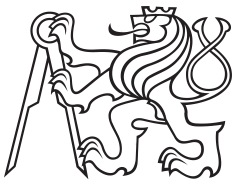
Diplomová práce



České
vysoké
učení technické
v Praze

F3

Fakulta elektrotechnická
Katedra kybernetiky

Učení pro vizuální rozpoznávání vlastností povrchu

Bc. Martin Vlašimský

Vedoucí: Ing. Jan Čech, Ph.D.
Obor: Kybernetika a robotika
Studijní program: Kybernetika a robotika
Srpen 2020

ZADÁNÍ DIPLOMOVÉ PRÁCE

I. OSOBNÍ A STUDIJNÍ ÚDAJE

Příjmení: **Vlašimský** Jméno: **Martin** Osobní číslo: **434670**

Fakulta/ústav: **Fakulta elektrotechnická**

Zadávající katedra/ústav: **Katedra řídicí techniky**

Studijní program: **Kybernetika a robotika**

Studijní obor: **Kybernetika a robotika**

II. ÚDAJE K DIPLOMOVÉ PRÁCI

Název diplomové práce:

Učení pro vizuální rozpoznávání vlastností povrchu

Název diplomové práce anglicky:

Training for Visual Surface Recognition

Pokyny pro vypracování:

Vizuální rozpoznání vlastností povrchu před vozidlem má důležitý význam pro bezpečnost řízení, vzhledem k nerovnostem a adhezi.
1) Navrhněte metodu, která za jízdy vozidla automaticky asociuje vnímané vlastnosti povrchu (z akcelerace nebo řídicích signálů) s obrazem pořízeným kamerou.
2) Natrénujte konvoluční neuronovou síť, která z takto anotovaných obrazových dat predikuje vlastnosti povrchu.
3) Proveďte experiment a kvantitativně vyhodnoťte přesnost predikce.
Pozn.:
1) Zaměřte se na učení nerovností povrchu, volitelně uvažujte i učení adhezních vlastností.
2) K experimentům použijte pokud možno instrumentovaný RC model auta, alternativně skutečný automobil vybavený synchronním sběrem dat z kamery a inerciální jednotky (např. mobilní telefon).

Seznam doporučené literatury:

[1] Olaf Ronneberger, Philipp Fischer, Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation. In Proc. Medical Image Computing and Computer-Assisted Intervention. 2015.
[2] Sebastian Thrun et al.; Stanley: The Robot that Won the DARPA Grand Challenge. Journal of Field Robotics 23(9), 661–692, 2006.
[3] C. Brooks, K. Iagnemma. Vibration-based terrain classification for planetary exploration rovers. IEEE Transactions on Robotics, 21(6), 185–1191, 2005.

Jméno a pracoviště vedoucí(ho) diplomové práce:

Ing. Jan Čech, Ph.D., skupina vizuálního rozpoznávání FEL

Jméno a pracoviště druhé(ho) vedoucí(ho) nebo konzultanta(ky) diplomové práce:

Datum zadání diplomové práce: **13.02.2020** Termín odevzdání diplomové práce: **14.08.2020**

Platnost zadání diplomové práce:
do konce zimního semestru 2021/2022

Ing. Jan Čech, Ph.D.	prof. Ing. Michael Šebek, DrSc.	prof. Mgr. Petr Páta, Ph.D.
podpis vedoucí(ho) práce	podpis vedoucí(ho) ústavu/katedry	podpis děkana(ky)

 © ČVUT v Praze, Design: ČVUT v Praze, VIC

III. PŘEVZETÍ ZADÁNÍ

Diplomant bere na vědomí, že je povinen vypracovat diplomovou práci samostatně, bez cizí pomoci, s výjimkou poskytnutých konzultací.
Seznam použité literatury, jiných pramenů a jmen konzultantů je třeba uvést v diplomové práci.

. _____
Datum převzetí zadání

_____
Podpis studenta

 © ČVUT v Praze, Design: ČVUT v Praze, VIC

Poděkování

Rád bych poděkoval Ing. Janu Čechovi, Ph.D., vedoucímu mé diplomové práce, za odborné vedení, cenné rady, trpělivost a vstřícnost při konzultacích.

Prohlášení

Prohlašuji, že jsem předloženou práci vypracoval samostatně a že jsem uvedl veškerou použitou literaturu.

V Praze, 13. srpna 2020

Abstrakt

Tato práce se zabývá automatickým vytvořením anotovaných dat a následně predikcí hrubosti povrchu před autem ještě dříve než na něj auto vjede.

Manuální anotace dat je velmi zdlouhavá a může dojít k chybě způsobené lidským faktorem. Proto se zde z nasbíraných dat automaticky vypočetla veličina pro popis hrubosti povrchu a ta se následně asociovala s vyfocenou fotografií z jízdy. Takto anotovaná data byla následně použita pro naučení a validaci konvoluční sítě architektury ResNet-50, využité pro predikci hrubosti povrchu před autem.

Pro experimenty bylo využito RC modelu auta, které bylo již dříve na ČVUT obohaceno mnoha senzory a výpočetní elektronikou. S tímto modelem bylo nasbíráno velké množství dat volnou jízdou. V této práci bylo využito dat z akcelerometru z osy z, která odpovídají hrubosti přejetého povrchu, a fotografií nasbíraných z kamery přimontované ke střeše modelu auta.

Provedené experimenty ukázaly schopnost sítě naučit se na automaticky anotovaných datech predikovat hrubost povrchu na trénovací množině. Predikce na validačních množinách pak téměř držela trend, jako tomu bylo u trénovacích dat, avšak je potřeba ještě dalších experimentů pro možnost plného využití této metody v praxi pro predikci hrubosti povrchu.

Klíčová slova: automatická anotace dat, konvoluční neuronová síť, ResNet-50, predikce hrubosti povrchu

Vedoucí: Ing. Jan Čech, Ph.D.
ČVUT v Praze
Fakulta elektrotechnická
Katedra kybernetiky
Karlovo námestí 13
121 35 Praha 2

Abstract

This work deals with the automatic creation of annotated data and the subsequent prediction of the surface roughness in front of the car before the car even enters the surface.

Manual data annotation is very lengthy and human can make mistakes. Therefore, the quantity for the description of the surface roughness was automatically calculated from the collected data and this was subsequently associated with the photograph taken from the ride. The annotated data were used then to learn and validate the convolutional network of the ResNet-50 architecture used to predict the surface roughness in front of the car.

The RC model of the car was used for the experiments, which was previously enriched with many sensors and computer electronics at CTU. With this model, a large amount of data were collected on a free ride. In this work, data from the accelerometer from the z-axis were used, which correspond to the roughness of the traversed surface and photographs collected from a camera mounted on the roof of a car model.

The performed experiments showed the ability of the network to learn and predict the surface roughness on a training set on automatically annotated data. The prediction on the validation sets then almost followed the trend as in the case of the training data, but further experiments are needed for the possibility of full use of this method in practice for the prediction of surface roughness.

Keywords: automated annotation, convolutional neural network, ResNet-50, prediction of surface roughness

Title translation: Training for Visual Surface Recognition

Obsah

1 Úvod **1**

2 Používané metody v literatuře **3**

3 Navrhovaná metoda **5**

3.1 Výpočet veličiny popisující povrch 7

3.2 Úprava fotografií 8

3.3 Předzpracování dat 10

3.4 Použití neuronové sítě 10

4 Testovací platforma **13**

4.1 Model auta 13

4.1.1 Kamera 14

4.1.2 Výpočetní zařízení 16

4.1.3 Inerciální měřicí jednotka (IMU) 17

4.1.4 Výsledná TOMI platforma . . 18

5 Experimenty **21**

5.1 Použití dat z akcelerometru 21

5.2 Sběr dat 26

5.3 Datasety 26

5.4 Referenční metoda 30

5.5 Navrhovaná metoda 33

5.6 Detailní analýza výsledků 34

5.7 Výsledky 43

6 Závěr **45**

Literatura **47**

Obrázky

2.1 Planetární rover - umístění senzoru, [BI05] 3

2.2 Robot Stanley, vítěz DARPA Grand Challenge v roce 2005, [ea06] 4

3.1 Fotografie z levé čočky kamery s vyznačením oblasti zájmu 6

3.2 Oblast zájmu z obrázku 3.1 po ortorektifikaci 6

3.3 Hrubost povrchu pro obrázek 3.2 6

3.4 Jízda po kamenné dlažbě - levá čočka 8

3.5 Jízda po kamenné dlažbě - ortorektifikovaný obraz 9

3.6 Jízda po kamenné dlažbě - výřez z ortorektifikovaného obrazu 3.5 9

3.7 Příklad konvoluční sítě pro regresní úlohu [MSE17] 11

3.8 Příklad učení reziduální konvoluční sítě [HZRS15] 11

4.1 Losi Desert Buggy XL-E 1:5 4WD, [citc] 14

4.2 ZED kamera, [citg] 15

4.3 Ukázka fotky z kamery ZED při konstantní jízdě TOMI 15

4.4 Jetson AGX Xavier, [citb] 16

4.5 Raspberry Pi 2 s modulem Navio 2, [citd] 17

4.6 Inerciální měřicí jednotka Xsens MTi-7-DK, [citf] 17

4.7 TOMI platforma z boku 18

4.8 TOMI platforma s namontovanými komponentami 18

5.1 Jízda střídavě po asfaltu a trávě - Navio 2 22

5.2 Jízda střídavě po asfaltu a trávě - Xsens 23

5.3 Jízda po kamenné dlažbě - Navio 2 24

5.4 Jízda po kamenné dlažbě - Xsens 24

5.5 Asfalt s hodnotou veličiny popisující hrubost povrchu 27

5.6 Tráva s hodnotou veličiny popisující hrubost povrchu 27

5.7 Lesní cesta s hodnotou veličiny popisující hrubost povrchu 28

5.8 Panelová cesta s hodnotou veličiny popisující hrubost povrchu 28

5.9 Asfalt s výmoly a tráva s hodnotou veličiny popisující hrubost povrchu 28

5.10 Mapa projeté trasy, [cita] 29

5.11 Ortorektifikovaný obraz asfaltu a jeho obrazový gradient 32

5.12 Ortorektifikovaný obraz lesní cesty a jeho obrazový gradient 33

5.13 Dataset 1 - Histogram rozložení hrubosti povrchů 34

5.14 Dataset 1 - Hodnoty MSELoss pro dané epochy 35

5.15 Dataset 1 - Predikce na trénovacích datech, $\rho = 0{,}982$ 35

5.16 Dataset 1 - Predikce na validačních datech, $\rho = 0{,}921$ 36

5.17 Dataset 2 - Histogram rozložení hrubosti povrch 37

5.18 Dataset 2 - Hodnoty MSELoss pro dané epochy 37

5.19 Dataset 2 - Predikce na trénovacích datech, $\rho = 0{,}963$ 38

5.20 Dataset 2 - Predikce na validačních datech, $\rho = 0{,}903$ 38

5.21 Dataset 3 - Histogram rozložení hrubosti povrch 39

5.22 Dataset 3 - Hodnoty MSELoss pro dané epochy 39

5.23 Dataset 3 - Predikce na trénovacích datech, $\rho = 0{,}484$ 40

5.24 Dataset 3 - Predikce na validačních datech, $\rho = 0{,}05$ 40

5.25 Dataset 4 - Histogram rozložení hrubosti povrchů 41

5.26 Dataset 4 - Hodnoty MSELoss pro dané epochy 42

5.27 Dataset 4 - Predikce na trénovacích datech, $\rho = 0{,}929$ 42

5.28 Dataset 4 - Predikce na validačních datech, $\rho = 0{,}478$ 43

Tabulky

5.1 Korelační koeficienty obrazového gradientu pro jednotlivé datasety . 32

5.2 Korelační koeficienty výstupu neuronové sítě pro jednotlivé datasety 33

5.3 Korelační koeficienty predikcí a reálných hodnot pro jednotlivé datasety 44

ctuthesis t1606152353

Kapitola 1

Úvod

V dnešní době se neuronové sítě používají více než kdy dříve nejen v robotice a počítačovém vidění, ale i v dalších odvětvích za účelem získání co možná nejvíce strojů, které by byly schopné fungovat autonomně. Jeden z mnoha příkladů se dá najít v automobilovém průmyslu, kde je žádoucí zajistit plnou spolehlivost automobilu a zároveň zajistit komfort pro spotřebitele. Proto se využívá nespočet senzorů, které mají v tomto těžkém úkolu nahradit smysly člověka a s pomocí neuronových sítí musí být schopny předcházet krizovým situacím, pomáhat při jízdě a při nespočtu dalších situací. Tyto úlohy je možné řešit s příchodem nových technologií a jejich miniaturizací.

Jedním z důležitých parametrů pro jízdu automobilem je hrubost povrchu, po kterém auto jede. Nejde zde pouze o různé typy povrchů, ale i o různé nerovnosti či výmoly v rámci jednoho daného povrchu. Pro různé typy povrchů je potřeba brát v potaz například různé brzdné dráhy. Při projetí velkého výmolu pak může dojít k poruše automobilu [ea06]. Je proto nezbytně nutné znát povrch a jeho hrubost ještě dříve, než se na něj najede. V této práci se budeme zabývat jedním z možných řešení této problematiky s pomocí kamery a akcelerometru.

Rozpoznání povrchů před autem jako jsou například asfalt, tráva, štěrk a dlažba se již predikuje. Příklad takových predikcí je vysvětlen v článku [BFP+20]. Naučení klasifikátoru tedy není takový problém. Mnohem problémovější část je získání anotovaných dat. Většina dat se musí anotovat ručně a to zabírá hodně času. Také může docházet k špatné nebo nepřesné anotaci způsobené lidským faktorem.

V případě rozpoznání hrubosti povrchu však není manuální anotace možná. Povrchů je mnoho a popis hrubosti těchto povrchů je spojitý. Proto se tato práce zabývá automatickou anotací nasbíraných dat.

S autem jsou volně projety různé trasy s různými povrchy a jsou při tom zaznamenány naměřené hodnoty z akcelerometru a kamery. Data pro osu z, naměřená pomocí akcelerometru, pak odpovídají hrubosti přejetého povrchu.

Následně se z těchto dat vypočte veličina, která udává hrubost povrchu. Ta se poté asociuje s obrazem, tedy s fotkami nasbíranými ze stejných jízd. Takto předpřipravená data se použijí pro učení konvoluční neuronové sítě, kde je právě každé fotografii přiřazena vypočtená hodnota veličiny pro popis hrubosti povrchu zobrazeného na fotografii.

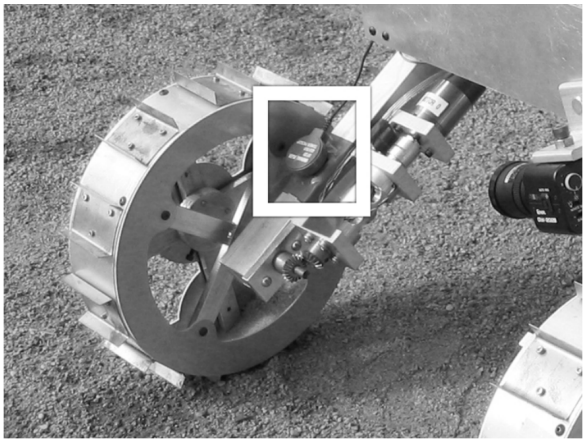
Ve výsledku jsme poté schopni při testování predikovat hrubost povrchů zobrazených na fotografii. Tato fotografie je však pořízena z jedoucího auta a díky tomu probíhá predikce pro povrch před autem, tedy jestě dříve než na tento povrch auto najede. Tím pádem by bylo možné následně využít této informace například k automatizovanému manévru automobilu pro jeho lepší stabilitu nebo pro vyhnutí se výmolům na silnici.

Kapitola 2

Používané metody v literatuře

Jako jeden z prvních případů použití predikce hrubosti nebo frikčních vlastností povrchu bylo v případech, kdy byl robot vystaven takovým okolním podmínkám, že člověk nemohl být v jeho blízkosti a nemohl tudíž sám tyto vlastnosti jasně určit.

Příkladem použití jsou u planetární rovery, které prozkoumávají povrchy na jiných planetách. Zde je nezbytné, aby byl rover vždy pojízdný. Pokud tedy například zapadne v písku nebo se stane jiná událost, která znemožní pohyb roveru, nemůže pak prozkoumávat rozlehlé oblasti. Proto je zapotřebí, aby rover dokázal odhadnout, na kterém povrchu jede, a předešel tedy možnému zaseknutí na jednom místě [BI05].



Obrázek 2.1: Planetární rover - umístění senzoru, [BI05]

Na obrázku 2.2 je zobrazen planetární rover s umístěním senzoru vibrací.

V tomto konkrétním případě se jedná o využití akcelerometru [BI05].

Druhým příkladem je pak autonomní vozidlo, robot Stanley. Stanley měl za úkol ujet 142 mil za méně než 10 hodin. K tomu bylo zapotřebí nespočet senzorů a výkonného výpočetního zařízení. Jedním ze senzorů byl akcelerometr, který měřil zrychlení osy z. Tato data byla přefiltrována pásmovou propustí, kdy se zbylé frekvence nacházely v oblasti rezonanční frekvence auta. Amplituda výsledného signálu pak představuje hrubost povrchu, přes který auto přejelo. Amplituda přefiltrovaných dat z akcelerometru se nejspíše mění lineárně se změnou rychlosti vozidla [ea06].

Obrázek 2.2: Robot Stanley, vítěz DARPA Grand Challenge v roce 2005, [ea06]

Kapitola 3

Navrhovaná metoda

Veškeré zkoumané metody se zaměřovaly na predikci hrubosti povrchu až v okamžik, kdy na něj robot vjel. To může být v mnoha případech ale pozdě a proto se tato práce zabývá predikcí hrubosti povrchu ještě než na něj robot nebo automobil najede. K tomu se využije obraz z kamery a data z akcelerometru z osy z.

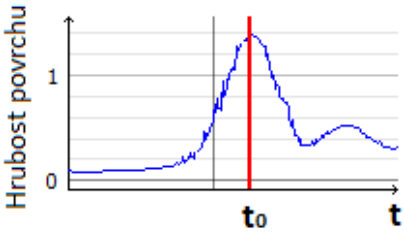
Pro získání relevantních informací o daném povrchu bylo nutné získat veličinu, kterou by výsledná neuronová síť byla schopna predikovat. V této práci byla predikovaná veličina určena s využitím STFT (krátkodobá Fourierova transformace) dat z akcelerometru z osy z. Zde byly provedeny experimenty s použitím pouze omezeného frekvenčního pásma z dat akcelerometru [ZZWL15].

Základem je tedy využití fotografie z kamery. V tomto obraze nás zajímá pouze jeho část, která odpovídá ploše, přes kterou projede auto. Fotografie s vyznačenou oblastí zájmu je zobrazena na obrázku 3.1.

Následně byl výřez ortorektifikován, což je zobrazeno na obrázku 3.2 společně s hodnotou hrubosti zobrazeného povrchu. Pro tento výřez, který má 1 metr na šířku a 1 metr na délku, je z akcelerometru vypočtena hodnota hrubosti povrchu zobrazeného na tomto obrázku. Graf s vyznačenou hodnotou pro tento obrázek je zobrazen na obrázku 3.3.

Obrázek 3.1: Fotografie z levé čočky kamery s vyznačením oblasti zájmu

Obrázek 3.2: Oblast zájmu z obrázku 3.1 po ortorektifikaci



Obrázek 3.3: Hrubost povrchu pro obrázek 3.2

3.1 Výpočet veličiny popisující povrch

Pro výpočet veličiny popisující hrubost povrchu bylo využito STFT. Výsledné hodnoty frekvencí jsou pak uvedeny jako komplexní čísla. Výpočet probíhal podle následujícího vzorce popsaném v [Cite]:

$$x[n] = \frac{\sum_t x_t[n]w[n - tH]}{\sum_t w^2[n - tH]}, \tag{3.1}$$

kde:
$x[n]$ = pole frekvencí po transformaci,
$x_t[n]$ = pole signálu v časové oblasti,
$w[n]$ = velikost okna,
H = velikost okna - překryv okna,
t = daný čas.

Dále se z výsledku rovnice 3.1 vyfiltrují nežádoucí frekvence s použitím horní propustě. Tím se odfiltruje působení gravitace. Po filtraci je upraven výsledek tak, aby byl vyjádřen pouze jedním číslem pro jednoduchou změnu pásmové propusti bez nutnosti změny kódu. Každá hodnota pro danou frekvenci byla po výpočtu z rovnice 3.1 ve tvaru komplexního čísla. To bylo dále převedeno na reálné číslo pro lepší manipulaci s daty a umocněno, aby byly všechny hodnoty kladné a zvětšil se rozsah hodnot pro daný povrch. Tím se získala větší citlivost na případné nerovnosti povrchu.

Výsledná reprezentace daného povrchu pak byla suma výsledných reálných čísel. Tento výpočet byl proveden podle následující rovnice:

$$x = \sum (x_{filt}[n]\overline{x_{filt}[n]})^2, \tag{3.2}$$

kde:
x = výsledné reálné číslo pro reprezentaci povrchu,
$\underline{x_{filt}[n]}$ = pole vyfiltrovaných frekvencí po Fourierově transformaci,
$\overline{x_{filt}[n]}$ = pole komplexně sdružených hodnot k $x_{filt}[n]$.

3.2 Úprava fotografií

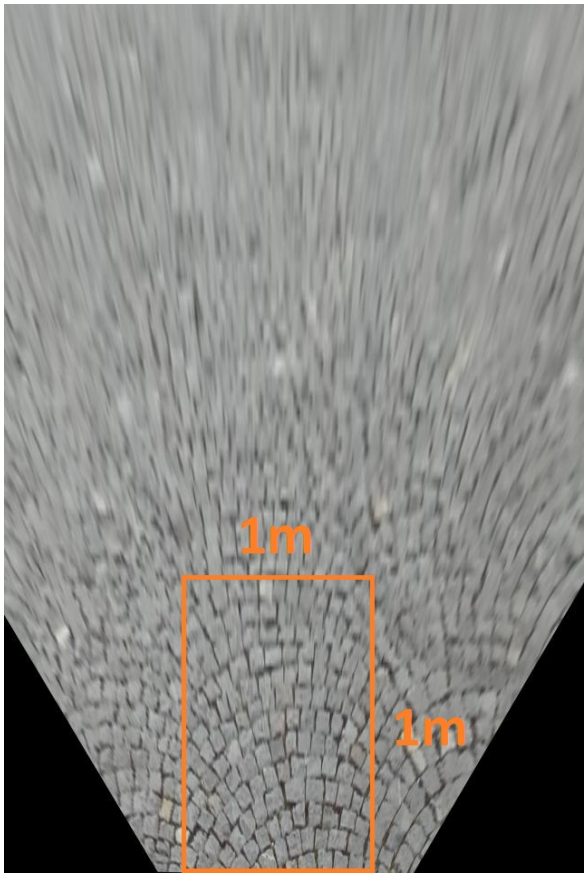
Další kritickou částí bylo zpracování fotografií, které byly následně společně se zpracovanými daty z akcelerometru použity pro učení neuronové sítě. Tento proces se dělil na dvě části. V první řadě bylo nutné fotografii ortorektifikovat. Příklad ortorektifikované fotografie je zobrazen na obrázku 3.5.

Při použití ortorektifikace vznikne obraz pohledu na povrch shora. Tím se umožní použití již natrénovaných neuronových sítí na těchto obrázcích i při jiném umístění kamery. Pokud by se tato transformace neprováděla, byly by naučené sítě závislé jen na jednom umístění kamery pod stejným úhlem.

Následně se vytvořil výřez obrazu z takto upravené fotografie, který odpovídal prostoru před autem, kde se následně konalo učení a predikce povrchu. V této práci se využívalo vždy jen fotografie z levé čočky kamery. Příklad takové fotografie je vidět na obrázku 3.4.

Obrázek 3.4: Jízda po kamenné dlažbě - levá čočka

Na ortorektifikované fotografii je patrné, že s rostoucí vzdáleností se ztrácejí využitelné informace fotografie. Úbytek těchto informací je zapříčiněn nižším rozlišením, což je patrné na obrázku 3.6. V tomto případě při použití platformy TOMI, kdy je kamera relativně nízko nad zemí, odpovídá nerozmazaná část fotografie přibližně vzdálenosti 4m před autem. Tato použitelná vzdálenost by se mnohonásobně zvětšila, kdyby byla kamera ve větší výšce nad zemí. Pro malé modely však toto nepřichází v úvahu, protože by kamera více oscilovala na nerovnostech a tím by se celý obraz rozmazal. TOMI je model auta vytvořený na ČVUT. Detailněji bude tato platforma probrána v kapitole 4.



Obrázek 3.5: Jízda po kamenné dlažbě - ortorektifikovaný obraz

Obrázek 3.6: Jízda po kamenné dlažbě - výřez z ortorektifikovaného obrazu 3.5

Po transformaci bylo nutné vzít v potaz pouze část fotografie a to tu, po které určitě auto projede. Je to nezbytné, protože se fotografie neanotují

ručně, ale využívá se vypočtených dat z akcelerometru. Z tohoto důvodu je velká část fotografie bezcenná pro učení. Výsledný výřez byl vytvořen v bezprostřední blízkosti před autem. Konkrétně to odpovídá délce 1m na délku a 1m na šířku. Tento výřez je zobrazen na obrázku 3.6 [PJ96].

3.3 Předzpracování dat

Kritickou částí této práce bylo předzpracování dat, které se skládalo z již vysvětlených částí obsahujících výpočet veličiny popisující hrubost povrchu a úpravu fotografií. Následně však bylo nutné správné přiřazení veličiny určitému obrázku. Tedy určení správného labelu pro vstup neuronové sítě.

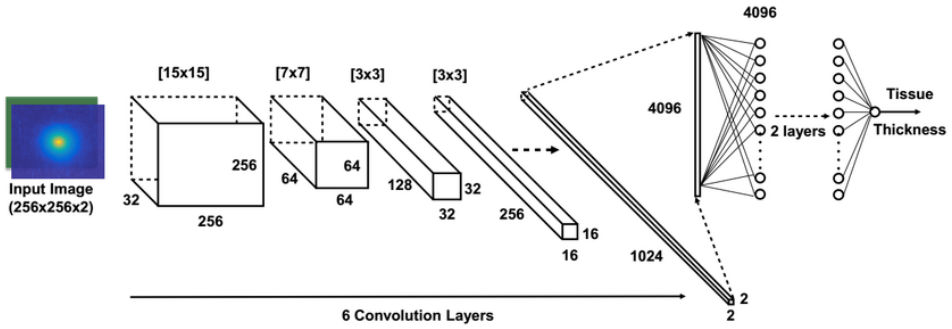
Protože rychlost sběru dat z akcelerometru pomocí Xsens dosahuje 100 Hz a kamera je schopna dosáhnout 30FPS, bylo nutné přiřadit ke každému snímku danou hodnotu a ne naopak. Odhad, kdy auto přejelo povrch zachycený na výřezu fotografie, byl vypočten s využitím znalosti času, kdy byl pořízen snímek a data z akcelerometru. Dále byla použita rychlost auta, která byla získána přepočtem z otáček kol.

Přiřazení vypočtené veličiny tedy bylo predikováno pomocí vypočtení času, kdy auto mělo ujet danou vzdálenost 1m. Jelikož však nemáme vypočtenou veličinu pro každý časový okamžik, bylo nutné najít z časového hlediska nejbližší hodnotu, která se následně použije jako predikovaná hodnota. Příklad určení dané hodnoty je zobrazen na obrázku 3.3, kde právě určená hodnota v čase t_0 odpovídá hrubosti povrchu zobrazeném na výřezu z ortorektifikované fotografie zobrazené na obrázku 3.2.

3.4 Použití neuronové sítě

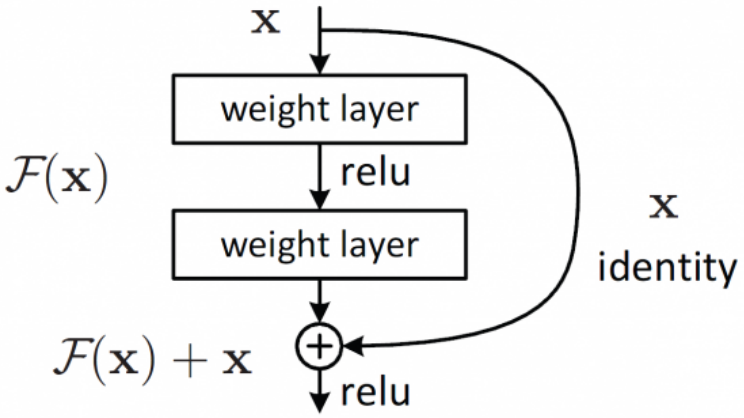
Jako poslední krok bylo naučení konvoluční neuronové sítě pro predikci veličiny, která reprezentuje hrubost povrchu [ZF14]. Příklad sítě pro učení regresní úlohy je zobrazen na obrázku 3.7.

Pro určení hrubosti z obrazu je zapotřebí použít architekturu konvoluční sítě, která bude připravena pro regresní úlohu. Jelikož se jedná o predikci hrubosti, není zde předem určeno, že jeden daný povrch musí být vždy hladký



Obrázek 3.7: Příklad konvoluční sítě pro regresní úlohu [MSE17]

nebo velmi hrubý. U asfaltu je například předpoklad, že tento povrch bude hladký a hodnota predikované hrubosti bude nízká. Musí se však počítat s tím, že v asfaltu se mohou vytvořit výmoly, které pak již hladké nejsou a jsou naopak velmi hrubé. Nedá se tedy předem určit jakých hodnot hrubosti by měl jaký povrch nabývat.



Obrázek 3.8: Příklad učení reziduální konvoluční sítě [HZRS15]

Například pro učící množinu o velikosti 6900 obrázků trvalo učení jedné epochy na jedné grafické kartě 45 vteřin při použití architektury sítě ResNet-50. Validace pak trvala pro 7258 obrázků 26 vteřin. Učení tedy pro 1000 epoch trvalo necelých 20 hodin.

Tato práce se zabývá regresní úlohou, tedy cílem je získat na výstupu neuronové sítě jednu hodnotu reprezentující hrubost povrch. K tomu byla využita architektura ResNet-50 [HZRS15]. Pro učení bylo použito následující nastavení sítě:

- Batch size: 20
- Learning rate: $1e^{-5}$
- Loss funkce: MSELoss
- Aktivační funkce: ReLu

Batch size, neboli počet obrázků použitých pro jedno učení, byl určen experimentálně. Testovaly se hodnoty 10, 20, 30 a 50. Jako nejvhodnější hodnota se nakonec vybrala hodnota 20. Při ostatních hodnotách byla vždy větší odchylka predikovaných hodnot od těch reálných.

Learning rate, tedy hodnota ovlivňující rychlost učení sítě, byl nastaven na $1e^{-5}$. Při experimentech vyšla tato hodnota jako nejlepší z hlediska času učení a výsledné odchylky.

Ztrátová funkce MSELoss je průměrná kvadratická odchylka predikované hodnoty od té reálné.

Aktivační funkce ReLu je jedna z nejvíce používaných funkcí, která všem záporným hodnotám přiřadí 0 a všem kladným hodnotám nechá jejich nezměněnou hodnotu.

Učení probíhalo výhradně na grafických serverech ČVUT, kde jsou k dispozici grafické karty NVIDIA GeForce GTX 1080Ti.

Kapitola 4

Testovací platforma

Pro sběr dat a testování bylo využito modelu auta Toyota Mini (TOMI), na jehož vývoji jsem se podílel jako jeden z autorů a vývoj probíhal v laboratoři Toyoty na ČVUT. TOMI bylo vytvořeno za účelem testování mnoha výzkumných projektů. Detailní popis auta a jeho hardwarová výbava je popsána v [BFP+20].

4.1 Model auta

Jedná se o model Losi Desert Buggy XL-E 1:5 4WD, tedy model je v měřítku 1:5 oproti skutečnému autu a má náhon na všechna 4 kola. Tento model byl vybrán tak, aby splňoval veškeré požadavky pro budoucí vývoj:

- Dostupnost auta a náhradních dílů
- Elektrický motor
- Silný motor z důvodu těžkého hardwaru
- Velký prostor pro hardware
- Přijatelná cena

Jedním z důležitých bodů při výběru byla dostupnost auta a náhradních dílů. Bylo nezbytné, aby byla rychlá dodávka nových dílů a tudíž aby nebyly dlouhé prostoje mezi objednáním a doručením potřebných dílů v případě jejich rozbití. Dále byl požadavek na elektrický motor, což umožňuje jízdu s autem v areálu ČVUT a zároveň v lese či jiných místech, kam modely se spalovacím motorem nemohou. Motor musel být silný, protože bylo potřeba zajistit, aby se auto pohybovalo svižně i s velkým množstvím těžkého hardwaru a aby bylo schopné rychlých rozjezdů a brzdění s prokluzem kol.

Velký prostor byl stěžejní pro prvotní pokusy, kdy se musel pod kapotu vejít notebook. To umožnilo rychlý výběr a test ostatních komponent. Následně byl notebook vyměněn za mini počítače, které tento prostor plně zaplnily.

V neposlední řadě šlo i o cenu modelu a jeho výbavu, se kterou byl prodáván. V tomto případě se jednalo pouze o základní model auta, který byl následně upraven.

Originální neupravené auto je vidět na obrázku 4.1. Tento základní model prošel mnoha vylepšeními z důvodu několika navazujících výzkumným projektů. V případě hardwaru se jedná zejména o přidané senzory, výpočetní techniku a komunikační techniku.

Obrázek 4.1: Losi Desert Buggy XL-E 1:5 4WD, [citc]

4.1.1 Kamera

Jedna z nejdůležitějších komponent pro tuto práci je kamera. Jelikož se jedná o kameru, která měla být přimontována na auto, bylo stěžejní, aby

Obrázek 4.2: ZED kamera, [citg]

byla dostatečně odolná pro případné kolize a především, aby byla schopna pořizovat kvalitní snímky při rychlých pohybech. Tyto parametry splňuje kamera ZED vyobrazena na obrázku 4.2. Kamera byla připevněna ke kovové konstrukci nad otvorem pro čelní sklo. Tato konstrukce byla pevně spojena se střechou auta. Tím se docílilo, že kamera viděla těsně před auto, ale i na větší vzdálenosti.

Pro tuto práci bylo nastaveno rozlišení 2560x720 při 30FPS, kde je v rozlišení zahrnuta fotografie z levé i pravé čočky. Ukázka fotografie je na obrázku 4.3. Na zmíněném obrázku je patrný černý okraj, který byl úmyslně vytvořen přelepením části čoček. Tuto úpravu bylo nutné improvizovaně vyřešit, kdy přelepení čočky fungovalo jako sluneční clona. Bez této úpravy se často stávalo, že byl snímek přesvětlen nebo naopak moc tmavý v závislosti na počasí. Tento problém přetrvával i v případě rychlých rozjezdů, kdy se kamera naklonila a v záběru bylo více světlého nebe. Tím se následně přesvětlil celý obraz. Potíže jsou způsobeny nastavením expozice, která je automaticky nastavena kamerou. Ruční nastavení expozice v tomto případě nepřichází v úvahu, protože z důvodu jízdy za různého počasí nemůže být nastavená hodnota expozice konstantní.

Nejdůležitější částí obrázku je povrch vozovky, odpovídající dolní polovině fotografie, jak je vidět na obrázku 4.3. Z tohoto důvodu se tedy zanedbalo použití horní poloviny fotografie právě přelepením části čoček. Tím se následně vyřešily časté problémy s nastavením expozice. Toto řešení bylo improvizované z důvodu ušetření velkého množství času, které by se jinak musel vynaložit na výrobu krytu nebo jiných řešení.

Obrázek 4.3: Ukázka fotky z kamery ZED při konstantní jízdě TOMI

4.1.2 Výpočetní zařízení

Jako hlavní výpočetní zařízení byl zvolen Jetson AGX Xavier, který byl vybrán především pro jeho velký výkon nezbytný pro klasifikace s použitím neuronových sítí s těmito vlastnostmi:

- Procesor (CPU)
  - 8-core Carmel ARM v8.2 64-bit CPU, 8MB L2 + 4MB L3
- Grafický procesor (GPU)
  - 512-core Volta GPU with 64 Tensor Cores
  - 11 TFLOPS (FP16)
  - 22 TOPS (INT8)
- Maximální příkon
  - 30 W

Tento počítač je hlavní komponentou celé sestavy. Běží s operačním systémem Linux Ubuntu 18.04 a je zobrazen na obrázku 4.4. Ukládají se zde veškerá nasbíraná data na externí SSD disk. Tato data se sbírají přímo ze senzorů připojených přes USB 3.0 a Jetson zároveň přijímá data poslaná přes sériovou linku RS232 od Raspberry Pi 2.

Obrázek 4.4: Jetson AGX Xavier, [citb]

Jako další komponenta bylo vybráno Raspberry Pi 2 s modulem Navio 2 zobrazené na obrázku 4.5. Tento počítač byl vybrán pro výčet jednotlivých senzorů a především pro komunikaci s kontrolérem motoru. Hlavním úkolem

tohoto počítače je komunikace s Arduinem, kde se posílají PWM signály z přijímače pro ovládání plynu a natočení kol. Dále se zde sbírají data například z GPS, IMU a v neposlední řadě i informace o rychlosti auta v RPM.

Obrázek 4.5: Raspberry Pi 2 s modulem Navio 2, [citd]

4.1.3 Inerciální měřicí jednotka (IMU)

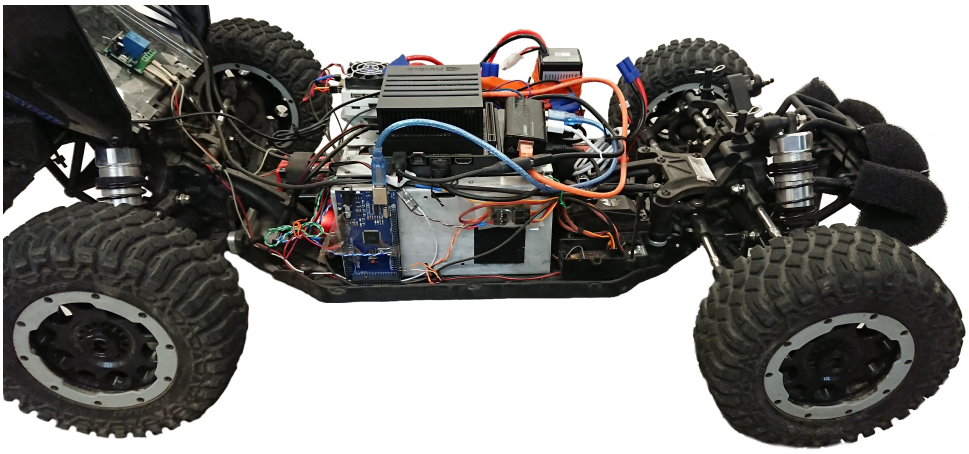
Jedna z klíčových meřených veličin pro tuto práci byla akcelerace v ose z. Proto bylo stěžejní vybrat kvalitní IMU a následně její montáž na konstrukci auta tak, aby nedocházelo k žádnému rušení. V prvé řadě se použila IMU jednotka integrovaná v modulu Navio 2, který byl pevně spojen s Raspberry a namontován na společné kovové konstrukci s Jetsonem. Po mnoha pokusech se však zjistilo, že jsou při roztočeném motoru data z akcelerometru značně zašuměná a jsou tudíž pro další zkoumání nepoužitelná. Problém byl ve vibracích, které vytvářel roztočený motor na připevněné konstrukci. Bohužel se Raspberry s modulem nemohlo kvůli své velikosti přemístit na jiné místo v autě. Z tohoto důvodu se pro sběr dat místo Raspberry vybrala jednotka Xsens MTi-7-DK zobrazená na obrázku 4.6. Tato jednotka se přimontovala přímo k podvozku auta a tím se vyřešily problémy s mechanickými vibracemi od motoru. Jednotka Xsens byla připojena přes USB k výpočetní jednotce Jetson, odkud probíhalo vyčítání dat. Naměřená data pak byla uložena v Jetsonu.

Obrázek 4.6: Inerciální měřicí jednotka Xsens MTi-7-DK, [citf]

Obrázek 4.7: TOMI platforma z boku

4.1.4 Výsledná TOMI platforma

Výsledná TOMI, zobrazená na obrázku 4.7, obsahuje již výše zmíněné komponenty, které jsou stěžejní pro tuto práci. V rámci vývoje TOMI v Toyota laboratoři byl tedy uzamčen diferenciál mezi přední a zadní nápravou a byl vytvořen záchranný systém, který vypnul motor v případě technických potíží. Dále byly vybrány a namontovány do auta mini počítače Jetson a Raspberry na kovovou konstrukci pro sběr dat ze senzorů, kamera pro různá rozpoznávání scény pře autem, další baterie pro napájení celé elektroniky kromě původní výbavy, GPS pro přesnou lokalizaci a IMU pro získání dat o pohybu auta. Vnitřek auta s veškerou výbavou je pak vidět na obrázku 4.8.



Obrázek 4.8: TOMI platforma s namontovanými komponentami

Oproti původně zakoupené platformě auta byly tedy přidány následující komponenty:

- Jetson AGX Xavier
- Raspberry Pi 2 s modulem Navio 2
- Arduino Mega
- Kamera ZED
- Baterie
- USB rozbočovač
- GPS
- IMU Xsens MTi-7-DK
- Záchranný systém

Kapitola 5

Experimenty

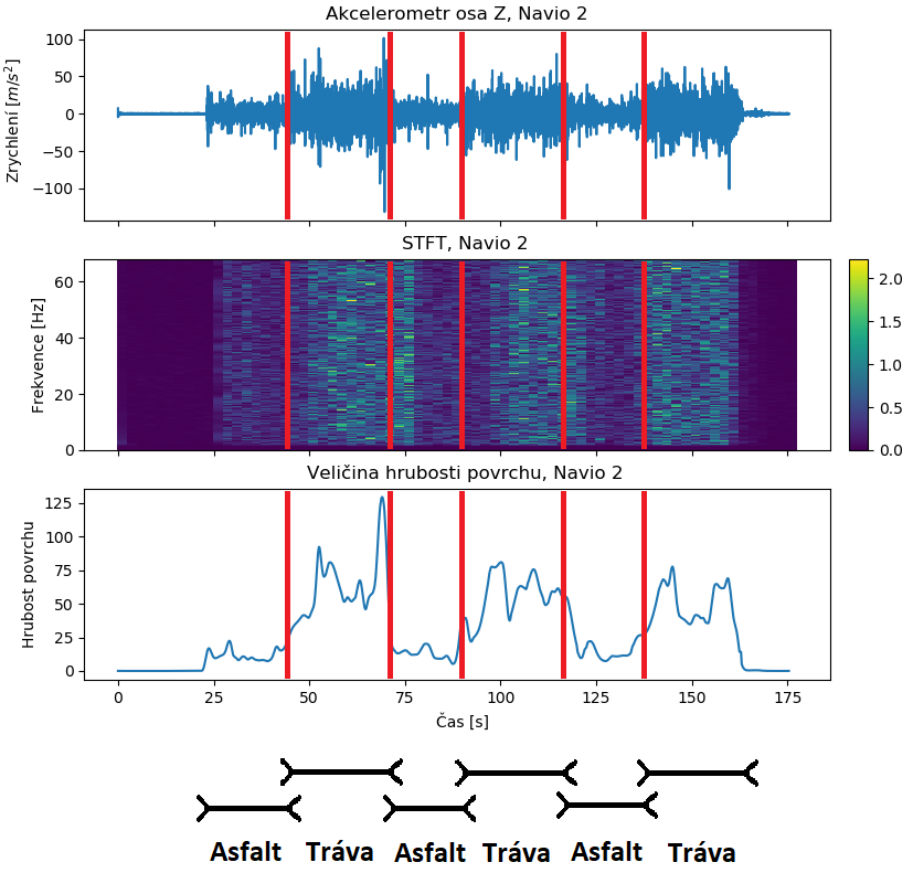
Pro všechny experimenty v této práci byla použita platforma Toyota Min, zkráceně TOMI. Pouze při testování IMU jednotek byl využit i osobní automobil. Provedené experimenty obnášely vždy sběr dat s TOMI, následné předzpracování nasbíraných dat, ruční analýzu dat a nakonec samotné učení neuronových sítí.

5.1 Použití dat z akcelerometru

Pro sběr dat bylo využito modelu RC auta, které obsahovalo IMU (inerciální měřicí jednotka) obsaženou v modulech Navio 2 a Xsens. Navio 2 je přídavný modul pro mini počítač Raspberry Pi. Komponenta Xsens MTi-7-DK je pak modul specializovaný na sběr dat z akcelerometru a gyroskopu. Detailnější popis použité platformy RC auta, jednotek Navio 2 a Xsens je rozepsán v kapitole 4.

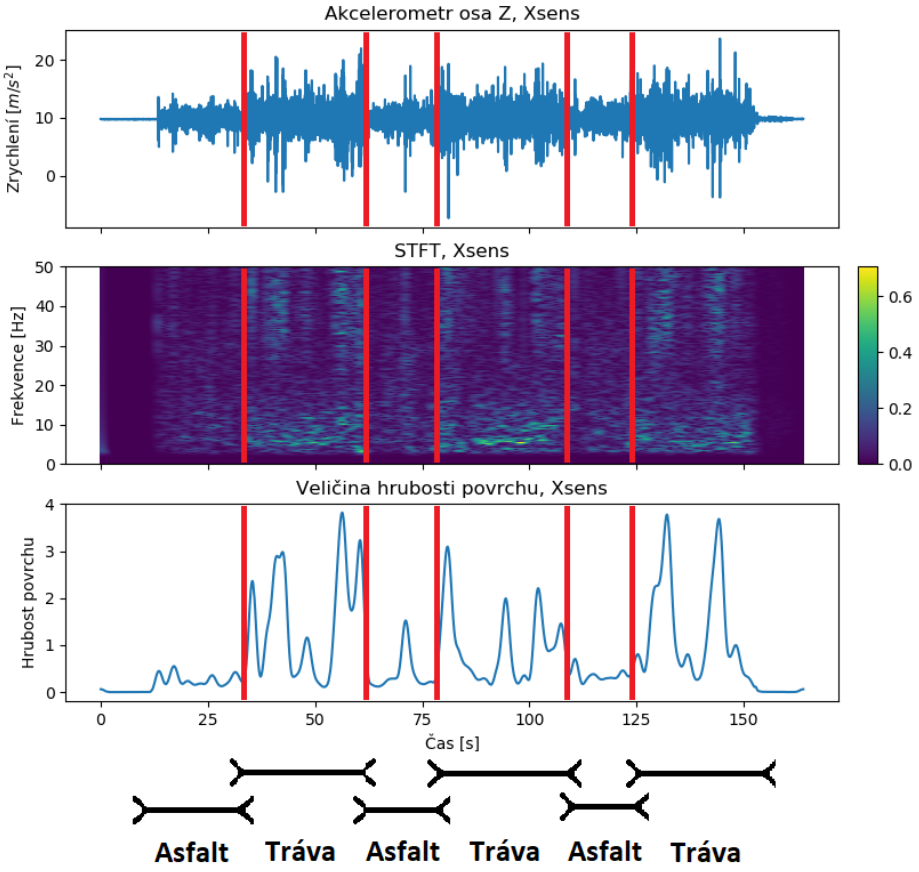
Jako hlavní předpoklad pro realizaci této práce bylo použití IMU a následná analýza naměřených dat z akcelerometru [BI05]. V prvé řadě bylo použito integrované IMU jednotky v modulu Navio 2. V tomto kroce bylo naměřeno mnoho dat s různými povrchy pro pozdější analýzu, ve které se hledala souvislost mezi drsností povrchu a daty z akcelerometru [ID02]. Konkrétně osy z, která je k povrchu kolmá. V této práci byly vybrány 2 jízdy, které byly vždy zaznamenány jak pomocí Navia, tak i Xsens IMU jednotek.

První jízda byla provedena na 2 površích. Jednalo se o asfalt a o trávu. U asfaltu byl předpoklad, že na akcelerometru nebudou velké oscilace, protože je tento povrch velmi hladký. U trávy se naopak předpokládaly větší výchylky z důvodu hrubšího povrchu. Data z jízdy naměřené na jednotce Navio jsou na obrázku 5.1. Naměřená data pomocí jednotky Xsens z jízdy po stejném povrchu jsou pak vidět na obrázku 5.2.



Obrázek 5.1: Jízda střídavě po asfaltu a trávě - Navio 2

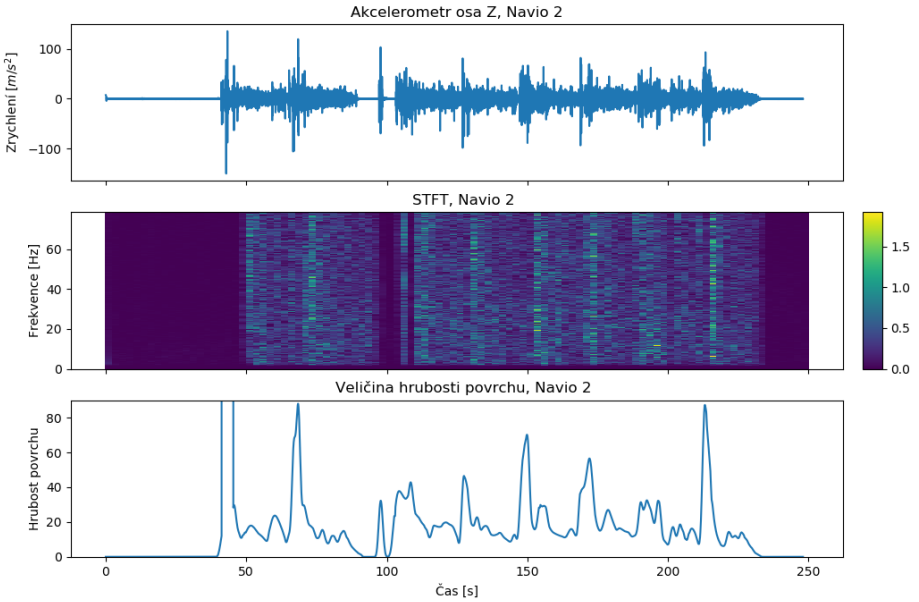
Na obrázku 5.1 je již na první pohled zarážející hodnota zrychlení. V časovém intervalu do 25 sekund, kdy je auto v klidu, je vše v pořádku a hodnota zrychlení tak dosahuje hodnoty gravitačního zrychlení. Bohužel však od 25 sekund dále jsou patrné velké oscilace, které dosahují hodnot až $\pm 100 m/s^2$. Tyto hodnoty jsou patrné od chvíle, kdy se auto začne pohybovat. Po mnoha testech, které zahrnovaly například i možnost elektromagnetického rušení od motoru, se nakonec ukázalo, že běžící motor rozvibruje kovovou konstrukci natolik, že se pak při měření kvůli jednotce Navio dostáváme na již zmíněné hodnoty. Z tohoto důvodu byla na auto namontována jednotka Xsens, která se vešla přímo na podvozek a tím se eliminovaly vibrace způsobené motorem zobrazené na obrázku 5.2.



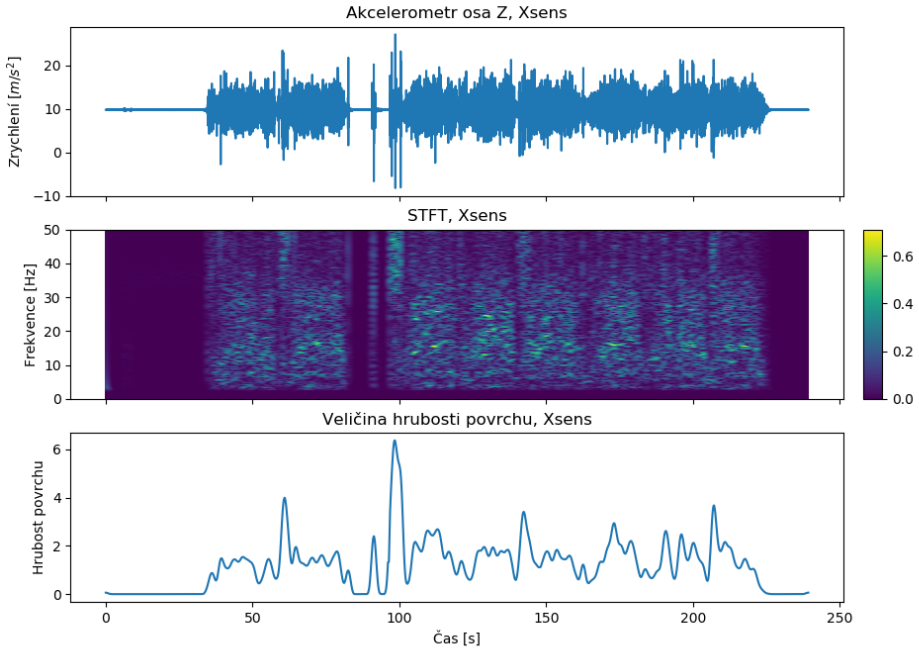
Obrázek 5.2: Jízda střídavě po asfaltu a trávě - Xsens

V časové oblasti však není jasně rozeznatelné, kdy jede auto po asfaltu a kdy po trávě. Proto bylo nutné převést signál do frekvenční oblasti s použitím STFT. Po transformaci se může signál zkoumat z hlediska obsažených frekvencí. Následně se tyto informace dají využít při zpracování signálu [Wel67]. Pro měření pomocí Navia je signál po transformaci zobrazen na obrázku 5.1.

Zde je celkem zřetelné, že v časových intervalech 25s - 40s, 70s - 90s a 115 - 135s je mnohem menší amplituda na všech frekvencích, což by odpovídalo hladkému povrchu. V tomto případě asfaltu. Naopak v intervalech 40s - 70s, 90s - 115s a 135s - 160s je vidět mnohem větší četnost frekvencí v signálu a tudíž se jedná o hrubší povrch. Zde se jedná o jízdu po trávě. Je však vidět, že nelze jednoznačně určit jaké frekvence v signálu převládají a pro asfalt nebo trávu jsou frekvence téměř rovnoměrně rozložené až do 70Hz.



Obrázek 5.3: Jízda po kamenné dlažbě - Navio 2



Obrázek 5.4: Jízda po kamenné dlažbě - Xsens

Oproti tomu ve zpracovaném signálu z Xsens, který je zobrazen na obrázku

5.2, není na první pohled až tak zřetelné, které části patří asfaltu a které trávě. Je zde však vidět, že v signálu převládají frekvence 5Hz - 15Hz.

Další jízdy byly provedeny na kamenné dlažbě. Tedy předpokladem je, že při konstantní rychlosti bude v signálu převládat relativně úzké pásmo frekvencí. Tento předpoklad si zdůvodníme tím, že kostky v dlažbě jsou téměř totožné a spáry mezi nimi jsou tedy vždy ve stejné vzdálenosti od sebe. Při konstantní rychlosti bychom pak v ideálním případě dostali silný signál na frekvenci odpovídající přejezdu přes spáry. Naměřená data na jednotce Navio jsou zobrazena na obrázku 5.3 a data nasbíraná na jednotce Xsens na obrázku 5.4.

Z dat naměřených pomocí Navia, zobrazených na obrázku 5.3, by se v časové oblasti předpokládalo, že auto přejelo přes několik velkých nerovností, které způsobily velké výchylky v datech z akcelerometru z osy z. Auto však jelo konstantní rychlostí po kamenné dlažbě a proto jsou tyto oscilace v našem případě velmi nežádoucí.

U dat nasbíraných pomocí Xsens se také jednalo o jízdu po kamenné dlažbě. Vypočtené hodnoty hrubosti povrchu až na výjimku v čase 100 sekund, kdy se auto po krátkém zastavení opět rozjelo, se pohybují ve stejně velkém intervalu. Tento signál je zobrazen na obrázku 5.4.

Ve frekvenční oblasti se z dat z Navia na obrázku 5.3 tentokrát nedá určit ani převládající frekvence v signálu a také to vypadá, že oproti první jízdě zobrazené na obrázku 5.1 je povrch druhé jízdy méně hrbolatý.

Naopak z dat nasbíraných pomocí Xsense, zobrazených na obrázku 5.4, je hned na první pohled patrné, že převládají frekvence mezi 10Hz a 30Hz. Tudíž je patrné, že se opravdu jedná o jeden povrch, ve kterém jsou všechny nerovnosti přibližně stejně vzdálené od sebe, a to odpovídá kamenné dlažbě.

Z uvedených obrázků grafů dat z Navia a Xsens je tedy patrné, že u Navia je zřetelné velké rušení, které je způsobeno vibracemi kovové konstrukce, na které je Navio přimontované, během motoru. Z tohoto důvodu jsou v datech velké oscilace. Při frekvenční analýze dat pak nelze jednoznačně odlišit jízdu po trávě od jízdy po kamenné dlažbě. Naopak u dat naměřených pomocí Xsens se dá rozlišit jak jízda po trávě oproti jízdě po kamenné dlažbě, tak i jízda po asfaltu. Z tohoto důvodu byla místo Navia použita jednotka Xsens pro sběr dat, která se následně zpracovávají.

5.2 Sběr dat

Fyzicky nejobtížnější a zároveň i nejkritičtější částí v této práci byl sběr pokud možno co nejideálnějších dat pro následné zpracování. Jedna nepřetržitá jízda mohla trvat maximálně 20 minut z důvodu kapacity akumulátorů. Zároveň bylo důležité, aby byla jízda plynulá, ideálně s konstantní rychlostí a bez častého zatáčení.

Rychlost nemohla být moc malá, ale ani moc vysoká. Při malé rychlosti by se některé nerovnosti nemusely zobrazit ve frekvenčním spektru a zároveň by se mohlo stát, že auto nebude schopné vyjet potřebný kopec. Naopak při větších rychlostech by auto mohlo nerovnosti, například spáry, jednoduše přeskočit a vůbec je nezaznamenat.

Také bylo nezbytné, aby auto jelo pokud možno konstantní rychlostí. Pokud by měnilo výrazně svoji rychlost, ovlivnilo by to výslednou veličinu reprezentující hrubost povrchu. Ta je totiž závislá i na rychlosti auta [ea06].

Aby bylo možné co nejlépe identifikovat hodnoty, které patří výřezu ortorektifikované fotografie, bylo potřeba, aby byla jízda bez zbytečných zatáčení. Při každém zatočení se může stát, že předzpracované hodnoty, které by měly odpovídat výřezu fotografie, budou ve skutečnosti pro úplně odlišný povrch i hrubost povrchu. Tím se pak již do učení neuronové sítě vkládají nepravdivé hodnoty považované za ty správné.

Pro co nejobecnější využití naučené sítě je zapotřebí nasbírat data pro daný povrch za různého počasí a denní doby. Proto probíhal sběr dat při slunečném dni, když bylo zataženo, v dopoledních hodinách, večerních hodinách a dokonce i po dešti.

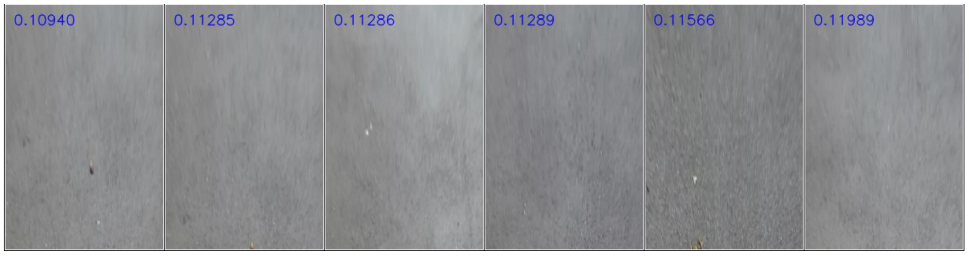
5.3 Datasety

Dataset, neboli soubory dat pro následné využití v učení neuronové sítě byl vytvořen z jedné i více jízd. Vytvoření jednotlivých datasetů bylo nedílnou součástí této práce a bylo velmi důležité, aby ke každému datasetu použitému pro učení existoval i dataset pro validaci. V každém z nich musel být stejný povrch, jinak by výsledky nebyly směrodatné.

Mezi první experimenty patřily relativně malé datasety s počtem obrázků řádově od pár stovek až po pár tisíc. Tato data byla nasbírána převážně na kamenné dlažbě, trávě, panelové cestě asfaltu a lesní cestě. Tyto datasety byly vytvořeny převážně pro manuální analýzu a zjištění, zda je výsledná veličina popisující hrubost povrchu vhodná pro další učení. Zejména se jednalo o hodnotu, která jednoznačně určí nejen hrubost povrchu, ale právě na základě hrubosti by bylo možné určit i daný typ povrchu. To by bylo možné ale jen v případě, že na daném povrchu nebudou nerovnosti, které nejsou pro daný povrch specifické.

Příklady jednotlivých povrchů a vypočtených hodnot jsou vidět na obrázcích 5.5, 5.6, 5.7 a 5.8. Modře vepsaná hodnota v obrázcích vždy určuje vypočtenou hodnotu veličiny popisující hrubost povrchu. Rychlost auta přitom byla konstantní a dosahovala hodnoty 5 km/h.

Předzpracovaná data byla seřazena podle vypočtené veličiny od nejmenších hodnot po největší. Jako nejhladší, tedy co nejméně hrbolatý, povrch byl určen asfalt. Vypočtené hodnoty pro asfalt nabývaly hodnot od 0,06 do 0,51. Jsou zde započítány i krajní hodnoty, které jsou zapříčiněny například velkou spárou v asfaltu. Příklady předzpracovaných dat s asfaltem jsou vidět na obrázku 5.5.



Obrázek 5.5: Asfalt s hodnotou veličiny popisující hrubost povrchu

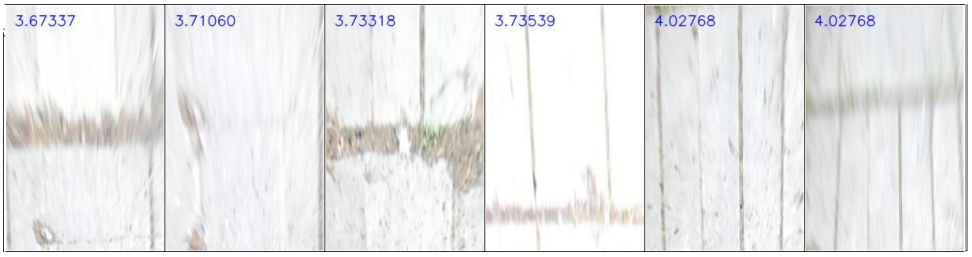
Dalším povrchem je travnatá plocha. Ta nabývala hodnot od 0,17 do 1,18 a příklady jsou zobrazeny na obrázku 5.6. Opět jsou zde zahrnuty krajní hodnoty, které mohly být způsobeny například přejezdem z hladkého povrchu na trávu nebo zpomalením v důsledku vjetí do trávy.

Obrázek 5.6: Tráva s hodnotou veličiny popisující hrubost povrchu

Předposledním testovaným povrchem v tomto datasetu byla lesní cesta. Ta byla druhá nejhrbolatější a nabývala tak hodnot od 1,11 do 2. Příklady jsou zobrazeny na obrázku 5.7.

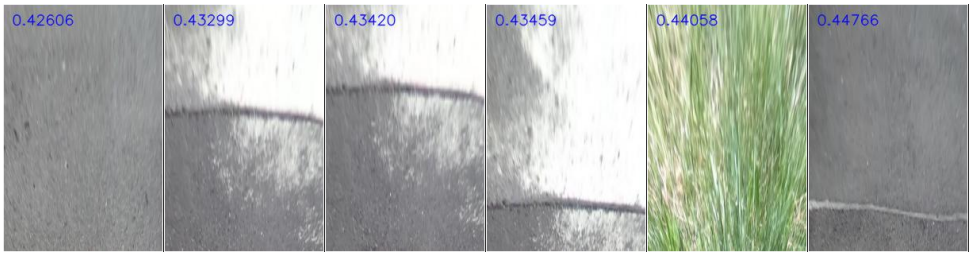
Obrázek 5.7: Lesní cesta s hodnotou veličiny popisující hrubost povrchu

Posledním a to nejhrbolatějším povrchem v tomto datasetu byla panelová cesta. Zde se pohybovaly hodnoty od 1,4 až do 6,1. Jde tedy o velmi velký interval hodnot, kterých tento povrch nabývá. Vysoké hrubosti nabývajících hodnot kolem 6 odpovídají v tomto případě přejezdu mezer mezi jednotlivými panely.



Obrázek 5.8: Panelová cesta s hodnotou veličiny popisující hrubost povrchu

Povrchy se zobrazenými hodnotami na obrázcích 5.5, 5.6, 5.7, 5.8 jsou klasické příklady pro tyto povrchy. V této práci však nejde pouze o predikci hrubosti, které odpovídají daným typům povrchů. Jde také o predikci nečekaných hrubostí na povrchu, který by měl být jinak hladký. Příkladem je asfalt, který by měl být ve většině případů hladký. Bohužel se ale stává, že se v asfaltu vytvoří výmoly, které je potřeba identifikovat a vyhnout se jim. Tento příklad je zobrazen na obrázku 5.9.



Obrázek 5.9: Asfalt s výmoly a tráva s hodnotou veličiny popisující hrubost povrchu

V dalším experimentu byla projeta trasa obsahující asfalt, panelovou cestu

a lesní cestu. Vždy se dbalo na dodržení konstantní rychlosti a co nejmenších změn směru jízdy. Jelikož vedla trasa skrz les, vyskytovalo se na cestě mnoho stínů a zároveň velmi mnoho světlých míst z důvodu prosvítajících slunečních paprsků skrz koruny stromů. Takto nasbíraná data se následně náhodně rozdělila na trénovací a testovací množinu. Tím se však mohlo stát, že budou dva sobě velmi podobné snímky jak v trénovací, tak i validační množině. Také nebylo zaručeno, že bude vždy dostatečně velký počet vzorků reprezentující daný povrch v obou množinách. Jako řešení tohoto problému by se daly vybírat fotografie, které by měly mezi sebou dostatečně velký časový rozestup, nebo nasbírat data z více jízd. Jedna jízda by se pak dala využít pro učení a druhá pro validaci.

Mezi poslední experimenty patří dataset pro učení z jízdy na jedné trase. Tato trasa je vidět na obrázku mapy 5.10 jako červeně vyznačená trasa. Jelikož bylo ale potřeba, aby nebyly stejné obrázky v učící i validační množině zároveň, jako tomu bylo v prvním experimentu, byla tato trasa projeta dvakrát. Pro učení se použila trasa projetá v protisměru hodinových ručiček a pro validaci po směru hodinových ručiček. Tím bylo zajištěno, že v obou množinách bude vždy dostatečné množství počtu fotografií daného povrchu. Zároveň tím, že byla trasa projeta dvakrát, se nemůže stát, že by byly jakékoli dva snímky téměř stejné. Každý z těchto dvou datasetů obsahoval přibližně 30000 obrázků. Jedná se tedy již o celkem velký dataset. Nasbíraná data s obrázky před zpracováním mají dohromady velikost 26 GB. To odpovídá velikosti záznamu 15 minut jízdy a délce trasy 1,8 km.

Obrázek 5.10: Mapa projeté trasy, [cita]

Z výše popsaných datasetů je tedy zřejmé, že bylo nasbíráno velké množství různorodých dat. Je také vidět, že se vypočtené hodnoty pro popis jednotlivých povrchů prolínají. To ukazuje na potřebu znalosti hrubosti povrchu před autem, například asfaltu, v případě, že je zde výmol, kterému je žádoucí se vyhnout.

5.4 Referenční metoda

Referenční metodou bylo v této práci použití obrazového gradientu místo využití neuronových sítí a dat z akcelerometru. Tato metoda byla vybrána kvůli její jednoduchosti a rychlosti na výpočet. Předpokladem je, že hladký povrch nebude mít silnou texturu. Hrubý povrch bude mít naopak velmi silnou texturu. Hrubší povrch by tedy měl mít mnohem větší hodnotu obrazového gradientu než hladký povrch. Výhodou je také to, že není potřeba učení klasifikátoru.

Výpočet průměrného gradientu $\overline{G}$ pro obraz probíhá podle následujícího vzorce [PYK08]:

$$\overline{G} = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \sqrt{\frac{\Delta I_x^2 + \Delta I_y^2}{2}}, \tag{5.1}$$

$$\Delta I_x = X(i+1, j) - X(i, j), \tag{5.2}$$

$$\Delta I_y = X(i, j+1) - X(i, j), \tag{5.3}$$

kde:
X = obrázek, pro který je počítán gradient,
ΔI_x = horizontální gradient pro každý pixel,
ΔI_y = vertikální gradient pro každý pixel.

Ačkoli skýtá tato metoda velké výhody v oblasti výpočetní náročnosti, její nevýhody jsou hlavně v neznalosti predikované fotografie. Není si tedy schopna spojit určitou barvu a tvary s možnou predikovanou hodnotou.

Jak je již patrné z názvu metody, využívá obrazový gradient k výpočtu hrubosti povrchu, jak je ukázáno v rovnici 5.1. Výpočet tedy využívá kontrastu mezi jednotlivými prvky v obraze. Čím jsou prvky kontrastnější, tím větší bude vypočtená hodnota. Tato metoda se využívá mimo jiné i při detekci hran v obraze.

Příklad využití této metody pro rozpoznání hrubosti povrchu je vidět na obrázku 5.11. Na tomto obrázku je zobrazen asfalt po výřezu z ortorektifikované fotografie a vypočtený obrazový gradient pro každý pixel této fotografie. Na tomto obrázku je vidět, že zde není moc textury, a proto nabývá obrazový gradient převážně nízkých hodnot. Naopak na obrázku 5.12 je příklad ortorektifikované fotografie lesní cesty a její obrazový gradient. Zde je již patrné, že oproti obrázku 5.11 nabývá obrazový gradient vyšších hodnot. To je způsobeno silnější texturou, která je obsažena v obrázku.

Z této fotografie, která vizualizuje obrazový gradient fotografie pořízené z auta, se následně vypočetla koncová hodnota pro popis hrubosti povrchu. Tento výpočet probíhal jako průměrná hodnota všech pixelů a následně se tato hodnota vynásobila koeficientem 0,057. Tímto se výsledná hodnota zmenšila, aby odpovídala skutečným hodnotám vypočteným v kapitole 3.1. Pro tento konkrétní příklad pak vyšla hodnota vypočtená pomocí obrazového gradientu 1.322 a reálná hodnota byla 0.142. Je tedy vidět, že odhad pomocí obrazového gradientu označil povrch jako hrubší, než ve skutečnosti je. Tato hodnota by spíše odpovídala hrubosti lesní cesty, jak již bylo zobrazeno na obrázku 5.7.

Pro porovnání výsledků predikcí bylo využito korelačního koeficientu ρ. Tento koeficient může nabývat hodnot od -1 do 1. Hodnota -1 znamená velmi silnou zápornou lineární kroelaci. Pro hodnotu 0 to znamená, že mezi daty není lineární korelace a hodnota 1 znamená velmi silnou kladnou lineární korelaci.

Pro ohodnocení jednotlivých korelačních koeficientů bylo využito následujícího rozdělení [Eva96]:

- Velmi slabý: 0,00 - 0,19
- Slabý: 0,20 - 0,39
- Střední: 0,40 - 0,59
- Silný: 0,60 - 0,79
- Velmi silný: 0,80 - 1,00

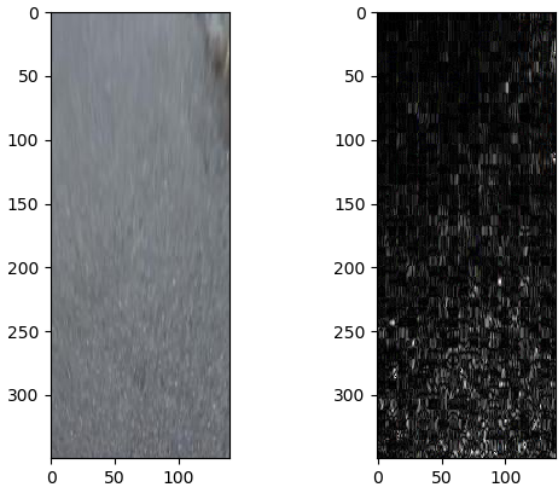
Jednotlivé korelační koeficienty byly vypočteny pro jednotlivé datasety. Dataset 1 obsahoval 425 obrázků. Dataset 2 obsahoval 33858 obrázků. Dataset 3 obsahoval 39678 obrázků. Dataset 4 obsahoval 8666 obrázků. Datasety 1, 2 a 3 obsahovaly následující typy povrchů: asfalt, trávu, lesní cestu a panelovou

	Korelační koeficient ρ
Dataset 1	0,282
Dataset 2	0,126
Dataset 3	0,012
Dataset 4	**0,294**

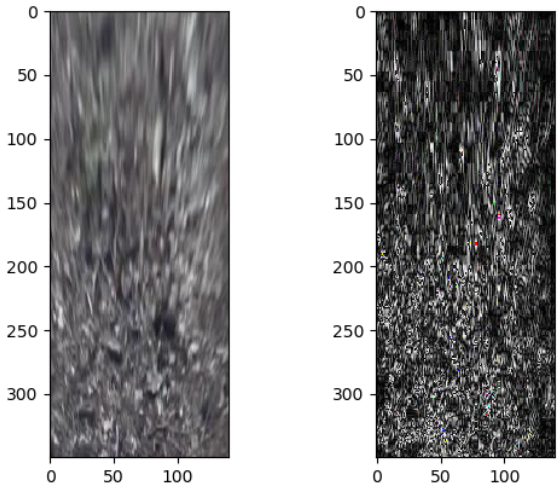
Tabulka 5.1: Korelační koeficienty obrazového gradientu pro jednotlivé datasety

cestu. Dataset 4 pak obsahoval asfalt, trávu a kamennou dlažbu. Výsledky jsou vidět v tabulce 5.1.

Z tabulky 5.1 je vidět, že nejlepší predikce pomocí obrazového gradientu byla pro dataset 4 a téměř stejné korelační hodnoty nabývala predikce na datasetu 3. Tyto dva datasety nabývaly slabé kladné korelace. Naopak u datasetu 2 a 3 můžeme říci, že se jedná o velmi slabou kladnou korelaci.



Obrázek 5.11: Ortorektifikovaný obraz asfaltu a jeho obrazový gradient



Obrázek 5.12: Ortorektifikovaný obraz lesní cesty a jeho obrazový gradient

5.5 Navrhovaná metoda

Tato metoda využila nasbíraných dat z akcelerometru, které byly převedeny do frekvenční oblasti s použitím STFT. Velikost okna pro STFT byla zvolena 5 vteřin. Následně byla ke každému obrázku přiřazena časově odpovídající hodnota. Takto předzpracovaná data se následně použila jako vstup pro neuronovou síť. Tento postup byl popsán v kapitole 3.

Pro porovnání výsledků predikce pomocí navrhované metody oproti reálným datům byla spočtena korelace mezi těmito daty. Dataset 1 obsahoval 425 obrázků. Dataset 2 obsahoval 33858 obrázků. U datasetu 1 a 2 byly obrázky rozděleny napůl na trénovací a validační množinu dat. Dataset 3 obsahoval 38675 trénovacích a 39678 validačních obrázků. Dataset 4 obsahoval 9844 trénovacích a 8666 validačních obrázků. Datasety 1, 2 a 3 obsahovaly typy povrchů asfalt, trávu, lesní cestu a panelovou cestu. Dataset 4 pak obsahoval asfalt, trávu a kamennou dlažbu. Výsledky jsou vidět v tabulce 5.2.

	Korelační koeficienty ρ	
	pro trénovací data	pro validační data
Dataset 1	**0,982**	**0,921**
Dataset 2	0,963	0,903
Dataset 3	0,484	0,050
Dataset 4	0,929	0,478

Tabulka 5.2: Korelační koeficienty výstupu neuronové sítě pro jednotlivé datasety

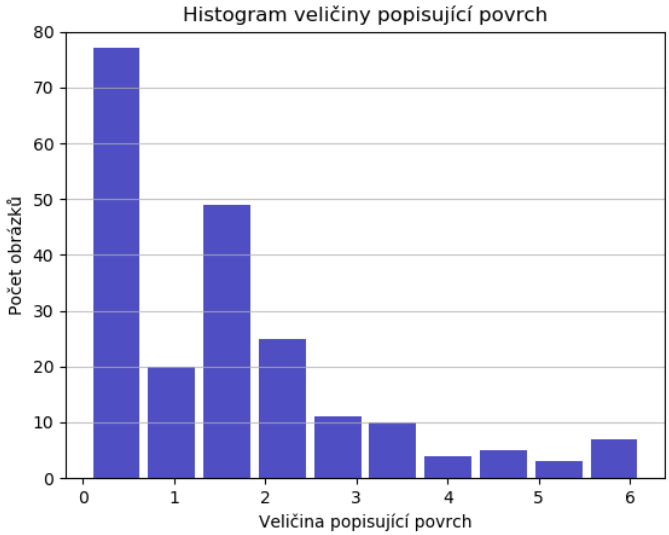
Z tabulky 5.2 je vidět, že nejlepší predikce pomocí navrhované metody byla pro dataset 1, kde byla velmi silná kladná korelace mezi predikovanými a reálnými hodnotami. Velmi silné kladné korelace bylo dosaženo i u datasetu 2. U datasetu 4 pak byla velmi silná kladná korelace pro data z trénovací množiny a střední kladná korelace pro data validační. Naopak u datasetu 3 byla sice středně kladná korelace na trénovací množině, ale na validačních datech byla kladná korelace velmi slabá.

5.6 Detailní analýza výsledků

Trénování a validace probíhaly v prostředí PyTorch a byla použita architektura ResNet-50. Více informací o použitém nastavení je popsáno v kapitole 3.4.

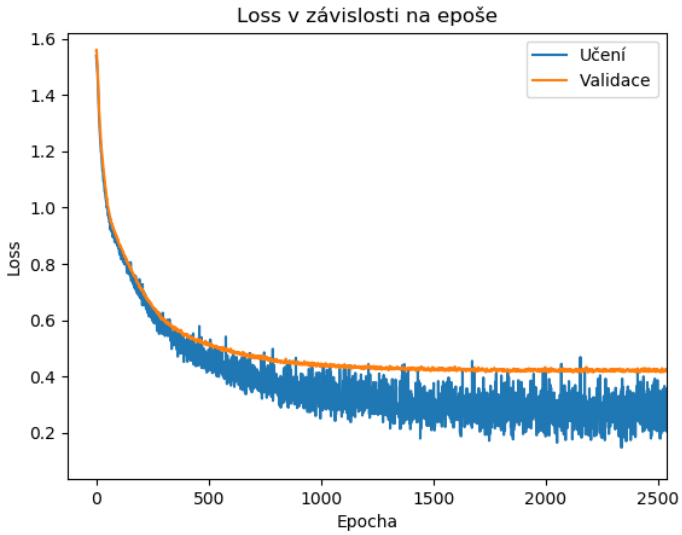
V této části práce bude detailně popsáno a zobrazeno na grafech rozložení veličiny popisující hrubost povrchu v daných datasetech, průběh učení neuronové sítě zobrazený na grafech s hodnotami z výstupu loss funkce a v neposlední řadě výsledky predikcí na učících a validačních datech.

Jako první bude rozebrán dataset 1. Tento dataset obsahoval nejméně obrázků a to 425. Polovina těchto obrázků byla použita na učení a polovina na validaci. Četnost jednotlivých hodnot pro popis hrubosti povrchu je zobrazena na obázku 5.13.



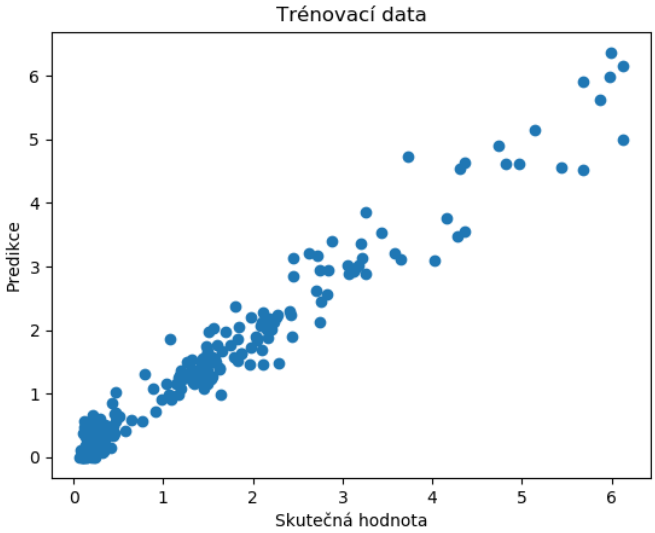
Obrázek 5.13: Dataset 1 - Histogram rozložení hrubosti povrchů

Pro tento dataset můžeme vidět hodnoty loss funkce v závislosti na epoše na obrázku 5.14. Zde je vidět, že hodnota loss funkce při učení dosáhne 0,2 a při validaci hodnoty 0,5.



Obrázek 5.14: Dataset 1 - Hodnoty MSELoss pro dané epochy

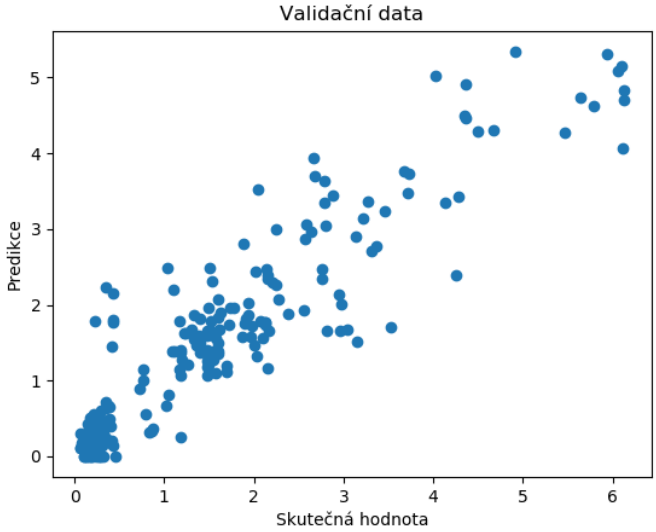
Na obrázku 5.15 jsou zakreslena data z výstupu neuronové sítě při trénování a skutečná data. Z tohoto grafu je vidět, že pro trénovací data je síť schopna naučit se správně predikovat hrubost povrchu.



Obrázek 5.15: Dataset 1 - Predikce na trénovacích datech, $\rho = 0{,}982$

Na obrázku 5.16 jsou zakreslena data z výstupu neuronové sítě při validaci a skutečná data. Zde je opět patrný trend, jako tomu bylo u trénovacích dat,

avšak výsledky nejsou predikované s tak velkou přesností, jako tomu bylo právě u trénovacích dat.



Obrázek 5.16: Dataset 1 - Predikce na validačních datech, $\rho = 0{,}921$

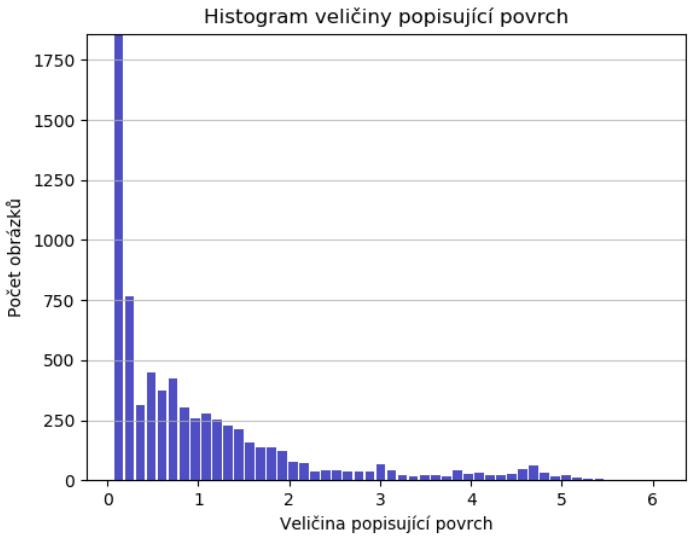
U tohoto datasetu se neuronová síť dokázala vcelku přesně naučit a následně predikovat hrubost povrchu. Bylo zde však velmi málo vstupních dat, to znamená, že s největší pravděpodobností by predikce na datech stejných povrchů, ale z jiných míst nebyla tak přesná. Tento dataset byl navíc rozpůlen na testovací a validační data, tím se mohlo stát, že se v obou množinách vyskytly velmi podobné obrázky.

Následně byl nasbírán dataset 2. V tomto datasetu se již nasbíralo mnohonásobně více obrázků, konkrétně 33858. Opět se zde ale použil pouze jeden dataset náhodně rozpůlen na testovací a validační množinu. Četnost jednotlivých hodnot pro popis hrubosti povrchu je zobrazena na obázku 5.17.

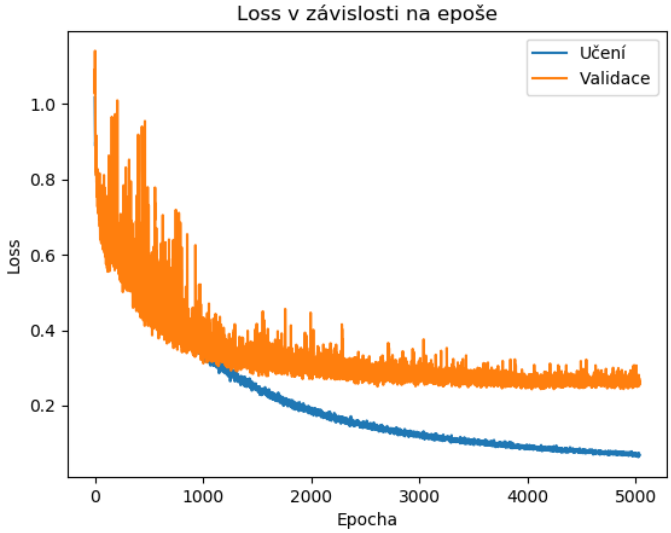
Při učení se hodnota loss funce dostala opět na hodnoty 0,2 a při validaci na 0,4. Graf hodnot loss funkcí je zobrazen na obrázku 5.18. Zde je vidět, že při validaci hodnota loss funkce nejdříve klesne tak, jak bychom předpokládali, dále se ale začne zvyšovat a tím se i zhorší možná predikce.

Na obrázku 5.19 jsou vidět výsledky predikce pro učící množinu. Opět, stejně jako tomu bylo u prvního datasetu, jsou výsledky predikcí až na několik výjimek celkem přesné.

Naopak u výsledků predikce pro validační množinu zobrazenou na obrázku 5.20 je vidět, že s rostoucí hrubostí povrchu je predikce méně přesná. Stále je



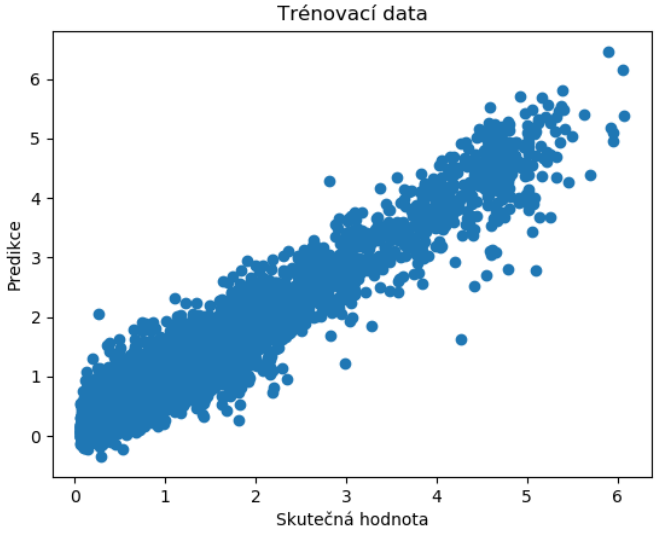
Obrázek 5.17: Dataset 2 - Histogram rozložení hrubosti povrch



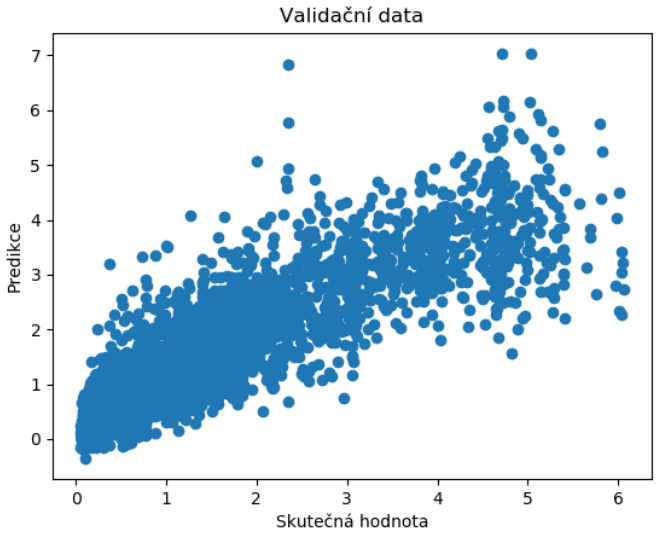
Obrázek 5.18: Dataset 2 - Hodnoty MSELoss pro dané epochy

však patrný stejný trend, jako tomu bylo u učící množiny a datasetu 1.

Jako další byl nasbírán a použit dataset 3. V tomto datasetu již byly provedeny 2 jízdy na stejné trase. Trénovací data byla projeta v protisměru hodinových ručiček a obsahovala 38675 trénovacích obrázků. Pro validační data byla trasa projeta v protisměru hodinových ručiček a obsahovala 39678 obrázků. Ačkoli se však jednalo o stejnou trasu, auto ji neprojelo po totožné trajektorii. Zajistilo se tím ale nasbírání stejného počtu obrázků reprezentují-



Obrázek 5.19: Dataset 2 - Predikce na trénovacích datech, $\rho = 0{,}963$

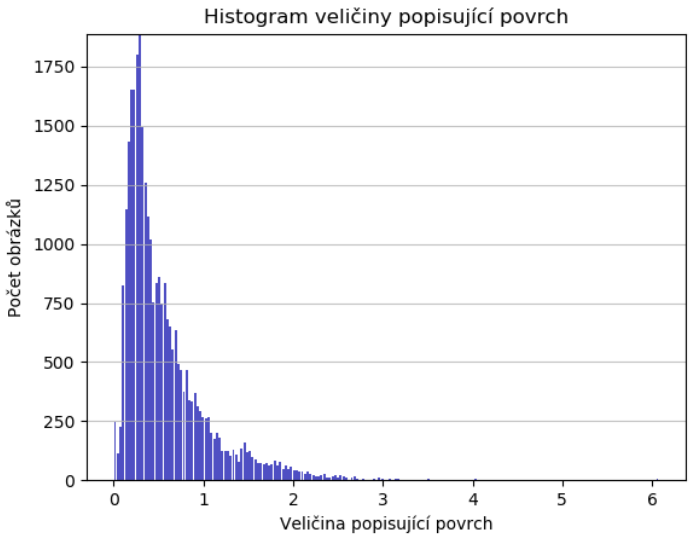


Obrázek 5.20: Dataset 2 - Predikce na validačních datech, $\rho = 0{,}903$

cích daný povrch jak v trénovací, tak i ve validační množině.

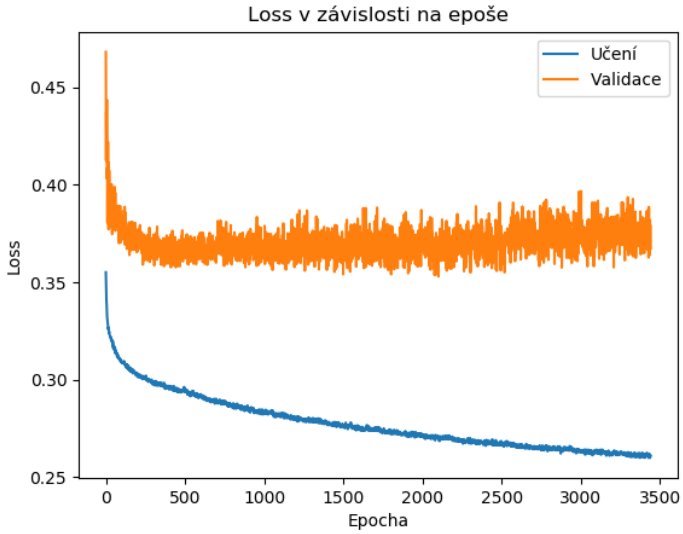
Rozložení hodnot veličiny popisující hrubost povrchu je vidět na obrázku 5.21. Jelikož jde již o velký dataset, je zde dobře vidět, že je zde mnohonásobně větší zastoupení hladších povrchů než těch hrubších.

Z grafu 5.22 je patrné, že hodnoty loss funkce pro učící množinu klesají až na hodnotu 0,26. U validační množiny pak dosahují hodnot 0,36. Je zde také



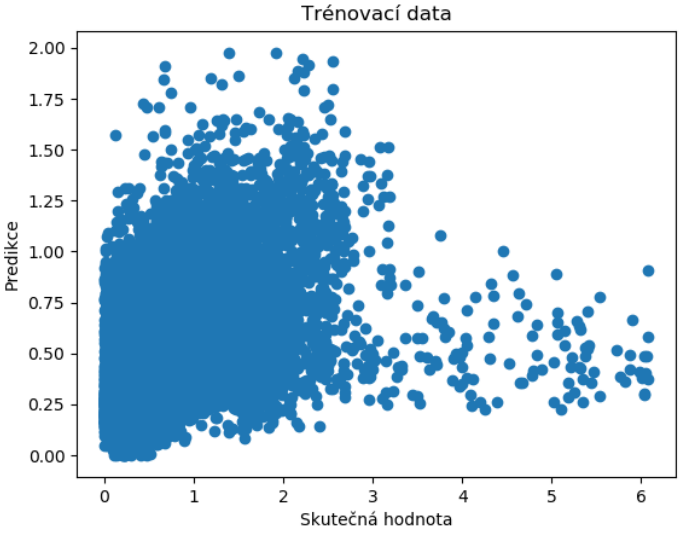
Obrázek 5.21: Dataset 3 - Histogram rozložení hrubosti povrch

patrné, že hodnoty pro učící množinu mohou dále klesat, je tedy možné, že se nedosáhlo globálního minima. U validační množiny je naopak patrné,že hodnoty již dosáhly lokálního minima, a v tomto případě se zdá, že by to mohlo odpovídat i globálnímu minimu.



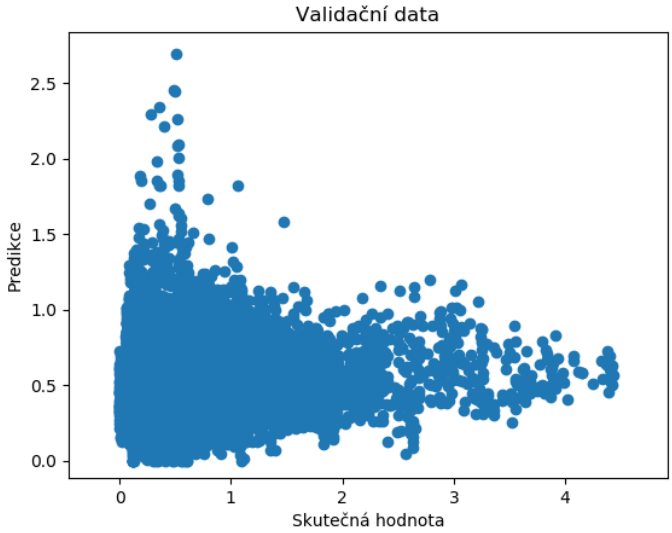
Obrázek 5.22: Dataset 3 - Hodnoty MSELoss pro dané epochy

Na obrázku 5.23 jsou zobrazeny výsledky predikcí na učící množině. Je zde patrné, že predikce dosahuje mnohem menších hodnot, než jsou skutečné hodnoty. Také je zde patrné, že převládají skutečná data s hodnotami od 0 do 3.



Obrázek 5.23: Dataset 3 - Predikce na trénovacích datech, $\rho = 0{,}484$

Na obrázku výsledných predikcí pro validační množinu, která je zobrazena na obrázku 5.24, je také patrné, že predikované hodnoty jsou menší než skutečné hodnoty. Jsou zde ale i predikce pro malé hodnoty, které dosahují naopak poměrně vysokých hodnot.



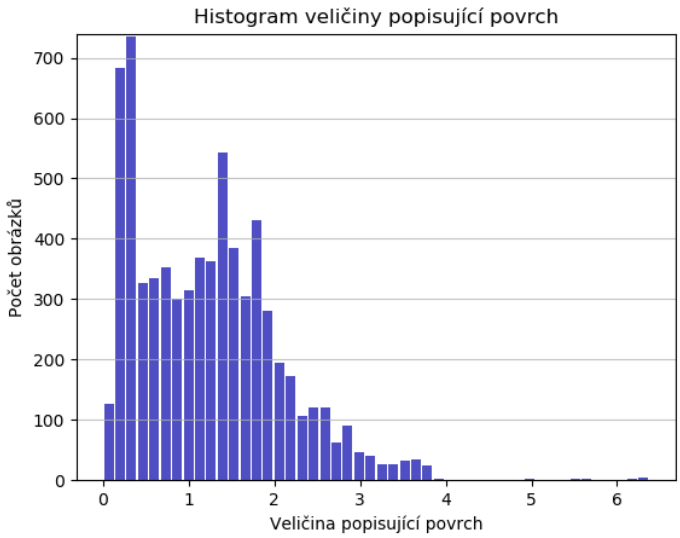
Obrázek 5.24: Dataset 3 - Predikce na validačních datech, $\rho = 0{,}05$

U tohoto datasetu bylo z grafů vidět, že predikce dosahovaly jak u trénovací, tak i u validační množiny mnohem menších hodnot, než byly skutečné hodnoty. Tento fakt mohl být nejspíše způsoben důsledkem výrazně nevyváženého datasetu. Bylo zde velké množství obrázků, kterým odpovídala nízká hodnota

hrubosti, a naopak velmi málo těch s velkou hrubostí. Neuronová síť se tak s největší pravděpodobností naučila převážně predikovat nízké hodnoty.

Posledním popsaným datasetem v této práci je dataset 4. Zde bylo stejně, jako tomu bylo u datasetu 3, využito dvou jízd po stejném povrchu, ale v různém směru. V tomto datasetu byla nasbírána data pro asfalt, trávu a kamennou dlažbu. Kamenná dlažba byla nasbíraná a použita pouze v tomto datasetu.

V tomto datasetu již byly reálné hodnoty pro popis hrubosti nasbírány o trochu víc rovnoměrněji. Četnosti obrázků s danou hodnotou hrubosti jsou zobrazeny na obrázku 5.25.

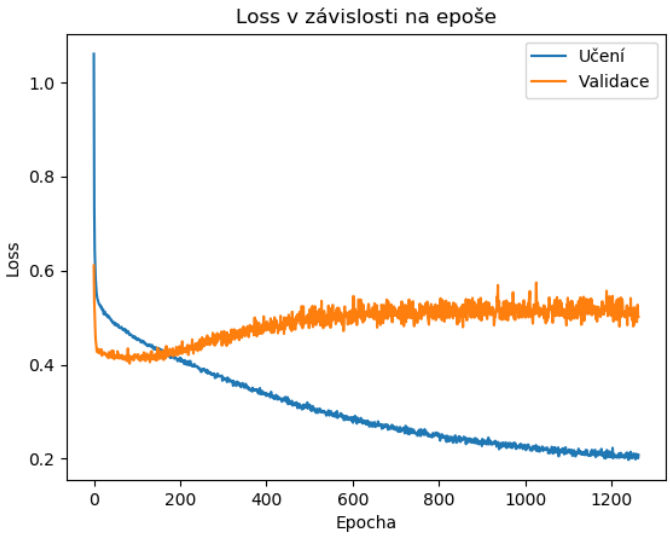


Obrázek 5.25: Dataset 4 - Histogram rozložení hrubosti povrchů

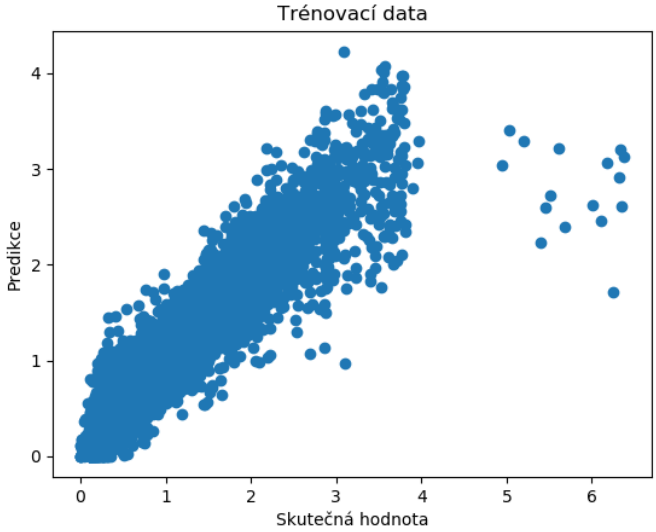
Na obrázku 5.26 jsou vidět hodnoty loss funkce pro tento dataset. Je zde patrné, že hodnoty pro učící množinu dosahují hodnot 0,2. Validační hodnoty pak nejdříve dosáhly hodnot 0,4, což bylo lokální minimum, a následně hodnota loss stoupla na 0,5.

Predikce na učící množině je zobrazena na obrázku 5.27. Až na několik predikcí, odpovídající velmi hrubému povrchu, byly zbylé predikce celkem přesné a bylo by možné přiřadit je k jednotlivým povrchům.

U predikcí na validační množině je naopak patrné, že predikce již není tak přesná. Je zde mnoho predikcí, které neuronová síť nedokázala predikovat jako velmi hrbolaté a také je zde patrný velký shluk predikcí, které hladké povrchy predikuje již jako hrubší povrchy.



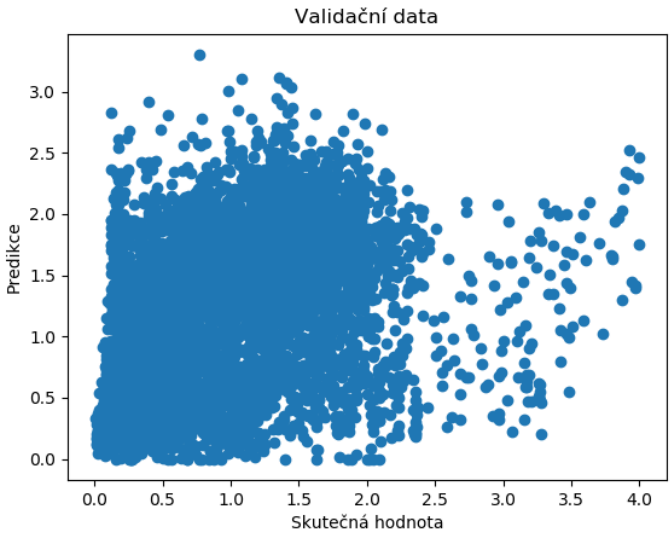
Obrázek 5.26: Dataset 4 - Hodnoty MSELoss pro dané epochy



Obrázek 5.27: Dataset 4 - Predikce na trénovacích datech, $\rho = 0{,}929$

U datasetu 4 bylo patrné, že predikce pro trénovací množinu byla lepší než tomu bylo u datasetu 3. Predikce na validační množině pak nedržely trend, který byl u datasetu 3. Přesto dosahovaly hodnoty korelačního koeficientu mezi predikovanými a skutečnými daty pro trénovací množinu 0.929 a pro validační množinu 0.478.

Z grafů predikcí na validačních sadách je zřejmé, že naučené konvoluční sítě měly problémy s generalizací na datech, které nebyly obsaženy v trénovací



Obrázek 5.28: Dataset 4 - Predikce na validačních datech, $\rho = 0{,}478$

množině. Pro lepší výsledky by se mohlo využít technik, které by predikce mohly zlepšit, jako jsou například regularizace a dropout. Také by zajisté pomohlo více dat s rovnoměrným rozložením hrubostí povrchů.

5.7 Výsledky

Pro predikci hrubosti povrchu bylo využito dvou odlišných přístupů. První, referenční metodou bylo využití obrazového gradientu k predikci hrubosti povrchu. Druhou, navrhovanou metodou, bylo využití dat nasbíraných pomocí akcelerometru jako vstup pro učení a predikci pomocí neuronové sítě. Obě metody však mají své výhody a nevýhody.

U první metody je výhodou rychlost zpracování bez nutnosti učení klasifikátoru. Nevýhodou je pak to, že se metoda není schopna naučit například stíny v obraze a vždy je bude určovat jako drsnější povrch. U druhé metody je výhodou, že se je naopak při dostatečném množství dat schopna naučit, že například stín na asfaltu nepředstavuje výmol a nebude tedy predikovat velké hodnoty hrubosti. Nevýhodou je pak nutnost velkého množství dat z důvodu generalizace predikce na různých místech a za různých světelných podmínek. Dále také nutnost učení sítě, které je při velkém množství dat velice zdlouhavé.

Experimenty v této práci proběhly na 4 datasetech. Datasety 1, 2 a 3 obsahovaly obrázky asfaltu, trávy, lesní cesty a panelové cesty. Dataset 4 pak obsahoval asfalt, trávu a kamennou dlažbu. Datasety 1 a 2 využily jedné jízdy, která byla náhodně rozpůlena na trénovací a na validační množinu. Datasety 3 a 4 byly tvořeny ze dvou jízd provedené na jedné trase, ale projety byly opačným směrem.

Pro relevantní srovnání obou použitých metod byla spočtena korelace mezi predikovanými a reálnými hodnotami pro každý z již zmíněných datasetů. Tyto údaje jsou zaznamenány v tabulce 5.3. Zde je patrné, že predikce s využitím obrazového gradientu byla ve všech datasetech horší, než predikce s využitím neuronové sítě ResNet-50 na validačních datech.

	Korelační koeficienty ρ		
	pro trénovací data	pro validační data	pro obrazový gradient
Dataset 1	**0,982**	**0,921**	0,282
Dataset 2	0,963	0,903	0,126
Dataset 3	0,484	0,050	0,012
Dataset 4	0,929	0,478	**0,294**

Tabulka 5.3: Korelační koeficienty predikcí a reálných hodnot pro jednotlivé datasety

Z tabulky 5.3 je vidět, že nejsilnější kladné korelace byly pro konvoluční neuronovou síť u datasetu 1 jak na trénovací, tak i na validační množině. Tyto hodnoty dosahovaly velmi silné kladné korelace.

U predikcí s využitím referenční metody obrazového gradientu byla nejsilnější kladná korelace u datasetu 4. Tato hodnota však dosahovala pouze slabé kladné korelace.

Kapitola 6

Závěr

Tato práce se zabývala predikcí hrubosti povrchu před autem s využitím kamery a akcelerometru. V experimentech byl použit model auta Toyota Mini, který mimo jiné obsahuje výpočetní zařízení Jetson AGX Xavier a Raspberry Pi 2. Pro tuto práci pak byla na auto namontována IMU jednotka Xsens MTi-7-DK. S tímto autem byla nasbírána veškerá data pro jejich následné zprocesování.

Cílem této práce byl sběr dat a následně vytvoření metody pro automatickou anotaci těchto dat. Následně natrénování konvoluční neuronové sítě na různých datesetech o různém počtu dat a vyhodnocení přesnosti predikcí hrubosti povrchu.

Jako referenční metoda pro predikci byl vybrán obrazový gradient spočtený pro fotografie z kamery. Korelační koeficient predikovaných hodnot dosahoval hodnot od 0,012 do 0,294. Pro navrhovanou metodu, využívající konvoluční neuronovou síť architektury ResNet-50, pak dosahoval korelační koeficient na trénovacích datech hodnot od 0,484 do 0,982 a na validačních datech od 0,05 do 0,921.

Experimenty provedené v této práci ukazují velmi slibné výsledky predikce hrubosti povrchu s využitím konvoluční neuronové sítě. Ve všech experimentech překonaly výsledky konvoluční neuronové sítě výsledky základní referenční metody založené na výpočtu obrazového gradientu. Nicméně u konvoluční neuronové sítě je problém s generalizací na validačních datech, tedy na datech, které síť neviděla. V literatuře však existují techniky jak zabránit prefitování. Přikladem je využití dropoutu nebo regularizace.

Veškeré experimenty proběhly na modelu auta Toyota Mini, avšak navrhovanou metodu by bylo možné využít i u osobních automobilů. Pro rozsáhlejší využití této metody je však zapotřebí dalších experimentů a zlepšení generalizace dat.

Literatura

[BFP+20] Michal Bahník, Dominik Filyo, David Pekárek, Martin Vlašimský, Jan Čech, Tomáš Haniš, and Martin Hromčík, *Visually assisted anti-lock braking system*, IEEE Intelligent Vehicles Symposium (IV) (2020).

[BI05] C. Brooks and K. Iagnemma, *Vibration-based terrain classification for planetary exploration rovers*, IEEE TransactionsonRobotics (2005), 1185–1191.

[cita] *Google maps*, `https://www.google.cz/maps/dir///@50.0397835,14.5276765,17z/data=!4m2!4m1!3e0?hl=cs&authuser=0`, cit. 2020-08-11.

[citb] *Jetson agx xavier developer kit*, `https://developer.nvidia.com/embedded/jetson-agx-xavier-developer-kit`, cit. 2020-08-11.

[citc] *Losi desert buggy xl-e 1/5 4wd electric buggy (black/yellow)*, `https://www.rcplanet.com/rc-vehicles/car-truck/losi-desert-buggy-xl-e-1-5th-4wd-rtr-black-yellow-los05012t1`, cit. 2020-08-11.

[citd] *Navio 2*, `https://robots.ros.org/navio2`, cit. 2020-08-11.

[Cite] *Scipy library - stft*, `https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.stft.html`, cit. 2020-08-11.

[citf] *Xsens mti-7-dk*, `https://www.newark.com/xsens/mti-7-dk/dev-kit-3d-motion-tracking/dp/69AC6243`, cit. 2020-08-11.

[citg] *Zed camera*, `https://www.stereolabs.com/zed`, cit. 2020-08-11.

[ea06] Sebastian Thrun et al., *Stanley: The robot that won the darpa grand-challenge*, Journalof Field Robotics 23(9) (2006), 661–692.

[Eva96] J.D. Evans, *Straightforward statistics for the behavioral sciences*, Brooks/Cole Publishing Company, 1996.

[HZRS15] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, *Deep residual learning for image recognition*, CoRR **abs/1512.03385** (2015).

[ID02] Karl D. Iagnemma and Steven Dubowsky, *Terrain estimation for high-speed rough-terrain autonomous vehicle navigation*, Unmanned Ground Vehicle Technology IV (Grant R. Gerhart, Chuck M. Shoemaker, and Douglas W. Gage, eds.), vol. 4715, International Society for Optics and Photonics, SPIE, 2002, pp. 256 – 266.

[MSE17] Jirapong Manit, Achim Schweikard, and Floris Ernst, *Deep convolutional neural network approach for forehead tissue thickness estimation*, Current Directions in Biomedical Engineering **3** (2017).

[PJ96] D. Pomerleau and T. Jochem, *Rapidly adapting machine vision for automated vehicle steering*, IEEE Expert **11** (1996), no. 2, 19–27.

[PYK08] Pushkar Pradham, Nicolas H. Younan, and Roger L. King, *16 - concepts of image fusion in remote sensing applications*, Image Fusion (Tania Stathaki, ed.), Academic Press, Oxford, 2008, pp. 393 – 428.

[Wel67] Peter D. Welch, *The use of fast fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms*, IEEE Transactions on Audio and Electroacoustics **15** (1967), 70–73.

[ZF14] Matthew D. Zeiler and Rob Fergus, *Visualizing and understanding convolutional networks*, Computer Vision – ECCV 2014 (Cham) (David Fleet, Tomas Pajdla, Bernt Schiele, and Tinne Tuytelaars, eds.), Springer International Publishing, 2014, pp. 818–833.

[ZZWL15] Y. Zhao, Z. Zou, L. Wu, and Y. Li, *Frequency detection algorithm for frequency diversity signal based on stft*, 2015 Fifth International Conference on Instrumentation and Measurement, Computer, Communication and Control (IMCCC), 2015, pp. 790–793.