikinformation." Accessed: 2022-05-11. (), [Online]. Available: https://api.trafikinfo.trafikverket.se/.
[35] ADASIS. "Advancing map-enhanced driver assistance systems leading to automated driving." Accessed: 2022-05-20. (), [Online]. Available: https://adasis.org.
[36] QGIS. "A free and open source geographic information system." Accessed: 2022-05-19. (), [Online]. Available: https://qgis.org.
[37] K. Kearney. "Interpshapefile, mathworks file exchange." Accessed: 2022-05-19. (), [Online]. Available: https://mathworks.com/matlabcentral/fileexchange/10278-interpshapefile.
[38] R. Salay, R. Queiroz, and K. Czarnecki, "An analysis of ISO 26262: Using machine learning safely in automotive software," *CoRR*, vol. abs/1709.02435, 2017. arXiv: 1709.02435. [Online]. Available: http://arxiv.org/abs/1709.02435.
[39] A. Busch, D. Fink, M.-H. Laves, Z. Ziaukas, M. Wielitzka, and T. Ortmaier, "Classification of road surface and weather-related condition using deep convolutional neural networks," in *Advances in Dynamics of Vehicles on Roads and Tracks*, M. Klomp, F. Bruzelius, J. Nielsen, and A. Hillemyr, Eds., Cham: Springer International Publishing, 2020, pp. 1042–1051, ISBN: 978-3-030-38077-9.

Appendices

A Simulink model

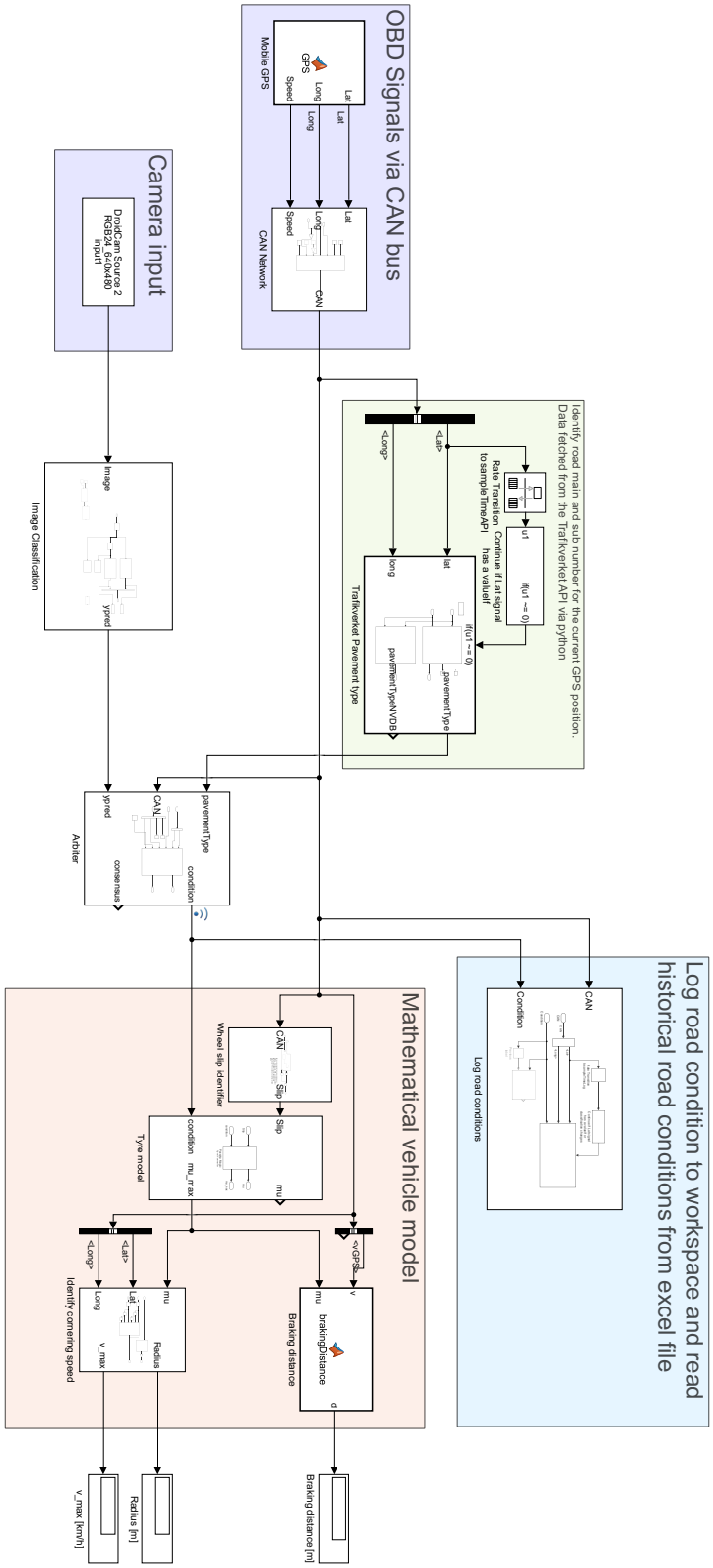


Figure A.0.1: *Top-level overview of the grip estimator implementation in Simulink. When viewed in landscape, information travel from left to right. Vehicle inputs are marked in purple and databases marked in green.*

B Arbiter switch case flowchart

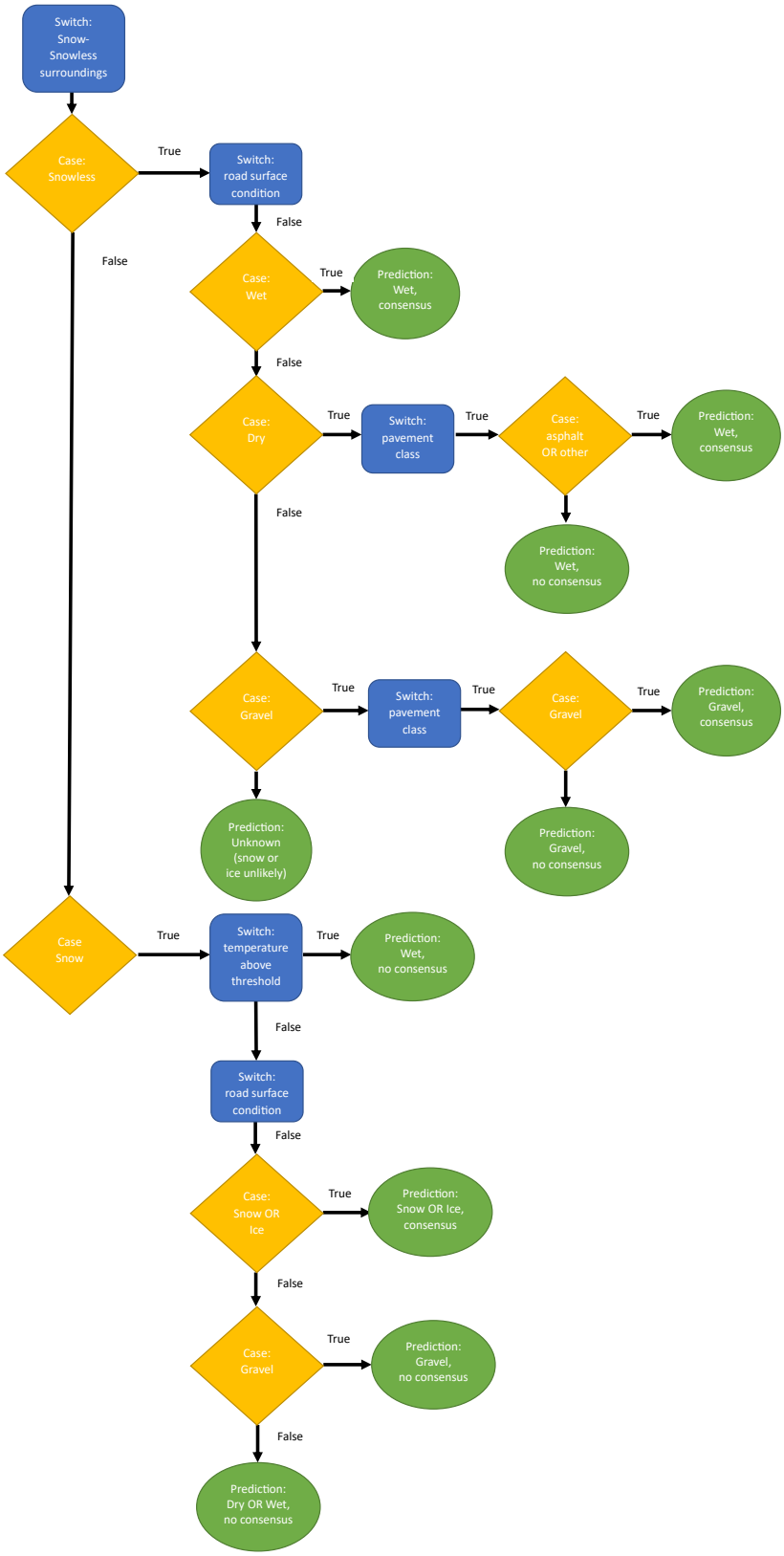


Figure B.0.1: *Switch case flowchart for the road surface condition arbiter, outputting predicted surface condition and status of consensus (green labels).*

C Brake distance measurements

Test scenario	**Velocity [km/h]**	**Estimated braking distance [m]**	**VBOX corrected braking distance [m]**	**Average distance [m]**	**Std. dev. [m]**	**Mean abs. dev. [m]**	**Average deceleration [g]**
Dry	100.1	39	43.68				
asphalt	101.7	39	43.54				
	101.2	39	44.08				
	100.96	39	43.57				
				43.72	0.249	0.181	0.90
Gravel	43.45	11	15.51				
	43.67	11	14.72				
	41.92	11	13.27				
	42.79	11	15.96				
	42,04	11	15.05				
				14.90	1.026	0.726	0.42
Gravel	61.13	24	34.17				
	61.98	24	35.72				
	61.52	24	33.94				
	61.9	24	34.71				
				34.64	0.792	0.580	0.41
Gravel	82.13	42	54.99				
	84.5	42	60.81				
	82.69	42	58.91				
	86.23	42	75.2				
				62,48	8.82	6.36	0.43
Low μ	85.2	284	225.7				
basalt tiles				225.7	-	-	0.14

Table C.0.1: Velocity-corrected braking distances at the Hällered proving ground, measured with a VBOX 3iSL data logger.