

Univerzita Karlova

Přírodovědecká fakulta

Studijní program: Geografie (magisterské studium)

Studijní obor: Kartografie a geoinformatika



Bc. Martina ŘÁDOVÁ

**KLASIFIKACE UAV HYPERSPAKTRÁLNÍCH OBRAZOVÝCH DAT S
VYUŽITÍM METOD HLUBOKÉHO UČENÍ**

**CLASSIFICATION OF UAV HYPERSPECTRAL IMAGES USING
DEEP LEARNING METHODS**

Diplomová práce

Vedoucí diplomové práce: Ing. Markéta Potůčková, Ph.D.

Praha, 2020

Vysoká škola: Univerzita Karlova

Fakulta: Přírodovědecká

Katedra: Aplikované geoinformatiky a kartografie

Akademický rok: 2020/2021

Zadání diplomové práce

pro Bc. Martina Řádová

obor Kartografie a geoinformatika

Název tématu:

Klasifikace UAV hyperspektrálních obrazových dat s využitím metod hlubokého učení

Zásady pro vypracování

Diplomová práce v obecné rovině řeší problematiku klasifikace krajinného pokryvu z UAV hyperspektrálních obrazových dat s vysokým prostorovým rozlišením (0,09 m resp. 0,03 m) s využitím konvolučních neuronových sítí (CNN). Navazuje na dosavadní výsledky projektu „Vývoj metod pro sledování změn vegetace a krkonošské tundry analýzou dat z multispektrálních, hyperspektrálních a LIDAR senzorů UAV“ řešeného katedrou aplikované geoinformatiky a Správou KRNP. Oblastí zájmu je experimentální lokalita o rozloze 100 m x 100 m na Luční hoře. Klasifikační legenda je dána projektem a zahrnuje 9 kategorií – 4 kategorie převážně zapojených alpínských trávníků (s dominantním zastoupením metlice trstnaté, metličky, smilky tuhé), brusnicová vegetace, borovice kleč, vřes obecný, vyfoukávané alpínské trávníky, suťová pole.

Diplomová práce by ve své rešeršní části měla podat obecný přehled o východiscích a metodách CNN. Prvotním cílem je z poznatků získaných z literatury sestavit přehledný rámec, který by byl podkladem pro vytipování vhodné klasifikační metody hyperspektrálních snímků s vysokým prostorovým rozlišením. Z výsledků autorů používajících vybranou metodu by měla vyplývat její vhodnost pro odlišení tříd krajinného pokryvu s blízkou spektrální odezvou. Dílčím úkolem je příprava dat z tundry pro jejich úspěšný vstup do zvolené architektury. Dalším dílčím cílem práce je odladění parametrů sítě tak, aby bylo dosaženo co nejvyšší přesnosti a zároveň nedošlo k jejímu přetrénování. Kvalita výsledku bude hodnocena na nezávislé sadě validačních bodů. Klasifikace bude provedena na datech pořízených v měsících červen, červenec a srpen 2020 a na jejich multitemporálním kompozitu. Bude diskutována otázka odlišení jednotlivých tříd v daných termínech vegetační sezóny. Výsledná přesnost klasifikace bude porovnána s výsledky jiných klasifikačních metod použitých v projektu (maximum likelihood, random forest).

Rozsah grafických prací: Dle potřeby

Rozsah průvodní zprávy: cca 70 stran

Seznam odborné literatury:

- Fu, G., Liu, C., Zhou, R., Sun, T., & Zhang, Q. (2017). Classification for high resolution remote sensing imagery using a fully convolutional network. *Remote Sensing*, 9(5), 498.
- Hu, W., Huang, Y., Wei, L., Zhang, F., & Li, H. (2015). Deep convolutional neural networks for hyperspectral image classification. *Journal of Sensors*, 2015.
- Li, S., Song, W., Fang, L., Chen, Y., Ghamisi, P., & Benediktsson, J. A. (2019). Deep learning for hyperspectral image classification: An overview. *IEEE Transactions on Geoscience and Remote Sensing*, 57(9), 6690-6709.
- Paoletti, M. E., Haut, J. M., Fernandez-Beltran, R., Plaza, J., Plaza, A. J., & Pla, F. (2019). Deep pyramidal residual networks for spectral-spatial hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 57(2), 740-754.
- Paoletti, M. E., Haut, J. M., Plaza, J., & Plaza, A. (2019). Deep learning classifiers for hyperspectral imaging: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 158, 279-317.
- Vali, A., Comai, S., & Matteucci, M. (2020). Deep learning for land use and land cover classification based on hyperspectral and multispectral earth observation data: A review. *Remote Sensing*, 12(15), 2495.

Vedoucí diplomové práce: Ing. Markéta Potůčková, Ph.D.

Datum zadání diplomové práce: 29.11.2019

Termín odevzdání diplomové práce: září 2021

Platnost tohoto zadání je po dobu jednoho akademického roku.

.....
Vedoucí diplomové práce

.....
Garant studijního oboru

V Praze dne 7. září 2021

Prohlášení

Prohlašuji, že jsem tuto diplomovou práci zpracovala samostatně, pod vedením školitelky Ing. Markéty Potůčkové, Ph.D., a že jsem uvedla všechny použité informační zdroje a literaturu.

Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

Svoluji k zapůjčení této práce pro studijní účely a souhlasím s tím, aby byla řádně vedena v evidenci vypůjčovatelů.

V Praze, dne 7. září 2021

Bc. Martina Řádová

Poděkování

Ráda bych tímto upřímně poděkovala vedoucí mé diplomové práce Ing. Markétě Potůčkové, Ph.D., za její čas, který mi věnovala, odborné vedení, podnětné rady a trpělivost, které velmi přispěly k dokončení této práce. Mé poděkování patří také všem osloveným odborníkům, za jejich čas, ochotu a cenné rady získané při společných konzultacích. Dále bych chtěla velmi poděkovat mé rodině a blízkým přátelům za podporu během studia a zpracovávání této diplomové práce.

Abstrakt

Uvedená diplomová práce “Klasifikace UAV hyperspektrálních (HS) obrazových dat s využitím metod hlubokého učení“ se v obecné rovině věnuje klasifikačním metodám hyperspektrálních snímků. Ve své rešeršní části diplomová práce podává obecný přehled metod konvolučních neuronových sítí. Na základě toho je sestaven přehledný rámec jako podklad pro vytipování vhodné metody pro tuto práci. Vybrány jsou dvě metody s otevřeným řešením v jazyce Python. Zvolenými metodami jsou Capsule Network a U-Net. Cílem práce je ověřit vhodnost těchto metod pro klasifikaci hyperspektrálních snímků krkonošské tundry s vysokým prostorovým rozlišením. Dílčím cílem je i příprava vstupních HS (54 pásem, 9 cm) dat do vhodné podoby pro vstup do sítě.

Vzhledem ke složitosti architektury nebylo dosaženo všech požadovaných výsledků u metody Capsule Network. Pro účely porovnání a ověření výsledků byla použita metoda U-Net. Ta dosáhla přesnějších výsledků oproti hodnotám získávaným tradičními metodami (SVM, ML, RF a další), kdy celková přesnost byla u U-Net vyšší než 90% a u ostatních zmiňovaných metod OA nepřesáhla 88%. Zejména třídy suť a kleč vyšly výrazně přesněji než všechny ostatní třídy (UA – user’s accuracy a PA – producer’s accuracy přes 99%).

Klíčová slova

Hluboké učení, konvoluční neuronové sítě, hyperspektrální snímky, Capsule Network, U-Net, klasifikace, mapování vegetace

Abstract

Diploma thesis "Classification of UAV hyperspectral images using deep learning methods" focuses on the classification methods, namely convolutional neural networks (*CNN*), of hyperspectral (*HS*) images. Based on a thorough literature review, a comprehensive overview on *CNN* utilisation in remote sensing is assembled as a basis for identifying suitable methods for the specific task of this thesis. Two methods with an open solution in programming language Python were selected - Capsule Network and U-Net. The main aim of this work is to verify the suitability of chosen methods for the classification of hyperspectral images. The images were acquired by sensors with high spatial resolution carried by a UAV over Krkonoše Mts. tundra. Important step was to prepare input HS data (54 bands, 9cm) to have suitable form for entering the network.

Not all the required results were achieved due to the complexity of the Capsule Network architecture. The U-Net method was used in purpose of comparing and verifying the results. Accuracies retrieved from the U-Net overcome results achieved by traditionally used machine learning methods (SVM, ML, RF, etc). Overall accuracy for U-Net was higher than 90% where other mentioned methods did not get over 88%. Especially classes block fields and dwarf pine achieved higher accuracies comparing to other classes (UA – user's accuracy a PA – producer's accuracy over 99%).

Key words

Deep learning, convolutional neural networks, hyperspectral images, Capsule Network, U-Net, classification, vegetation mapping

Obsah

Seznam obrázků	10
Seznam tabulek	10
Seznam grafů.....	10
Seznam příloh.....	11
1 Úvod a cíle práce.....	10
2 Literární rešerše a úvod do problematiky.....	12
2.1 Hluboké učení	13
2.2 Hyperspektrální data	19
2.3 Přehled architektur hlubokého učení pro klasifikaci HS dat	21
2.4 Klasifikace HS dat v tundře	31
2.5 Klasifikace HS dat v krkonošské tundře.....	32
3 Zájmové území a vstupní data.....	35
3.1 Zájmové území	35
3.1.1 Luční hora.....	36
3.2 Senzor	38
3.3 Použitá data.....	38
3.3.1 Předzpracování dat	38
3.3.2 Vstup do klasifikace	39
3.3.2.1 Volně dostupné datové sady	39
3.3.2.2 HS data HeadWall Nano-Hyperspec	40
3.3.3 Třídy	40
4 Metodika	42
4.1 Capsule network	42
4.2 Konkrétní Capsule Network	45
4.3 U-Net	49
5 Praktická část	51

5.1	Příprava dat	51
5.2	Práce s kódy	52
6	Výsledky	55
7	Diskuse	63
8	Závěr	66
9	Zdroje	68
9.1	Publikace	68
9.2	Internet	71
10	Přílohy	73

Seznam obrázků

Obrázek 1: Vymezení zájmového území.....	36
Obrázek 2: Luční hora – tvar (a) a útvary (b) (Zdroj: Nebeske	37
Obrázek 3: Dron s kamerou (Zdroj: Červená, L., 2020)	38
Obrázek 4: Základní schéma sítě Capsule Network	44
Obrázek 5: Dynamic routing algoritmus	45
Obrázek 6: Schéma Capsule Network Paoletti, M., et al. (2019a)	48
Obrázek 7: Schéma 1D Conv-Capsule sítě Zhu, K., et al. (2019).....	48
Obrázek 8: Schéma 3D Conv-Capsule sítě Zhu, K., et al. (2019).....	49
Obrázek 9: Architektura U-Net	50
Obrázek 10: Chybová matice po trénování pokus 5.....	57
Obrázek 11: Výsledné snímky trénovacích klasifikací	58

Seznam tabulek

Tabulka 1: Parametry krkonošských dat.....	40
Tabulka 2: Legenda.....	41
Tabulka 3: Trénovaná síť IP s různými parametry.....	55
Tabulka 4: Nastavení vah tříd	56
Tabulka 5: Ladění parametrů pro ostrou klasifikaci	58
Tabulka 6: Parametry pro vstup do klasifikace	59
Tabulka 7: Výsledky klasifikací.....	61
Tabulka 8: Chybová matice srpen	62
Tabulka 9: Chybová matice září.....	62
Tabulka 10: Chybová matice multitemporální kompozit.....	62

Seznam grafů

Graf 1: Hierarchie Umělé inteligence Zdroj: vlastní	12
Graf 2: Modely NN, Zdroj: JavaTpoint (2021)	14
Graf 3: DL a jeho dělení Zdroj: vlastní	15
Graf 4: Modely CNN. Zdroj: vlastní	16
Graf 5: Další dělení DL. Zdroj: vlastní	16
Graf 6: Opakovaně řešené problémy CNN. Zdroj: vlastní.....	17
Graf 7: Porovnání přístupu Paoletti a Zhu	46

Seznam příloh

<i>Příloha 1: Parametry data setů a HS dat</i>	<i>73</i>
<i>Příloha 2: Přehled NN. Zdroj: https://www.asimovinstitute.org/neural-network-zoo/.....</i>	<i>74</i>
<i>Příloha 3: Taxonomy DCNN. Zdroj: https://arxiv.org/ftp/arxiv/papers/1901/1901.06032.pdf</i>	<i>75</i>
<i>Příloha 4: Snímky metličky křivolaké (vlevo polovina června) (vpravo srpen).....</i>	<i>75</i>
<i>Příloha 5: Snímky smilky tuhé (vlevo druhá polovina července) (vpravo konec srpna).....</i>	<i>76</i>
<i>Příloha 6: Snímky metlice trsnaté (vlevo konec června) (vpravo začátek srpna)</i>	<i>76</i>
<i>Příloha 7: Výsledný snímek klasifikace za období srpen.....</i>	<i>77</i>
<i>Příloha 8: Výsledný snímek klasifikace za období září.....</i>	<i>78</i>
<i>Příloha 9: Výsledný snímek klasifikace multitemporální kompozit</i>	<i>79</i>

1 Úvod a cíle práce

Metody hlubokého učení (*Deep learning – DL*) zaznamenaly v posledních letech výrazný rozvoj v mnoha oborech využívajících informační technologie. Týká se to i aplikací spojených se zpracováním obrazu, mezi něž patří i dálkový průzkum Země (*DPZ*). Vedle metod zpracování obrazu se vyvíjejí a zpřístupňují i metody pořizování obrazových dat. Hyperspektrální (*HS*) skenery snímají zemský povrch v řádu stovek úzkých na sebe navazujících spektrálních pásem a jsou v současné době využívány na družicových, leteckých i UAV (*Unmanned Aerial Vehicle*) platformách. Díky svému spektrálnímu rozlišení skýtají velký potenciál pro získání podrobných informací o krajinném povrchu a jeho vlastnostech. Klasické přístupy strojového učení (*Maximum likelihood – ML*, *Support vector machine – SVM*, *Random forest – RF* a další) využívají spektrální příznaky obrazových dat. Metody DL, konkrétně konvoluční neuronové sítě (*Convolutional Neural Networks – CNN*), při klasifikaci obrazu vedle příznaků spektrálních, pracují díky své architektuře i s prostorovými vztahy. Otázkou tak je, do jaké míry tyto složitější a výpočetně náročnější přístupy mohou v konkrétní úloze zvýšit kvalitu informace výtěžné z HS dat oproti klasickým přístupům.

Motivací pro téma této diplomové práce je účast v projektu „Vývoj metod pro sledování změn vegetace a krkonošské tundry analýzou dat z multispektrálních, hyperspektrálních a LIDAR senzorů UAV“, probíhajícího od roku 2019. Řešení zajišťuje Katedra aplikované geoinformatiky a kartografie Přírodovědecké fakulty UK ve spolupráci se Správou Krkonošského národního parku a Botanického ústavu AV. Na čtyřech experimentálních plochách (Bílá louka, Luční hora, Úpská rašeliniště a kar Úpské jámy) jsou mimo jiné, 4x za sezónu pořizována HS data senzorem Headwall Nano—Hyperspec neseným UAV s prostorovým rozlišením 3cm a 270 hyperspektrálními pásmy. Klasifikační legenda je sestavena ve spolupráci s botaniky z Botanického ústavu a zahrnuje vybrané druhy travních a keříkových společenstev specifických pro oblast krkonošské tundry, kosodřevinu a suťová pole. Samotná klasifikace byla dosud realizována několika metodami strojového učení (ML, SVM, RF na pixelové a objektové úrovni), které ukázaly své limity při klasifikaci heterogenních společenstev (např. alpské zapojené a vyfoukávané trávníky) a druhů s podobným spektrálním projevem (Kupková, L., et al., 2020). Diplomová práce si proto klade za cíl zodpovědět otázku, zda metody hlubokého učení díky odlišné filozofii a práci s příznaky mohou docílit vyšší přesnosti klasifikace než použité metody strojového učení.

Diplomová práce ve své rešeršní části podává obecný přehled o východiscích a metodách CNN. Prvotním cílem je z poznatků získaných z literatury sestavit přehledný rámec, který by byl podkladem pro vytipování vhodné klasifikační metody HS snímků s vysokým prostorovým

rozlišením. Z výsledků autorů používajících vybranou metodu by měla vyplývat její vhodnost pro odlišení tříd krajinného pokryvu s blízkou spektrální odezvou. Dílčím cílem je i příprava dat z tundry pro jejich úspěšný vstup do zvolené architektury a následné nalezení nejvhodnějších vstupních parametrů sítě tak, aby bylo dosaženo co nejpřesnějších výsledků.

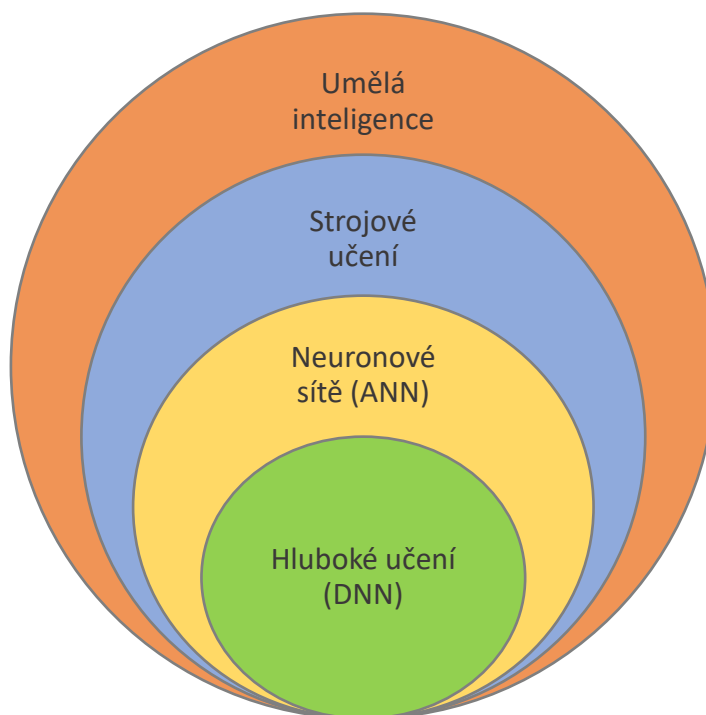
Preferovanou implementací vybrané metody je otevřené řešení v jazyce Python. V metodách hlubokého učení se jedná o hojně využívaný programovací jazyk s množstvím otevřených platforem pro sdílení nejen samotných architektur, ale i poznatků a řešených problémů. Python využívá i z rešerše vytipovaná metoda Capsule Network, která nejlépe odpovídá stanoveným požadavkům adaptace sítě pro HS data, nebo U-Net, která je zvolena jako příklad v aplikacích zpracování obrazu často používané plně propojené CNN, s cílem porovnat výsledky těchto rozdílných architektur a jejich přínos pro klasifikování HS snímků.

Během práce bude nutné zodpovědět několik otázek. Jak bude získaná architektura vypadat – budou nutné její úpravy? Bude možné ji použít na podrobných datech z tundry? Jaké budou vstupní parametry? Bude nutné vstupní data nějak upravovat, například formát, velikost apod.? Zásadní otázkou k zodpovězení bude, do jaké míry je zvolená metoda přínosem. Zda pomocí ní došlo ke zpřesnění klasifikace, oproti jiným klasifikačním metodám (např. MLC, RF, SVM) použitých na konkrétních datech z krkonošské tundry. Pro praktickou část práce jsou využita UAV HS data pořízená senzorem Headwall Nano—Hyperspec během vegetační sezóny roku 2020 v experimentální ploše Luční hora.

2 Literární rešerše a úvod do problematiky

Rešerše má za cíl podat přehled současné literatury se zaměřením na DL a jeho využití pro klasifikaci HS dat. Snaží se jít od obecného přehledu problematiky ke konkrétním studiím a metodám, které jsou uvedeny v následujících kapitolách. Pozornost je věnována i konkrétním sítím, jejich zdrojům, dostupnosti a možnostem implementace. Rozebrány jsou specifické vlastnosti HS dat zároveň s jejich parametry a možnostmi klasifikace. V neposlední řadě jsou zmíněny články zabývající se specifickým územím tundry, nejprve obecně a následně konkrétně v oblasti Krkonoš. Hlavním předmětem zájmu jsou senzory a metody, které jsou použity pro výzkum těchto území.

Původ umělé inteligence (*Artificial Intelligence – AI*) spadá do 50. let minulého století, kdy si hrstka průkopníků položila otázku, zda by mohli přimět počítače přemýšlet, jinými slovy automatizovat intelektuální úkoly, které jsou běžně prováděny lidmi (Chollet, F., 2018). Tato otázka je aktuální dodnes a stále se snažíme zjistit její dosah. Hierarchii metod AI přehledně popisuje Graf 1.

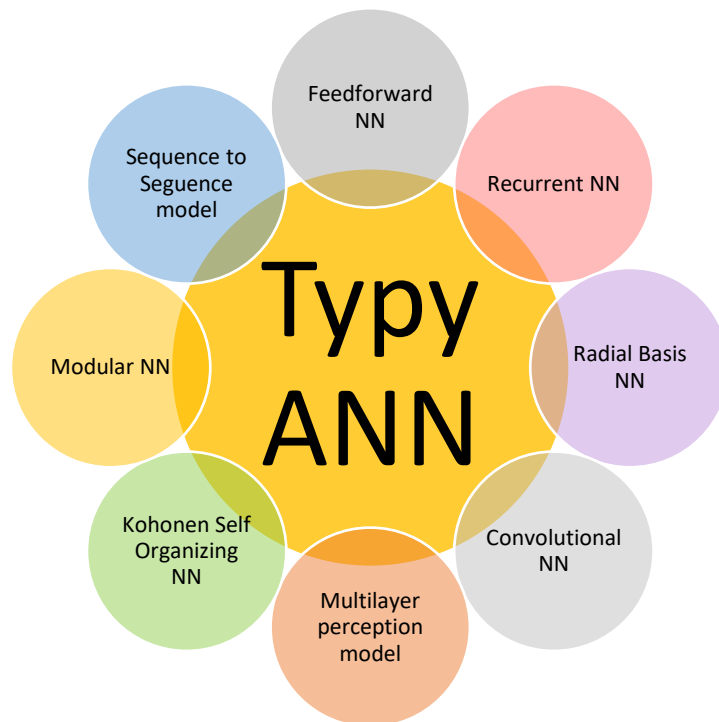


Graf 1: Hierarchie Umělé inteligence Zdroj: vlastní

Podoborem AI je strojové učení (*Machine Learning – ML*). ML vychází z otázek: Mohl by jít počítač dál za hranici toho, co ho naučíme a ukážeme mu, jak provádět? Mohl by se sám naučit, jak splnit konkrétní úkol? Mohl by nás počítač překvapit? Mohla by mu k učení se stačit pouze nová data bez zásahu člověka? (Chollet, F., 2018). Strojové učení se zabývá počítačovými algoritmy, které se automaticky zlepšují díky získaným zkušenostem a používání dat. Tvoří modely na základě trénovacích dat s cílem něco predikovat nebo rozhodnout, aniž by k tomu byly výslovně programovány (Wiki – ML, 2020). Ačkoliv ML úzce souvisí se statistikou, je mezi nimi několik významných rozdílů. Podstatným je tendence strojového učení vypořádat se s velkými a komplexními datasety.

2.1 Hluboké učení

Hluboké učení je součástí širší rodiny metod strojového učení a je založeno na umělých neuronových sítích (*Artificial Neural Networks – ANN* nebo *Neural Networks – NN*). ANN jsou sítě o mnoha vrstvách sestavených z neuronů, které se využívají ke klasifikacím, predikcím apod. Jedná se o systém, který je schopný se učit a napodobuje chování neuronových buněk ve zvířecím mozku. Mnoho buněk v neuronech živých organismů je schopno pracovat současně na dosažení výsledku, tento princip se snaží napodobit i ANN. Kombinují různé typy vstupních dat a dokáží pracovat s nelineárními vztahy mezi vstupními a požadovanými výstupními daty. Neuronová síť sestává z několika částí, záleží vždy na konkrétním modelu (viz Graf 2 a Příloha 2). Jednotlivé vrstvy ANN jsou tvořeny množstvím neuronů (uzlů), ve kterých se odehrávají výpočty. Neurony jsou navzájem propojeny váhovými koeficienty, které mají vliv na sílu signálu přicházejícího z předchozích neuronů. Váhy mají také vliv na samotnou aktivaci neuronu. Součástí procesů spojených s použitím ANN je takzvaná adaptivní fáze, kdy jednotlivé váhy jsou schopny se měnit podle tzv. učícího algoritmu (v případě “učení s učitelem”/řízené klasifikace). Na základě trénovacího souboru, ve kterém známe správné zařazení vstupních dat (v případě DPZ např. pixelů) do jednotlivých klasifikačních tříd, se váhy na základě předložených vzorů mění. Algoritmus zpětného šíření (*Back propagation*) slouží ke zjištění chyb mezi výstupními a trénovacími daty. Zjištěné chyby jsou použity k přenastavení vah v síti. Cílem je dosáhnout co nejmenší možné chyby. Čím více dat může algoritmus trénovat, tím je větší pravděpodobnost přesného výsledku.



Graf 2: Modely NN, Zdroj: JavaTpoint (2021)

Pojem hluboký v DL znamená počet vrstev, které jsou v modelu – neboli hloubka modelu. V současné době moderní modely obsahují desítky až stovky navazujících vrstev, které jsou učený automaticky z trénovacích dat. Každá z těchto vrstev může být specializována na zachycení specifického aspektu nebo příznaku. Ve vrstvách probíhá i nelineární zpracování za účelem naučit se více úrovní reprezentací dat (*multiple levels of data representations*). Těmi mohou být *low*, *mid* a *high* level příznaky. Základem jsou *low-level* příznaky, které jsou tvořeny nejmenšími detaily obrazu jako jsou základní tvary, například linie a body. Na jejich základě jsou vystavěny *high-level* příznaky (velké/komplexní objekty a tvary). DL se stává i důležitým nástrojem pro analýzu big dat. Dle Zhu, X., et al. (2017) je považováno za průlomovou technologii roku 2013, zejména v případech čím dál tím větších, rozmanitějších a komplexnějších datasetů (Pouyanfar, F., et al., 2018). Důvodem rychlého vzestupu DL je, že nabízí lepší výsledky v mnoha problémech a také dělá jejich řešení jednodušší díky automatizovanému hledání příznaků (*feature engineering*).

V níže zobrazeném grafu (Graf 3) je snaha shrnout a ujasnit jednotlivé rámce, které patří do DL a seskupit je do přehlednějších kategorií. Úkol to není vůbec snadný, protože jak je i zmiňováno v literatuře, obecně problémem hlubokého učení je jeho nejednotný rámec a struktura. Graf je sestaven z informací na základě prostudované literatury a dalšího hledání.

DL - deep learning

Frameworks and libraries

- Keras
- TensorFlow
- Caffe
- Theano
- Torch
- PyTorch
- MXNet
- Deeplearning4j
- Microsoft Cognitive Toolkit

Architectures

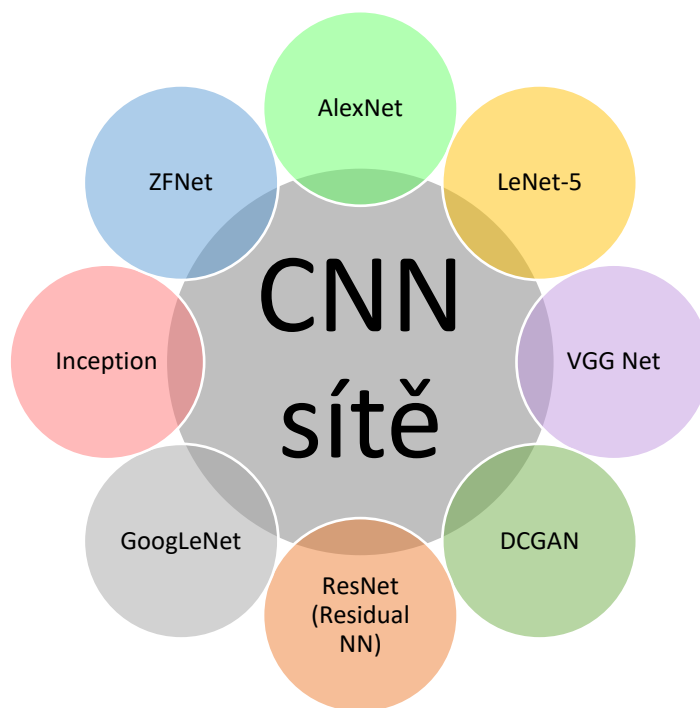
- CNN
- DNN - Deep NN
- RNN - Recurrent NN
- DBNs - Deep Belief Networks
- Autoencoder (stacked, sparse)
- Adversarial Networks
- RBMs - Restricted Boltzman Machines
- LSTM - Long Short-Term Memory
- RBF - Radial Basis Function Network

Algorithms

- (Batch) Gradient Descent
- Stochastic Gradient Descent
- Back Propagation through Time
- Momentum
- Levenberg–Marquardt algorithm

Graf 3: DL a jeho dělení Zdroj: vlastní

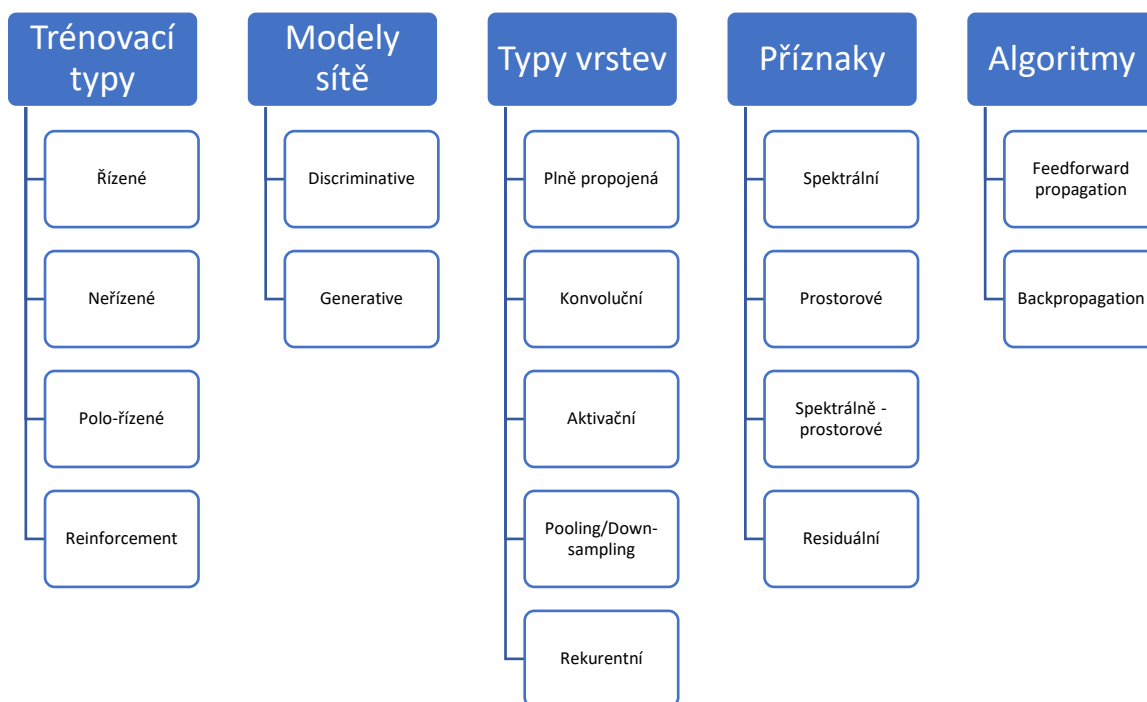
CNN jsou progresivní metodou v řadě aplikací. Jedná se o vícevrstvé algoritmy hlubokého učení, jejichž charakteristickým znakem jsou více než dvě skryté vrstvy (*hidden layers*). Proto se nazývají hluboké CNN (*deep CNN – DCNN*). Úkolem CNN je zredukovat obraz tak, aby bylo jednodušší ho zpracovat, ale bez ztráty příznaků. Výhodou je, že není třeba přílišné předzpracování jako u ostatních klasifikačních algoritmů (Vali, A., et al., 2020). CNN může být trénováno tak, aby lépe porozumělo náročnosti obrazu a samo se učilo. Architektura redukuje počet zapojených parametrů a znovu použitelnost vah pro lepší přizpůsobení se obrazu. Stejně jako NN i CNN sestávají z několika typů vrstev. S přibývajícimi vrstvami architektura zachycuje také *high-level* příznaky, což je i cílem CNN. Jedním z rozdílů mezi NN a CNN je, že CNN mají neurony uspořádány ve třech dimenzích – šířka, výška a hloubka. Dalším rozdílem je, že architektury konvoluční neuronové sítě předpokládají, že vstupními daty jsou obrazy, které umožňují zakódovat určité vlastnosti do architektury. Ty napomáhají redukovat množství parametrů v síti. Modely CNN jsou vypsány v Graf 4



Graf 4: Modely CNN. Zdroj: vlastní

Výše uvedená teorie je graficky doplněna nejen o jednotlivé modely daných sítí, ale i o jejich další dělení (Graf 5). Například zajímavým je dělení podle příznaků, které je i inspirací pro následující rozdělení článků.

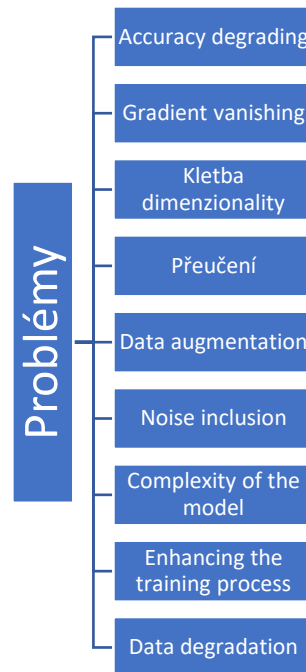
Většina dále rozebraných metod je implementována v programovacím jazyce Python, menší množství poté v SW Matlab. Pro Python je dostupné i množství knihoven (Graf 3), které



Graf 5: Další dělení DL. Zdroj: vlastní

jsou otevřené a lze je využít k inspiraci nebo jako podklad k práci. Pro tuto práci je vybrána TensorFlow v aplikaci Keras a PyTorch.

Graf 6 ukazuje opakovaně řešené problémy při práci s CNN. Jejich komplexní rozbor a možná řešení přesahují rámec této práce. Podrobněji budou některé problémy a jejich překonávání uvedeny v metodice této diplomové práce.



Graf 6: Opakovaně řešené problémy CNN. Zdroj: vlastní

Úvodní náhled do problematiky DL poskytuje článek Vali, A., et al. (2020), kteří se zabývají metodami DL, které v současné době překonávají ostatní přístupy strojového učení. Jsou předmětem velkého zájmu a využití i v oblasti klasifikace *land use* a *land cover* s využitím hyperspektrálních a multispektrálních (MS) snímků. Autoři poskytují přehled výzkumných trendů, zásadních pojmů, charakteristik vstupních dat a dostupných datových sad. Podrobněji rozebírají nejmodernější metody, aby lépe porozuměli, jak je možné tyto techniky adaptovat a přizpůsobit kontextu dálkového průzkumu Země. V návaznosti na to se snaží zodpovědět některé výzvy a problémy. Jednou z velkých výzev je omezené množství trénovacích dat (*limited ground truth*). Pro vyřešení tohoto problému navrhuje několik metod, například rozšíření (např. Geometrické transformace) existujících dat (*data augmentation*), *Generative adversarial networks* (GAN), využití již natrénované CNN (*transfer learning*) a neřízené učení (*unsupervised learning*). Zaměřují se i na další úkoly a výzvy, ke kterým se v literatuře snaží najít množství přístupů a řešení. Poukazují i na fakt, že chybí zavedený rámec DL na poli DPZ, který by modely optimálně kategorizoval a seskupoval. Studie je i velmi cenným a bohatým zdrojem dalších článků věnujících se konkrétním metodám nebo problematice DL obecně.

Nezapomínají zmínit ani možné novinky a směry rozvoje do budoucna. Jednou z budoucích cest je přechod na cloudové platformy nejen dat, ale i samotných výpočetních úkonů.

Pouyanfar, S., et al. (2018) poskytují přehled hlubokého učení z různých perspektiv včetně historie, výzev, příležitostí, algoritmů, rámců, aplikací a dalších. Také představují a porovnávají známé platformy otevřených zdrojových kódů – Caffè, DeepLearning4j, TensorFlow, Theano, a PyTorch. Podávají i komplexní přehled historických a současných nejmodernějších přístupů v různých oblastech DL (vizuální, audio, textové zpracování a další). Autoři neopomenuli ani rozebrat problémy, kterým DL čelí (nedostatek trénovacích dat, black box a další), aby tyto výzvy mohly být dále zkoumány a řešeny pro budoucí výzkum. V neposlední řadě zmiňují i tři hlavní výhody CNN – sdílení parametrů, méně interakcí (*sparse interactions*) a rovnocenné reprezentace (*equivalent representations*).

Vzhledem k předpokladu implementace vybrané metody v programovacím jazyce Python, je velmi přínosným článkem Hluboké učení s Python (Chollet, F., 2018). Přehledně vysvětluje, co je hluboké učení, umělá inteligence a strojové učení. Podrobně se věnuje DL, krátce pojednává o historii a hardwaru, jakých úspěchů se doposud v tomto oboru dosáhlo a proč zrovna nyní se stává velmi používaným. Jsou rozebrány i jednotlivé větve hlubokého učení – řízené, neřízené, polo řízené a učení pomocí zpětné vazby (*reinforcement learning*). Poskytuje náhled i do matematického pozadí NN – 1D, 2D a 3D. Rozebírá anatomii NN, jejich vrstvy, modely, konfigurace procesu a další. Představuje i Keras (DL rámec pro Python), který poskytuje pohodlný způsob, jak definovat a trénovat téměř jakýkoliv druh modelu hlubokého učení. V neposlední řadě se věnuje trénovacím, validačním a testovacím data setům i ve spojitosti s některými problémy jako jsou přeučení sítě (*overfitting*) a nedotrénování (*underfitting*). Nakonec jsou navrženy možné budoucí cesty, kterými by se DL mohlo vydat. Jednu z možných cest autor vidí ve větší samostatnosti architektur, které by měly být schopny se ještě více samostatně učit bez zásahu člověka. V návaznosti na to by se většina práce a dat měla přenést na cloudová úložiště, kde by měly být informace, data a modely snáze dostupné a sdílené.

Stejně jako předešlí autoři i Zhu, X., et al. (2017) se věnují DL v DPZ a jejich komplexnímu přehledu a seznamu zdrojů. Vyzdvihují hluboké učení jako jeden z rychle rostoucích trendů v big data analýzách a užitečný nástroj ke zpracování HS dat. Předmětem zájmu jsou především řízené hluboké neuronové sítě (*Deep Neural Network – DNN*) jejichž hlavním modelem jsou CNN. Avšak i samotné CNN zahrnují nespočet metod a přístupů, přičemž ty nejúspěšnější architektury z nich jsou schopny získávat *high-level*, hierarchické a abstraktní příznaky, které jsou robustnější vůči nelineárnímu zpracování (*nonlinear processing*). Zastoupeny jsou

i dostupné zdroje včetně tutoriálů a otevřených zdrojů DL rámců. Důležitým faktem je, že hluboké učení otevírá možnosti pro nové aplikace, ale i nástroje pro zpracování velko-měřítkových problémů.

Součástí DL je i plně propojená CNN zvaná U-Net. Původně byla navržena pro segmentaci biomedicínských snímků. Díky dobrým výsledkům byla následně metoda převzata i do dalších oborů. Hlavní myšlenka metody U-Net je postavena okolo problému segmentace, kde je nutné převést příznaky jednotlivých obrazových elementů (*feature map*) na vektor, ale opět z nich obraz zpět zrekonstruovat. Klíčem jsou naučené obrazové feature mapping (U-Net, 2019). Více o metodě a její architektuře v kapitole 4.3.

Obecný princip metody U-Net s využitím nad biomedicínskými snímky ve svém článku představují i Ronneberger, O., et al. (2015). Jak tvrdí, metoda závisí na využití augmentaci trénovacích dat pro jejich lepší využití. Ukazují, že síť může být učena end-to-end cestou s využitím pouze malého množství trénovacích dat. Staví na plně konvoluční architektuře, která je rychlá a překonává většinu metod ve své oblasti. Rozšiřují a upravují architekturu tak, aby byla schopna pracovat s malým množstvím snímků, a přesto byla schopna přesných výsledků. Metodu testují na medicínských snímcích včetně jejich specifických hodnotících parametrů.

Využití metody U-Net v DPZ, především v oblasti mapování vegetace, se v posledních letech zvýšilo a zabývá se jím i několik autorů. Například autoři Zhao, X., et al. (2019) se zabývají využitím metody U-Net pro zemědělské účely, konkrétně pro hodnocení zasazování rýže (*rice lodging assessment*). Pro sběr MS snímků využívají UAV. Z výsledků vyplývá, že lepší přesnosti dosáhli při využití rovnou nasnímaných snímků bez využití ručně získaných příznaků.

Oproti tomu autoři Wagner, F., et al. (2019) se snaží získat informace o druhové skladbě lesa na regionální úrovni. K tomuto účelu využívají architekturu metody U-Net nad snímky World-View-3 s velmi vysokým rozlišením 30 cm. Podařilo se jim dosáhnout OA přes 95%.

Stejně tak i Kattenborn, T., et al (2019) využívají RGB data vysokého rozlišení získaná pomocí UAV společně s využitím trénovacích dat pro mapování a segmentaci snímků vegetace. Jejich dosažená přesnost metodou U-Net se pohybuje nad 84%.

2.2 Hyperspektrální data

Hyperspektrální data jsou definována jako data s velkým množstvím spektrálních pásem (ve stovkách), která jsou velmi úzká (méně než 10nm), a navazují na sebe. HS data se vyznačují množstvím spektrální informace, kterou obsahují. V kombinaci s obsaženou prostorovou informací umožňují podrobnější popis zemského povrchu. Mimo množství výhod a bohatství informací mají HS snímky i množství problémů. Jedná se především o vysokou rozměrnost dat

a omezenou dostupnost trénovacích vzorků. Vliv na přesnost klasifikátorů, mimo výše zmíněné překážky, má i vysoká mezitřídní variabilita a podobnost klasifikovaných objektů a jevů. V poslední době je prezentováno množství DL architektur s cílem vyřešit nebo zmírnit tato omezení. Pohledů na HS data je velké množství, stejně jako článků, které se této problematice věnují. Pro účely této práce je vybráno několik článků, které jsou rozebrány z pohledu použití dat, předzpracování, vstupu do klasifikace, rozlišení, senzorů apod.

Jak vyplývá z tabulky (Příloha 1), většina autorů současných moderních metod používá pro prezentování výsledků stejné datové sady, což umožňuje alespoň částečně jejich porovnání. S tím souvisí i využití standardních hodnotících metod jako jsou OA , AA a kappu koeficient. V následujících článcích je zmíněno i případné předzpracování a transformace dat společně s počtem pásem vstupujících do sítě. Bez jedné metody jsou všechny implementovány v programovacím jazyce Python. To vede k závěru, že tento jazyk s jeho otevřenými knihovnamy je vhodný pro zpracování HS dat.

Li, S., et al. (2019) se věnují HS snímkům a jejich charakteristice. Poukazují na jejich specifickou (např. množství spektrálních pásem), kvůli níž je i jejich klasifikace tradičními metodami strojového učení náročná. Navíc se HS zobrazování často potýká s vnitřními nelineárními vztahy mezi zachycenou spektrální informací a odpovídajícími materiály. Efektivně řešit nejen nelineární problém, ale také automaticky se učit *high-level* příznaky z komplexních HS snímků dokáže DL. Učící proces je zcela automatický, což dělá hluboké učení mnohem vhodnější pro zvládnání různých situací (například velkou prostorovou variabilitu spektrálních podpisů nebo malé množství trénovacích dat). V článku jsou rozebrány hlavní výzvy klasifikace HS dat a výhody DL pro jejich řešení. Autoři se snaží vybudovat rámec, který by ukázal dosavadní úspěchy DL klasifikací. Jejich rozdělení je založeno na informacích získávaných z HS dat – spektrální, prostorové a spektrálně – prostorové. V neposlední řadě berou v úvahu fakt, že počet trénovacích vzorků je většinou velmi omezený a zahrnují několik strategií, jak tento problém řešit a zlepšit tak i výsledky klasifikace. V některých aplikacích je hluboké učení považováno za černou skříňku (*black box*), kde vnitřní informace sítě jsou ne vždy zcela jasné. Proto je prozkoumání vnitřních vztahů a vlastností velmi užitečné pro analýzu výkonu sítě a další navrhování hlubokých architektur. Nakonec je několik nejmodernějších metod, jako jsou například 3D-CNN, *Gabor-CNN*, *Siamese CNN* a DFFN (*deep feature fusion network*) porovnáno nad reálnými HS daty.

2.3 Přehled architektur hlubokého učení pro klasifikaci HS dat

Následující články se věnují problematice CNN pro klasifikaci HS dat. Snaha je podat co nejucelenější přehled postupů a následně rozebrat konkrétní architektury DL od těch základních až po ty méně tradiční. Z načtené literatury je z desítek až stovek zmiňovaných metod vybráno 20 příkladů. Jejich výběr je založen na informacích z přečtených článků a snahou je postihnout rozsah metod hlubokého učení vztažených k CNN. Cílem je touto cestou poznat jednotlivé přístupy a vytipovat nejvhodnější pro tuto práci. Vzhledem k rozsáhlosti tématu je vybrán jen daný počet příkladů, které jsou rozděleny do 4 kategorií podle příznaků (spektrální, prostorové, spektrálně-prostorové, residuální). Jsou přidány ještě další dva parametry, a to způsob trénování (řízené, neřízené, polořízené, reinforcement) a typ architektury (CNN, DNN, 3DCNN, ResNet, RNN a další).

Klasifikace nebo také sémantická segmentace obrazu je proces, kterým se dosahuje automatického vyhodnocení dat DPZ. Zapojení výpočetní techniky do vytváření klasifikačních pravidel, a následně i samotné klasifikace, má několik výhod. Počítačové zpracování je schopno rychle a opakovaně zpracovat i velké objemy dat a lépe zvládá větší dynamický rozsah dat (Halounová, L., et al., 2009). „*Podstata klasifikace spočívá v rozdělení naměřeného datového souboru do určitých tříd podle určitého příznaku (textura, barva, tvar, ...) a na základě určitých rozhodovacích pravidel (klasifikátoru). Klasifikátor je pak sada pravidel (algoritmus), která na základě příznaků rozhoduje o přiřazení zkoumaného pixelu do dané třídy. Třídami jsou v DPZ nejčastěji jednotlivé typy krajinného pokryvu v různé úrovni podrobnosti. Příznak je veličina, kterou můžeme pro jednotlivé objekty měřit, a která by měla především jednoznačně odlišovat jednotlivé třídy. Rozeznáváme čtyři základní příznaky: spektrální, prostorový, časový a polarizační.*“ (Červená, L., 2016); (Kolář, J., 1990).

První soubor článků vychází ze spektrálních příznaků, které Kolář, J. (1990) charakterizuje jako „*odrazivé a vyzařovací vlastnosti zkoumaného povrchu. Mohou jimi být hodnoty intenzity záření naměřené v jednotlivých spektrálních intervalech, jejich lineární kombinace nebo celková intenzita ve všech kanálech.*“

Nejčastější úlohou v DPZ je v dané lokalitě určit v různé míře podrobnosti typ krajinného pokryvu, jeho stav a případně dynamiku. V získávání těchto informací se tradiční CNN metody potýkají s řadou problémů, jako je například slabá generalizační schopnost modelu nebo automatizované sestavování trénovacích data setů. Následující metody se snaží těmto problémům předcházet různými postupy, které jsou níže popsány. Všechny dále zmíněné články v této kategorii využívají řízené klasifikace.

Autoři Hu, Y., et al. (2018) navrhuje nový typ DCNN, který je primárně určen pro družicová multispektrální data Landsat-8 OLI a hyperspektrální data. Předkládají nový technický rámec, který kombinuje dva typy vrstev pro extrakci informace o krajinného pokryvu. Prvním typem vrstvy je *cascaded cross-channel parametric pooling*, která je v podstatě konvoluční vrstvou a má za úkol začleňovat spektrální informace z jednotlivých pásem. Druhou použitou vrstvou je *global average pooling* jejímž úkolem je zachovat pouze průměrné hodnoty každého použitého spektrálního pásma. V dalším kroku metoda sestavuje DCNN model, který je schopný automaticky sestavovat trénovací data sety a klasifikovat obrazy. Použití DNN zjednodušuje aplikaci modelu a zlepšuje jeho schopnost generalizace. Na druhou stranu tato metoda vyžaduje větší počítačový výkon než tradičnější klasifikační metody.

Oproti předešlé metodě, autoři Hu, W., et al. (2015) využívají DCNN ke klasifikaci hyperspektrálních snímků přímo ve spektrální oblasti. Metoda využívá pět vrstev (vstupní, konvoluční, *max pooling*, plně propojenou a výstupní) s váhami, které jsou realizovány na každém spektrálním podpisu, aby byly rozlišitelné od ostatních.

Další variantu CNN zvanou plně konvoluční síť (*Fully Convolutional Network – FCN*) využívají Fu, G., et al. (2017) pro přesnou klasifikaci snímků s vysokým rozlišením. V jejich přístupu kombinují dvě fáze, segmentační a klasifikační. V prvním kroku pomocí *Atrous convolution* zlepšují hustotu výstupních klasifikačních map (*density of output class maps*) a následně přidávají strukturu vynechané vrstvy (*skip-layer structure*), aby byl model schopen klasifikovat obrazy různých rozlišení. V poslední části dále vylepšují výstupní klasifikační mapu s použitím *CRFs (Conditional Random Fields)* postprocesingu. V porovnání s *patch based* (k odstranění šumu z obrazu využívá tato technika *patches* (bloky), do kterých daný snímek rozdělí, a pak je zpracovává odděleně s cílem odhadnout pravou hodnotu pixelu) přístupy berou jako vstup celý obraz.

Rekurentní neuronové sítě (*Recurrent Neural Networks – RNN*) jsou další variantou neuronových sítí, které jsou značně využívány, protože extrahují kontextuální informaci ze sekvence. V článku Wu, H., et al. (2017) studují schopnost RNN klasifikovat HS data pomocí kontextuální informace, kterou získávají z dat. Konkrétně jsou HS data zpracovávána jako spektrální sekvence (*spectral sequences*) a RNN modelují závislosti mezi různými spektrálními pásmy. Navíc navrhuje použít konvoluční rekurentní neuronovou síť (*Convolutional Recurrent Neural Network – CRNN*), která je hybridem CNN a RNN, a získat tak více rozlišitelných příznaků (*discriminative features*) pro klasifikaci. V CRNN je nejprve učeno několik konvolučních vrstev k získání *middle-level* (například prostorové pyramidy nebo spektrální závislosti) a lokálně neměnných (*locally-invariant*) příznaků ze vstupních dat.

Zároveň s nimi jsou použity i *pooling* vrstvy, které díky podvzorkování (*subsampling*) dělají sekvenci kratší. Následně jsou použity rekurentní vrstvy, které z příznaků získaných z konvolučních vrstev dále těží spektrálně-kontextuální informace. Nakonec je výsledek klasifikace ještě zlepšen pomocí prostorově založených příznaků.

Další skupina článků využívá prostorových příznaků, které Kolář, J. (1990[20]) definuje jako „*plošnou homogenitu naměřené veličiny, která určuje hranice objektu a tím následně i jeho tvar a velikost. Důležitým prostorovým příznakem jsou parametry vyjadřující vnitřní texturu objektu.*“

Tématu se věnují Rezaee, M., et al. (2018) s článkem o adaptaci známé CNN – AlexNet pro mapování mokřad z DPZ snímků. Snaží se řešit problém s nedostatkem příznaků a trénovacích dat. První vrstvy získávají základní atributy jako jsou hrany a poslední vrstvy těží *high-level* příznaky, například objekty, větší tvary a texturu. Začlenění *high-level* příznaků do klasifikace má velký význam, protože napomáhají snižovat možnou záměnu mezi spektrálně podobnými třídami mokřadu. I z tohoto důvodu jsou využívány i prostorové příznaky (tvar, velikost a umístění), na kterých závisí výše zmíněné *high-level* příznaky.

Oproti předešlému článku autoři Krizhevsky, A., et al. (2017) využívají ke klasifikaci ImageNet. Významným příspěvkem tohoto článku je, že se jim podařilo natrénovat jednu z největších CNN sítí současnosti. Klasifikovali 1,2 milionu snímků s vysokým rozlišením do 1 000 různých tříd. Využili neuronovou síť, která se skládá z pěti konvolučních vrstev, následovaných max-pooling vrstvami a třemi plně propojenými vrstvami. Ke snížení přeučení v plně propojených vrstvách využívají novou metodu zvanou *dropout*, která se ukázala být užitečná. Jejím cílem je zabránit přeučení sítě (*overfitting*). Toho se snaží dosáhnout tím, že zabrání utlumení či vypnutí slabších spojení během procesu učení.

Následující článek (Springenberg, J., et al., 2015) se zabývá otázkou, které složky v CNN jsou pro rozpoznání objektů opravdu nezbytné. Došli k závěru, že max-pooling vrstvy mohou být jednoduše nahrazeny konvolučními vrstvami se zvýšeným krokem bez ztráty přesnosti. Jejich navržená architektura sestává pouze z malých konvolučních vrstev (jejich pooling velikost je menší než 5), které dokážou snížit počet parametrů, a tím zajistit určitou pravidelnost sítě. Ke znázornění příznaků, které jsou naučeny CNN, slouží nově představená varianta dekonvolučního přístupu (*deconvolution approach*).

Velkou a významnou roli poslední doby hraje společné využití spektrálních a prostorových příznaků pro klasifikaci. Definice jednotlivých příznaků jsou zmíněny výše. Tyto příznaky jsou většinou získávány třemi cestami. Mapováním z *low-level* příznaků na *high-level* příznaky skrz hluboké síť. Nebo také přímou extrakcí hlubokých příznaků z originálních dat či spojením

dvou oddělených hlubokých příznaků. Charakteristika těchto příznaků je na začátku kapitoly 2.1 Hluboké učení.

Na pomezí využití těchto příznaků stojí článek, ve kterém je poprvé využito užitečnosti a efektivnosti GAN (*Generative Adversarial Network*) pro klasifikaci HS snímků. V navržené metodě (Zhu, L., et al., 2018) je několik CNN trénováno dohromady. Generativní CNN se snaží vytvářet falešné vstupy, které jsou co nejvíce reálné. Snahou druhé rozlišující CNN je klasifikovat skutečné a falešné vstupy. Tato řízená klasifikace využívá dobře navrženou 1D – GAN jako spektrální klasifikátor a robustní 3D – GAN jako spektrálně prostorový klasifikátor. Díky “soupeřícímu“ (*adversarial*) trénování je zlepšena generalizační schopnost rozlišující CNN, což je důležité při omezeném množství trénovacích vzorků. Vysoká rozměrnost vstupních dat je redukována pomocí metody hlavních komponent (*principal component analysis – PCA*) a napomáhá ustálení trénovacího procesu.

Následující dva články vycházejí ze zpracování jedno rozměrných dat. Autoři Zhang, H., et al. (2017) navrhují novou DCCNN (*dual-channel convolutional neural network*) metodu pro přesnou klasifikaci HS dat s využitím spektrálně-prostorových příznaků. Hierarchické texturální příznaky jsou automaticky extrahovány pomocí jedno-dimenzionální CNN a dvou-rozměrná CNN je použita k získávání hierarchických prostorových příznaků. Nakonec je použit *softmax regression* klasifikátor, který spojuje texturální a prostorové příznaky, a nakonec předpovídá výsledky klasifikace. Jednoduchá metoda rozšíření dat je použita k řešení problému s omezeným množstvím trénovacích vzorků.

Oproti tomu hlavní myšlenkou Luo, Y., et al. (2018) je rozpoznat data za využití vzájemného vztahu mezi konvolučními výsledky a spojením jednorozměrných dat do dvou rozměrných, podobných obrazu, s cílem prohloubit strukturu sítě. Tyto kroky by měly napomoci síti lépe získávat a rozlišovat příznaky. Navržená síť zvaná *HSI-CNN* v prvním kroku získává strukturálně-prostorové příznaky z cílových pixelů a jejich sousedů. Poté jsou jednorozměrné příznakové mapy získané konvoluční operací z texturálně – prostorových příznaků spojeny do dvourozměrných matic. Tyto matice považované za obrázky jsou vstupem do klasické CNN, která obsahuje pět typů vrstev (vstupní, konvoluční, pooling, plně propojenou a výstupní).

Další články se věnují již o něco složitějším modelům. Například Yu, S., et al. (2017) navrhují účinnou CNN architekturu k posílení diskriminační schopnosti (*discriminative capability*) pro klasifikaci HS snímků. V modelu jsou originální data použita jako vstup a finální výstupy z CNN jsou očekávané výsledky související s třídami (*class-related*). Pro lepší využití informace z HS dat používají 11 konvolučních vrstev společně

s pooling vrstvou a většími *dropout* hodnotami. S problémy spojenými s analýzou HS snímků se vypořádává pomocí *end – to – end* metody. Kvůli nedostatku trénovacích dat se snaží předejít problému přeučení například pomocí rozšíření dat, většími *dropout* hodnotami a dalšími přístupy.

Další řízená *end – to – end* síť je navržena autory Wang, W., et al. (2019). Jejich AUSSC (*alternately updated spectral–spatial convolutional*) síť s *recurrent feedback structure* je navržena k učení se očištěných texturálních a prostorových příznaků pro klasifikaci HS dat. Jejich síť zahrnuje bloky, ve kterých každá vrstva slouží zároveň jako vstup a výstup pro ostatní vrstvy. AUSSC je schopna na základě stanovených parametrů několikrát očistit texturální a prostorové příznaky. Výhoda navržené architektury spočívá v rozkládání velkých 3D konvolučních jader (*kernels*) na tři menší 1D konvoluční jádra. Tím šetří velké množství parametrů a snižuje přeučení.

Příkladem další *end – to – end* sítě, která by měla řešit běžné problémy spojené s CNN, je *deconvolution network with skip architecture* s cílem naučit se texturální a prostorové příznaky (Ma, X., et al., 2018). Navržená síť vychází ze dvou větví. V prostorové větvi je navržena vrstva, která vybírá pásma a omezuje množství parametrů a poukazuje na problém přeučení. Ztracené informace z pooling vrstvy jsou obnoveny pomocí *unpooling* a *deconvolution* operací a učí se hierarchicky prostorové reprezentace po pixelech. V texturální větvi je kontextuální hluboká síť použita k získávání hlubokých spektrálních příznaků.

Hlavním smyslem neřízeného učení je získat užitečné příznaky z neoznačených dat, odhalit a odstranit nadbytečné (*redundancies*) a zachovat pouze podstatné aspekty dat (Mou, L., et al., 2017). V tomto článku je prezentována nová architektura – *fully Conv-Deconv network*, která může být učena v *end – to – end* směru. Jedná se o neřízené učení spektrálně – prostorových příznaků z HS obrazů. Nejprve je pomocí kodéru (*encoderu*) (konvoluční subsítě) vstupní 3D hyperspektrální *patch* transformováno do nižšího dimenzionálního prostoru. Následující krok reprodukuje počáteční data pomocí dekóderu (dekonvoluční subsítě). Některým problémům předchází pomocí *residual learning* a nové *unpooling* operaci. Zabývají se i důkladným prozkoumáním naučených příznaků.

Poslední tři články této kategorie se zabývají 3D povahou HS dat a jejich prostorovými vlastnostmi (pravidelnost a plynulost). Je navrženo množství 3D operátorů extrahujících příznaky, které jsou navrženy k získávání spojených texturálně-prostorových informací. Současné výzkumy ukazují, že využití těchto propojených informací může výrazně zlepšit výsledky klasifikací. Problémem zůstává, že kvůli velkému množství získaných příznaků může

architektura trpět takzvanou kletbou rozměrnosti (*curse of dimensionality*), zejména pokud je k dispozici malý trénovací data set.

První, kdo řeší výše zmíněnou problematiku jsou autoři Zhu, J., et al. (2018), kteří navrhují M3DF3 (*multiple 3D feature fusion framework*) pro klasifikaci hyperspektrálních obrazů. Nejprve jsou 2D *Gabor* povrchové příznaky rozšířeny na 3D (*3DSF*) oblasti, které vyhovují SS (*spektrálně-prostorové*) struktuře HS dat a jsou aplikovány přímo na originální data. Následně jsou spojeny dohromady tři metody extrakce 3D prvků (3D metody získávání příznaků) za použití *multitask sparse representation framework*, aby plně využili potenciál spojených násobků 3D příznaků. Zmíněné metody jsou konkrétně: 3D *morphological profile*, 3D *local binary pattern* a navržená 3DSF, která charakterizuje HS snímky ze tří různých pohledů – morfologie (tvary), místní závislosti (*local dependence*) a hladkost tvaru (*shape smoothness*).

Následující metoda se zaměřuje na přesnou klasifikaci HS dat pomocí 3D konvoluční neuronové sítě. Tato architektura se nespolehá na předzpracování ani na postprocessing a využívá celé HS krychle (*cubes*), díky čemuž je schopna extrahovat hluboké texturně-prostorové příznaky efektivně (Li, Y., et al., 2017). Na základě toho potřebuje méně parametrů oproti ostatním hlubokým metodám. Proto je model lehčí, méně náchylný k přeučení a jednodušší k učení. 3D-CNN v této metodě se skládá pouze z konvolučních a plně propojených vrstev. Účinnost klasifikace je ovlivněna třemi faktory: počtem jader, spektrální hloubkou jader a prostorovou velikostí vzorků ke klasifikování.

Autoři Zhong, Z., et al. (2017) navrhují řízenou *end – to – end* SSRN (*spectral-spatial residual*) síť s nepřetržitými učícími bloky (*learning blocks*), které berou v úvahu i charakteristiky hyperspektrálních snímků. Za rozšíření konvoluční vrstvy v CNN jsou považovány navržené spektrální a prostorové residuální bloky, které získávají rozlišující SS příznaky z krychlí HS snímků. Na každou konvoluční vrstvu zavádí *batch normalization* s cílem regulovat trénovací proces a zlepšit klasifikační výkon trénovaného modelu.

Hluboké modely si k získání příznaků osvojují velké množství hierarchických vrstev, což může vést k některým negativním efektům jako jsou přeučení, mizení sklonu (*gradient vanishing*) a snižování přesnosti. Se snahou předejít výše zmíněným problémům je navržena DFFN síť pro klasifikaci HS snímků. Autoři Song, W., et al. (2018) uvažují silnou doplňkovou a stále i korelovanou informaci mezi různými hierarchickými vrstvami. K dalšímu zlepšení klasifikace tak využívají spojených výsledků z různých hierarchických vrstev. Pomocí zbytkového učení (*residual learning*) je optimalizováno několik konvolučních vrstev jako totožné mapování (*identity mapping*), které může usnadnit učení hluboké sítě a prospívat

z rostoucí hloubky. Díky tomu je možné vystavět opravdu hlubokou síť a získat více rozlišujících příznaků hyperspektrálních obrazů.

Další možností je jít hlouběji a více do šířky než u ostatních hlubokých sítí pro klasifikaci HS dat s metodou CCNN (Lee, H., et al., 2017). Tuto metodu nazývají kontextuální hluboké CNN. Výhoda tohoto postupu spočívá v jeho schopnosti optimálně zkoumat lokální kontextuální interakce díky společnému využití lokálních prostorově – spektrálních vztahů mezi jednotlivými sousedícími pixelovými vektory. Tohoto společného využití prostorově – spektrálních informací je dosaženo pomocí mnohaměřítkové (*multi-scale*) konvoluční filtrační bance, která slouží jako výchozí komponent navrhované CNN. Získané mapy prostorových a texturálních příznaků jsou následně spojeny dohromady. Tento výsledek je následně vstupem pro plně konvoluční síť. Navrhovaná síť je složena ze tří klíčových částí: plně konvoluční síť z devíti vrstev, mnohaměřítková filtrační banka a zbytkové učení.

V neposlední řadě Chen, Y., et al. (2017) představují jednoduchou a efektivní DPN (*Dual Path Network*) metodu. Autoři došli k závěru, že ResNet umožňuje znovu využití příznaků a DenseNet umožňuje prozkoumání nových příznaků, což je obojí důležité pro učení se dobrých reprezentací. S cílem využít výhody obou topologií cest, navržená síť sdílí společné příznaky, zatímco zachovává flexibilitu pro objevování nových cest architektury. DPN se také těší vyšší účinnosti parametrů, nižší výpočetní náročnosti a menšímu využití paměti. Je též přívětivější k optimalizaci oproti moderním klasifikačním sítím. Tato architektura dědí výhody ze zbytkových a hustě propojených (*densely connected*) cest, umožňující opětovné efektivní použití příznaků.

Následující část rešerše je věnována užšímu výběru metod, které jsou podrobněji rozebrány a z nich je následně snaha vybrat jeden nejvhodnější přístup.

I autoři Shrestha, A., et al. (2019) se zabývají DL a jeho rolí v klasifikaci snímků DPZ. Poskytují důkladný přehled neuronových sítí a hlubokých neuronových sítí i s jejich pochopitelným vysvětlením. Poskytují hlubší pohled na známé tréninkové algoritmy a architektury a poukazují na jejich případné nedostatky. Příklady takových problémů jsou například nenalezení globálního, ale lokálního minima ztrátové funkce, přeučení a trénovací čas pro velké problémové datasey. Jsou vypsány i některé nejmodernější způsoby, jak překonat tyto výzvy pomocí různých optimalizačních metod a dojít tak i ke zlepšení přesnosti a snížení výpočetního času. Autoři také prostudovali několik současných studií a prezentují jejich nedostatky, implementace a zlepšení tréninkového procesu. V neposlední řadě rozebírají i různé typy hlubokých architektur jako jsou DCN, DRN, RNN, *reinforcement learning*,

variational autoencoders a další. Vše výše zmíněné je přehledně shrnuto v tabulkách, které poskytují přehled toho, jak spolu různé aspekty DL souvisejí.

K doplnění celkového obecného obrazu o vybrané problematice slouží článek od autorů Scott, J., et al. (2017), kteří se věnují tématu využití DCNN pro klasifikace krajinného pokryvu ze snímků DPZ s vysokým rozlišením. Konkrétně se zaměřili na jeden z hlavních problémů, se kterým se CNN potýkají. Ačkoliv je k dispozici velké množství leteckých a satelitních snímků, chybí dostatek bohatých trénovacích informací a dat. S cílem předejít tomuto problému jsou použity dvě techniky: přenos poznatků (*transfer learning – TL*) a rozšíření dat (*Data augmentation – DA*) navržené speciálně pro zlepšení trénovací sítě DPZ snímků. DA je schopno využívat různé vlastnosti DPZ snímků pro rozšíření malých trénovacích data setů a ke zlepšení robustnosti DCNN. Využití TL s případnými úpravami (*fine-tuning*) a DA bude získávat na důležitosti s rostoucí komplexností modelů DL sítí. Vybrané techniky jsou testovány na známých data setech UCM (UC Merced) a RSD. Dosahují vysokých přesností okolo 97-98% OA (*overall accuracy*) s CaffeeNet, GoogleNet a ResNet.

Stejně tak jako ostatní autoři i Paoletti, M., et al. (2019a) se snaží přijít s metodou, která předejde překážkám, které se vyskytují u architektur CNN. Jedním z již zmíněných problémů je omezené množství dat pro trénovací účely. Druhým z nich je efektivní využití vztahů mezi spektrálně – prostorovými prvky HS dat, které jsou klíčovým faktorem pro řešení vysoké komplexnosti HS dat. Částečně tuto překážku řeší hluboké architektury CNN díky jejich potenciálu efektivně charakterizovat spektrálně – prostorové vlastnosti. Nicméně i tyto architektury se potýkají s výzvami týkajícími se slučování parametrů sítě. To má za následek omezení výkonu klasifikace v případě náročných scénářů. Autoři představují novou metodu založenou na spektrálně – prostorové kapslové síti s cílem dosáhnout vysoce přesné HS klasifikace a zároveň výrazně snížit složitost navržené sítě. Konkrétně je na základě Hintonových kapslových sítí vyvinuto rozšíření CNN modelu, které přeformulovává koncept kapslových jednotek tak, aby se staly spektrálně – prostorovými celky specializovanými na klasifikaci dálkově snímaných HS dat. Zpracování HS dat navrženým modelem je následující. Nejprve jsou data zpracována kódovací (*encoder*) sítí, která sestává ze tří vrstev (klasická konvoluční, primární kapslová a hustá kapslová), které extrahují příznaky a fungují jako klasifikátor. Následuje fáze sítě dekodovací (*decoder*), která je složena z několika plně konvolučních vrstev, a která zlepšuje klasifikaci pomocí rekonstrukce dat. Výhodou navrženého modelu je jeho schopnost naučit se spektrálně-prostorové vlastnosti HS dat s ohledem na jejich odpovídající prostorové pozice ve scéně, jejich přidružené spektrální popisy a jejich možné transformace. Představená metoda je schopna charakterizovat vstupní HS data

na vyšším abstrakčním stupni, což nakonec umožňuje podstatně snížit počet konvolučních vrstev a složitost modelu za použití omezeného množství trénovacích dat. Výhodou je také krátká doba potřebná k dosažení téměř optimálního výkonu. K ověření kvality navržené metody slouží porovnání s několika nejmodernějšími klasifikačními metodami (SVM, RF, MLP, deep MLP, 2-D CNN, 3-D CNN, SSRN, DFCNN) nad známými HS data sety (Indian Pines, Salinas Valley, Kennedy Space Center, IP scene). Metoda dosahuje stabilně lepších výsledků než 3D CNN metody s průměrným zlepšením okolo 2 % celkové přesnosti (*overall accuracy – OA*), průměrné přesnosti (*average accuracy – AA*) a kappa koeficientu. Použitým rámcem je PyTorch a programovacím jazykem je Python.

Následující studie se zabývá zlepšením klasifikace HS dat za pomoci S-CNN metody (řízená extrakce hlubokých funkcí založená na Siamese konvoluční neuronové síti). Cílem je zlepšit výsledky klasifikace za situace, že k trénování je k dispozici pouze omezené množství trénovacích dat. Konkrétně je metoda trénována ke zvýšení oddělitelnosti mezi různými třídami. Skládá se v první fázi z klasické CNN s pěti vrstvami, která přímo extrahuje hluboké prvky z HS krychlí. Následuje *siamese* síť složená ze dvou CNN, trénovaná k učení se prvků, které vykazují nízkou a vysokou mezitřídní variabilitu. Navržená metoda je řízená s *margin ranking loss* funkcí, která umí získávat více diskriminačních funkcí pro klasifikaci, což je jednou z důležitých charakteristik (Liu, B., et al., 2018). Vybraná metoda extrakce prvků je založena pouze na malém množství parametrů jako jsou rychlost učení (*learning rate*), parametr okraje (*margin parameter*) a množství epoch (*number of epochs*). Hlavními přínosy této studie je zaprvé CNN zahrnující konvoluční, max-pooling a tři plně konvoluční vrstvy navržené k přímé extrakci hlubokých spektrálně – prostorových prvků. Za druhé, S-CNN je trénováno s malým množstvím označených vzorků. Díky uvážení oddělitelnosti různých tříd, může síť lépe extrahovat prvky pro danou klasifikaci, což vede k lepším výsledkům. Nevýhodou S – CNN jsou zvýšené nároky na výpočet. K ověření přesnosti navrhované metody jsou vybrány tři známé HSI data sety (Pavia Center, University of Pavia, Indian Pines) a jejich extrahované prvky jsou použity jako vstupní data pro SVM klasifikátor. Porovnány jsou i další metody – EMPs (*Extended Morphological Profiles*), LDA (Linear Discriminant Analysis), řízené CNN, a CNN-PPF. Výsledky ukazují, že kombinace S-CNN a SVM dosahuje lepších výsledků než tradiční metody. Implementace vybrané metody je v TensorFlow knihovně.

Díky jejich dobrým výsledkům se CNN dostaly do vedoucí pozice mezi nejmodernějšími metodami hlubokého učení. Hned z několika důvodů, jako je vnitřní komplexnost HS dat, vysoká dimensionalita, nadbytečnost a šum vedoucí k neschopnosti zobecnit diskriminační spektrálně-prostorové prvky, a v neposlední řadě i přidávání vrstev, jsou prezentovány nové

výzkumy v oblasti DL se zaměřením na klasifikaci HS dat. Snaží se těmto překážkám předcházet a zlepšit jejich výsledky. Za účelem zmírnění těchto problémů je prezentován článek od autorů Paoletti, M., et al. (2019b), který navrhuje nový model ResNet založený na *pyramidal bottleneck residual units* s využitím spektrálních i prostorových informací. Tento nový hluboký model obsahuje tři typy vrstev (konvoluční (CONV), *Batch normalization* (BATCH-NORM), *Nonlinearity*), které tvoří navrženou architekturu. Důležitou částí je několik bloků konvolučních vrstev, které mají architekturu *diabolo* (úzkého hrdla), kde výstupní vrstvy jsou větší než ty vstupní. Díky tomu je možné naučit se robustnější spektrálně – prostorovou reprezentaci z HSI kostek. Toto je možné díky zvyšování počtu spektrálních cest v původní HS kostce krok za krokem na každém bloku. Vytváří se tak iluze pyramidy, kde čím hlubší jsou reziduální jednotky tím více mapových jednotek může být získáno. I přes množství kladů jsou tyto pyramidové jednotky výpočetně náročné, což vede k přidávání zrychlujících technik pro snížení výpočetního času (přidání *Graphical Processing units – GPU*). Výsledky jsou jako v předchozích článcích porovnány nad čtyřmi známými HS data sety (Indian Pines, University of Pavia, Salinas Valley a Kennedy Space Center) s 10 různými klasifikačními přístupy (SVM, Random forest (RF), MLP, Extreme learning machines (ELM), kernel ELM, 1-D CNN, 2D CNN, 3-D-CNN, SSRN, DFCNN). Celkově *DPRN* dosahuje vyšších přesností a nižší standardní odchylky než SSRN, což znamená, že nejistota v třídách je výrazně snížena bez ohledu na uvažovanou prostorovou velikost. Pro sestavení architektury je použit programovací jazyk Python. Do budoucna autoři poukazují na několik nevyřešených problémů, na které se do budoucna zaměří (snížení výpočetní komplexnosti, vývoj účinnějších paralelních implementací navrhované metody).

Poslední z navržených článků této kapitoly se zabývá metodou *MMFN* (*multiscale deep middle-level feature fusion network*) (Li, Z., et al., 2019). Skládá se ze dvou nezávislých tréninkových fází. V první se slučují spektrální a prostorové informace a efektivně se extrahují spektrálně – prostorové prvky fúze. Druhá fáze slučuje doplňkové, ale korelované informace mezi regiony různých rozměrů a efektivně extrahuje multiměřítkové příznaky fúze, které mohou vést k dalšímu zlepšení klasifikace. *MMFN* funguje lépe pro učení se spektrálně prostorových prvků než jedno měřítkových sítí. Z výsledků vyplývá, že metoda pracuje lépe, když jsou tréninkové vzorky malé. Oproti ostatním metodám má *MMFN* v mnoha případech řadu výhod. Například optimální *middle-level* příznaky jsou užitečné pro druhou fázi tréninku s extrakcí více diskriminačních prvků. Prvky spojené v konvoluční vrstvě mohou napomoci síti k získání silných doplňkových a souvisejících informací mezi víceúrovňovými funkcemi. Další je, že *residual learning* může pomoci síti udržet vyšší přesnost i za předpokladu, že je vrstva

hlubší, a dělá síť robustnější. Nevýhodou navrhované metody je dlouhý čas trénování i z toho důvodu, že trénování sítě je rozděleno na dvě části. Stejně jako předchozí metody i tato je porovnána nad několika známými data sety (IP, UP, KSCC, Salinas) s dalšími nejmodernějšími architekturami (SVM, ResNet, Stacked autoencoders (SAE), 3-D CNN, Two-CNN, SSRN).

2.4 Klasifikace HS dat v tundře

Tato část rešerše se obecně věnuje využití nejen hyperspektrálního DPZ ve specifickém území tundry v různých částech celého světa.

Bratsch, S., et al. (2016) nevyužívají HS snímky ke zkoumání vegetace tundry, ale zkoumají přínos terénní spektroskopie. V oblasti Brooks Mountain Range na Aljašce zkoumají možnosti pozemní spektroskopie rozlišovat mezi čtyřmi vegetačními společenstvími. Tato studie představuje potenciální příklad, jak zlepšit mapování tundry pomocí DPZ. Výsledky poukazují na fakt, že HS data získaná pozemním průzkumem v kombinaci s *SPLS* (*sparse partial least squares*) a *LDA* (*linear diskriminant analysis*) metodikami mohou úspěšně rozlišovat mezi vegetačními společenstvími tundry.

Davidson, S., et al. (2016) se věnují území Aljašky a otestování kombinací odrazivosti a vegetačních indexů získaných z pozemních měření v kombinaci se satelitními daty a tím získání přesných vegetačních map. Další otázkou, kterou se zabývají, je, zda by tato metodika byla přenositelná i do jiných lokalit za účelem vývoje systematické metody mapování vegetačních společenstev. Vysokých přesností dosáhli s využitím lineární diskriminační analýzy.

Pro oblast Seward Peninsula (Aljaška) je navržena metoda spojení HS, MS, radarových a terénních dat pomocí řízených a neřízených klasifikačních technik (Langford, Z., et al. 2017). Cílem je vyvinout za pomoci CNN vegetační mapy vysokého rozlišení ve vybraném území. Nejlepších výsledků dosahuje kombinace HS a optických dat zpracovaných neřízeným shlukováním.

Stejnému území se věnují i Langford, Z., et al. (2018). Snaží se vyvinout metodu, která by eliminovala vlivy působící na nekvalitní mapování polárních oblastí jako jsou oblačnost, polární tma a nedostatečná dostupnost DPZ obrazů s vysokým rozlišením. Navrhují fúzi HS, MS a terénních dat za využití řízené a neřízené klasifikace. Díky tomu je získána vegetační mapa s vysokým rozlišením 5 m. CNN modely založené na označených datech získaných neřízenou klasifikací dosáhly nejlepších výsledků.

Interpretace vegetačních a nevegetačních objektů v tundře/tajze ze snímků s průměrným rozlišením je komplikovaná, protože je složité rozlišit tyto objekty v rámci jednoho pixelu. Proto je navržena subpixelová klasifikace snímků Terra ASTER ekotonu tundry/tajgy v Tuliok

River v Rusku (Mikheeva, A., et al. 2017). Je prokázáno, že tato metoda umožňuje určit polohu hranice objektů v jednotlivých ekotonech a jejich četnost na základě kvantitativních kritérií, což poskytuje mnohem přesnější charakteristiku vegetace ve srovnání s přístupem automatické interpretace snímků po pixelech. *MTMF (mixture tuned match filtering)* adaptivní filtrační algoritmus je vybrán pro subpixelovou klasifikaci obrazů.

Zagajewski, B., et al. (2014) používají pro mapování vegetačních ekosystémů vysoko v horách metodu kombinující HS obrazová a pozemní data. Cílí také na porovnání různých vegetačních klasifikačních metod a na vytvoření mapy, která bude odpovídat skutečné vegetaci. Zkoumaným územím je Dolina Gasienicowa v tatranském národním parku. Jsou porovnány dvě klasifikační metody: ML a NN. Výsledky klasifikací jsou porovnány s vegetační mapou získanou během terénního mapování. Porovnány jsou oba přístupy – řízený i neřízený. Neřízená klasifikace je příliš jednoduchá, aby byla schopna rozlišit jednotlivé vegetační druhy, a proto není vhodná pro tyto účely. Nejlepších výsledků tak bylo dosaženo s řízenou klasifikací a použitým algoritmem ML. NN se ukázaly být vhodné k separaci a klasifikaci vegetačních druhů. Podrobnější porovnání mezi ML a NN nebyla specifikována. Pro tyto účely je použit *SNNS (Stuttgart neural network)* simulátor.

Další studie se zabývá územím Krkonošského národního parku na polské straně. Článek je založen na Airborne Prism Experiment (APEX) s rozlišením 3.12 m, simulovaném environmentálním mapování a EnMAP datech. Zabývá se alpínskou a subalpínskou vegetací. Výše zmíněná data umožňují klasifikaci 21 vegetačních komunit, které jsou zgeneralizovány do 8 vegetačních tipů. Nejlepších výsledků je dosaženo pomocí lineární funkce a SVM *linear kernel function* (Marcinkowska, A., et al. 2017).

2.5 Klasifikace HS dat v krkonošské tundře

Vědecké články této kapitoly se věnují postupům klasifikace vegetačního pokryvu v krkonošské tundře na území České republiky a Polska. Jsou vybrány především na základě stejného zájmového území. Cílem jejich použití je podat rámcový přehled využití a typů klasifikace leteckých dat ve specifickém území alpínské (krkonošské) tundry. Středem zájmu jsou použité klasifikace a jejich parametry.

Důležitý článek, který se shoduje v mnoha parametrech s touto prací, je od autorů Kupková, L., et al. (2017). Jejich cílem je porovnat vhodnost leteckých HS dat a Sentinel-2A dat pro klasifikaci vegetace v KRNAP (*Krkonošský národní park*). Zkoumaným územím jsou nejvyšší partie Krkonoš, které se nachází nad linií stromů (okolí Luční a Labské boudy). Tyto partie jsou pokryty reliktní arкто-alpínskou tundrou. Hlavním předmětem klasifikace jsou alpínské druhy travin, nachází se tu i drobné křoviny, kleč, alpská vřesoviště a kamenné sutě.

V této studii jsou použity senzory AISA Dual a APEX. Pixelovým (ML, SVM, NN) i objektovým přístupem byla klasifikována HS data ze sensoru AISA Dual s prostorovým rozlišením 1-3 m a 494 spektrálními pásmy a data ze sensoru APEX, která mají prostorové rozlišení 3-5 m a 288 spektrálních pásem. Data Sentinel 2A jsou vynechána, protože nejsou předmětem zájmu. Nejlepší výsledek byl zaznamenán u dat ze sensoru AISA Dual pixelovou klasifikací SVM (84,3 %). Výsledky s daty ze sensoru Apex nebyly o moc horší.

Obdobným tématem se zabývají i Marcinkowska, A., et al. (2018), kteří se věnují aplikování leteckých HS dat pro mapování horské vegetace. Vyklasifikovali 22 tříd na základě leteckých Prism Experiment (APEX) HS obrazů s prostorovým rozlišením 3.12 m a 288 spektrálními pásmy. Klasifikačním algoritmem bylo SVM s výsledným průměrným OA 84%.

Stejně tak i Marcinkowska, A., et al. (2014) se zaměřují na ověření potenciálu HS dat pro mapování horských vegetačních ekosystémů. Zkoumaným územím je KRNAP v okolí Szrenica Góra (nedaleko Svinských kamenů a Vosecké boudy) na česko-polské hranici. Použita jsou HS data pořízena senzorem APEX s 288 spektrálními pásmy a prostorovým rozlišením 1.75 m. Ve spojení s terénním průzkumem bylo vytvořeno 15 tříd vegetace, na které byla aplikována SVM klasifikace. Téměř všechny třídy byly správně rozpoznány. OA vyšla 79.13 % a kappa koeficient 0.77.

Oproti ostatním se Suchá, R., et al. (2016) zabývá porovnáním vhodnosti použití MS dat s rozdílným spektrálním a prostorovým rozlišením pro klasifikaci krkonošské tundry nad linií stromů. Navrhují dvě legendy – podrobná s 12 třídami a zjednodušená s 8 třídami. Na letecké ortorektifikované snímky s vysokým prostorovým rozlišením 12.5 cm a 4 spektrálními pásmy je použita objektově orientovaná klasifikace. Satelitní data WorldView-2 (WV-2) (rozlišení 2 m a 8 spektrálních pásem) jsou zpracována objektově orientovanou klasifikací a pixelově orientovanou klasifikací. Ta je použita i na Landsat data 30 m se 7 pásmy. Jako algoritmy pro pixelově orientovanou klasifikaci jsou porovnány MLC, SVM a NN. Pro objektovou klasifikaci je použit SVM algoritmus. Nejlépe vyšla klasifikace ortofot s OA 83.56 % objektově orientovanou klasifikací oproti datům Landsat a WV2, která mají menší prostorové rozlišení. To potvrzuje fakt, že čím vyšší prostorové rozlišení, tím je většinou dosaženo vyšší OA.

Porovnání tří klasifikací: SVM, RF a ML se věnuje Jedrych, M., et al. (2017) s použitím snímků Sentinel-2 a HS EnMAP. Zájmovým územím je alpská vegetace Krkonoš podél česko-polské hranice. Stejně jako v předchozích studiích i zde nejlépe vyšla klasifikace SVM s 82 % pro Sentinel-2 data a 83 % pro EnMAP data.

Méně zaměřena na metodiku klasifikace, ale více na stav vegetace je studie od autorů Jarocińska, A., et al. (2016). Zabývají se analýzou stavu nelesních horských rostlinných společenstev – luk a pastvin. Územím zájmu je přeshraniční biosférická rezervace krkonošské tundry zařazená do UNESCO. HS snímky jsou pořízeny skenerem APEX s 288 spektrálními pásmy a prostorovým rozlišením 1,75 – 3,5 m. Na data je aplikována klasifikace SVM s *linear kernel*. Je hodnocen především stav vegetace pomocí spektrální odrazivosti a indexů, takže podrobné výsledky samotné klasifikace nejsou rozebrány.

3 Zájmové území a vstupní data

Tato část práce se věnuje vymezení zájmového území s jeho charakteristikami. Dále jsou popsána vstupní data včetně jejich získání a předzpracování pro vstup do klasifikace. Není opomenuta ani legenda s jejími třídami, které budou klasifikovány.

3.1 Zájmové území

Předmětem zájmu této práce je území arкто-aplínské tundry ve vrcholových partiích Krkonoš. Rozkládá se na ploše 47 km² na české i polské straně hor nad hranicí lesa s průměrnou nadmořskou výškou nad 1400 m. Podle specialistů na severskou biogeografii je tundra charakterizována třemi základními kritérii. Jedná se o bezlesé území s celoročním průměrem teplot pod bodem mrazu a trvalým permafrostem (Štursa, J., 2013). Soukupová, L., et al. (1995) charakterizuje území krkonošské tundry průměrnými ročními teplotami mezi 0 °C a +1 °C, značným vlivem činností větru na charakter území, průměrným ročním úhrnem srážek 1400 mm a sněhovou pokrývkou, která leží na povrchu v průměru 180 dnů v roce.

Díky unikátní mozaice ekosystémů, které představují ojedinělou kombinaci severských, středoevropských a alpínských společenstev, je toto území přírodní klenotnicí Krkonoš (KRNAP, 2018). Soukupová, L., et al. (1995) člení krkonošskou tundru do tří vegetačních zón: kryoeolická (lišejníky vrcholů), kryovegetační (tundrové traviny) a niveoglacigenní (květná tundra karů). Vegetace zde má zastoupení v podobě mechů, lišejníků, klečových porostů, alpínských vřesovišť, vegetace rašelinišť a výrazný podíl alpínských trávníků. Z nich mají hlavní zastoupení smilkové louky označované jako hercynská poušť, které jsou převážně přirozeného původu. Na loukách se vyskytují i třtina chloupkatá, bezkoleneček modrý, metlice a další (KRNAP, 2018).

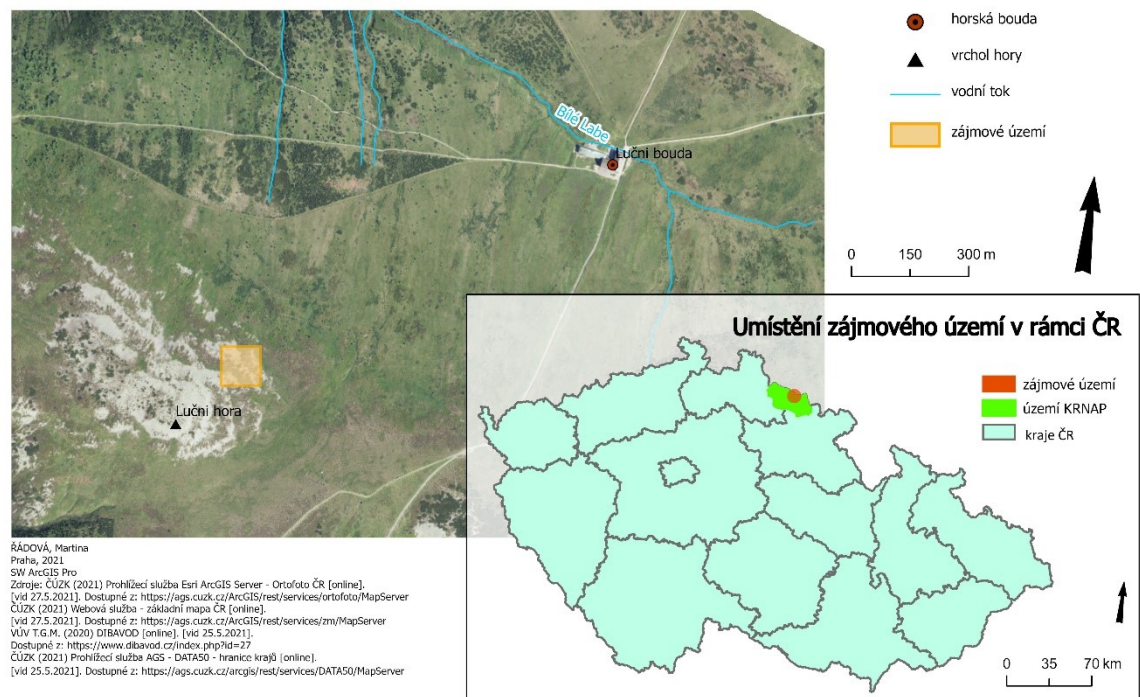
Toto konkrétní složení a zastoupení typů vegetace úzce souvisí s vývojem a charakterem krkonošského pohoří. Čtvrtohorní ochlazení vedlo k rozšíření skandinávského ledovce jižním směrem, který sice hřbety Krkonoš nepřekonal, ale umožnil proniknutí severské tundry do střední Evropy. Mimo jiné došlo i k celkovému ochlazení podnebí, což ovlivnilo vývoj reliéfu Krkonoš i jejich přírody. Při následném oteplení vznikl z horních partií Krkonoš izolovaný ostrov bezlesé krajiny, který napomohl ke stabilizaci reliktních severských a alpínských druhů a jejich následnému izolovanému vývoji (Štursa, 2013). Svůj význam hrály pro zachování tundrového reliéfu i specifické geologické podmínky, které zachovaly rozsáhlé plochy zarovnaných povrchů (KRNAP, 2016). Díky své specifičnosti a jedinečnosti je toto území označováno jako křižovatka biodiverzity (Jeník a Štursa, 2003 **Chyba! Nenalezen zdroj odkazů.**).

Již od středověku začal území Krkonoš ovlivňovat další faktor. Tím bylo proniknutí lidí do vrcholových partií hor díky dvěma zemským obchodním stezkám. Okolo nich začaly koncem 16. století vznikat nejstarší horské boudy. Jednou z nich je i Luční bouda jejíž založení roku 1623 (KRNAP, 2015) souvisí s tzv. budním hospodařením. To mělo za následek první zásahy do nenarušených hřebenových partií. Zasaženo bylo do přirozené hranice alpského lesa, porostů kleče a zvětšování alpského bezlesí. Ačkoliv některá stanoviště zanikla, jiná mohla vzniknout s leckdy bohatší druhovou rozmanitostí (např. květnaté horské louky) (Štursa, Wild 2014). Větší či menší zásahy v rámci budního hospodářství pokračovaly až do roku 1963, kdy Krkonoše získaly statut národního parku (Štursa, 2013). V dnešní době se park dělí do 4 chráněných zón, přičemž ty nejcennější partie spadají do první, nejvíce chráněné bez lidského zásahu.

3.1.1 Luční hora

Klasifikovaným územím (Obrázek 1) je plocha v okolí vrcholu Luční hory o rozloze 100 x 100 m, v nadmořské výšce 1485 až 1545 m.n.m a přibližně kilometr od Luční boudy. Luční hora je druhý nejvyšší vrchol České republiky a díky klimatickým a půdním poměrům se jedná o jedno z velmi unikátních a rozmanitých stanovišť nejen u nás, ale i v Evropě.

Zájmové zemí - Luční hora



Obrázek 1: Vymezení zájmového území

Její výjimečnost spočívá v jejím tvaru, který je stejně jako Sněžka trojúhelníkového půdorysu. Všechna tři úbočí jsou odlišná (Obrázek 2 **Chyba! Nenalezen zdroj odkazů.**a). Jihozápadní svahy spadající do Dlouhého dolu jsou nejvyšší, prudké a erozní. Jsou tak charakteristické množstvím vodopádů, kaskád a díky strmosti i zemními a sněhovými lavinami. Jihovýchodní strana je závětrná a tvoří se zde sněhová pole. Pro tuto práci podstatná severní strana spadá do údolí Bílého Labe. Ačkoliv jsou její svahy nejnižší a nejmírnější, z geomorfologického hlediska jsou velmi pozoruhodné. Díky mírnějšímu charakteru svahů se zde mohly uplatnit třídící mrazové procesy. Dlouhodobým působením mrazu vznikly kryoplanační terasy, které tvoří horizontální linie tvořené úlomky hornin na svazích hory. Na temeni Luční hory je možné pozorovat i jev zvaný tříděné půdy. Ty vznikají vlivem tlaku, způsobeného střídáním tání a mrznutí, na různě velké úlomky hornin, které jsou následně tříděny. Pokud začne působit i gravitace ve fázi tání půdního ledu, vznikají brázděné půdy (Obrázek 2b). Otázkou, která zůstává je, zda všechny tyto útvary jsou pouze pozůstatky dob minulých nebo procesy jejich vzniku stále fungují. Ať tak či onak, Luční hora je unikátem v mnoha oblastech nejen u nás, ale i v Evropě, díky svému bohatství, rozmanitosti a množství



Obrázek 2: Luční hora – tvar (a) a útvary (b) (Zdroj: Nebeske

endemitů (KRNAP, 2016).

3.2 Senzor

Jak uvádí Červená, L. et al. (2020), zpracovávané snímky jsou zachyceny řádkovou kamerou Headwall Nano-Hyperspec. Kamera je společně s IMU a GNSS přijímačem nesena dronem DJI Matrice 600 Pro. Tato hyperspektrální kamera snímá celkem 270 pásem s pravidelným rozestupem 2,24 nm na intervalu vlnových délek od 398,784 nm do 1 001,84 nm ve viditelném a blízkém



Obrázek 3: Dron s kamerou (Zdroj: Červená, L., 2020)

infračerveném spektru. Radiometrické rozlišení je 12 bit, šířka záběru 640 Px a ohnisko 17 mm. Pro pořízení dat je použito několik softwarů, díky nimž může samotný let, po jeho naplánování, probíhat automaticky. Pro naplánování letu slouží SW Universal Ground Control (UgCS), vytvořená trasa se posílá do tabletu s aplikací UgCS pro DJI. Dron létá nad stanovenými polygony s nastavenými parametry jako jsou například výška a rychlost letu. Toto nastavení je možné v SW Hyperspec III. Při požadované výsledné velikosti pixelu 0,03m je výška letu 68,9 m nad terénem.

3.3 Použitá data

Snímky byly pořízeny výše zmíněnou kamerou ve vymezené oblasti v okolí Luční hory (viz kapitola 3.1.1 Luční hora) a dále předzpracovány do podoby základního vstupu do architektury. V následující části je popsán postup předzpracování dat. Jsou specifikovány parametry vstupující do klasifikace. Důležitá část je věnována i několika variantám vstupních dat, které jsou pouštěny do architektury. Krátce jsou zmíněna i data, která jsou původním vstupem do použitých metod.

3.3.1 Předzpracování dat

Červená, L., et al. (2020) popisují předzpracování dat v několika fázích. První je již v terénu, kde se provádí kalibrace *dark reference* (měření vnitřního šumu přístroje). Měří se před samotným letem. Dark reference je jediným vstupem do procesu převedení všech data (*data cubes*) z DN hodnot na zář. Druhým krokem radiometrické korekce je přepočítání všech *cubes* na odrazivost. K tomu se používají měření DN hodnot referenčního koberce, který se umísťuje v terénu do snímané plochy. Jeho odrazivost je známa. Pro výše zmíněné radiometrické korekce je použit SW SpectralView – Hyperspec v3.1.0. Nevyrovnané hodnoty v datech mohou být způsobeny nerovnoměrným osvětlením snímaného území z důvodu

oblačnosti. K ověření přesnosti radiometrických korekcí slouží SW ENVI 5.5. V první řadě je žádoucí redukovat množství šumu v datech. V rámci řešení projektu bylo vyzkoušeno několik metod jako například PCA nebo spektrální převzorkování, z nichž nejlépe vychází spektrální převzorkování dat se sloučením pěti sousedních pásem (Červená, L., et al., 2020).

Další úpravou jsou geometrické korekce. Nejprve jsou sloučeny cubes do jednotlivých letových řad. K vytvoření těchto mozaik jsou využity i informace získané během letu systémy GNSS (*globální navigační satelitní systém*) a IMU (*inerciální měřící jednotka*). Následuje ortorektifikace s využitím rastrového DMT. Kvůli polohovým nepřesnostem je třeba ještě data manuálně zpracovat v ArcGIS Desktop. S využitím vlíčovacích bodů nalezených v ortofotu zájmového území pořízeném z jiného senzoru (v našem případě RGB camera Sony A7 ILCE-7 24,3 Mpx) a *spline* transformace byly jednotlivé řady snímků geometricky zpřesněny. Pro zachování radiometrie byly převzorkovány metodou nejbližšího souseda. Bylo využito přibližně 10 vlíčovacích bodů na jednu letovou řadu. Letové řady byly nakonec mozaikovány do jednoho snímku. Velikost pixelu výsledné mozaiky jsou 3 cm. Hodnocení geometrické přesnosti proběhlo na základě několika rovnoměrně rozmístěných bodů zaměřených geodetickým GNSS přístrojem přímo v terénu. RMSE (*root mean square error*) u manuálně upravených dat vychází mezi 8 a 13 cm (Červená, L et al., 2020).

3.3.2 Vstup do klasifikace

3.3.2.1 Volně dostupné datové sady

Data, která vstupují do klasifikace, sestávají ze dvou typů. Nejprve bylo snahou ozkoušet zvolené architektury na stejných datech jako byly v původních článcích (Paoletti, M., et al., 2019a; Zhu, K., et al., 2019). To znamená, že byly použity volně dostupné datasey Salinas Valley (SV), Kennedy Space Center (KSC) a Indian Pines (IP) (získané z následující webové stránky:

http://www.ehu.eus/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes#Salinas).

Ty jsou použity v obou testovaných kódech. Tyto datasey nejsou upravované, jsou použity ve formě, ve které byly obdrženy. Dostupná data, obsahující dva soubory, jsou ve formátu *mat*, tedy data v binárním formátu. Prvním z nich je hyperspektrální snímek s daným počtem spektrálních pásem a prostorovým rozlišením. Druhý, který má v názvu *_gt*, obsahuje matici čísel s trénovacími daty. Tedy čísla od jedné do celkového počtu tříd daného datasetu. Pixely, které jsou nezařazené do tříd, mají přiřazenou hodnotu 0. Konkrétní parametry datasetů jsou vypsány v Příloha 1.

3.3.2.2 HS data HeadWall Nano-Hyperspec

Druhým testovaným datasetem jsou data získaná senzorem Nano-Hyperspec, viz kapitola 3.2. Jedná se o hyperspektrální snímky o původním spektrálním rozlišení 269 pásem a prostorovém rozlišení 3cm. S ohledem na porovnatelnost výsledků uvedených v Kupková, L., et al. (2020), byly použity snímky s prostorovým rozlišením redukováným na 9 cm a spektrálním rozlišením redukováným na 54 pásem (průměrování 5 po sobě následujících pásem). Klasifikace byla provedena na datech pořízených ve třech termínech vegetační sezóny 2020 a jejich multitemporálním kompozitu. Vzhledem k jejich rozměru bylo nutné snímky před vstupem do vybrané architektury upravit. Tento postup je popsán v kapitole 5.1. Nelze opomenout ani trénovací data. Ta byla získána ve formě polygonů, které bylo třeba převést na raster s velikostí buňky a umístěním stejným jako hlavní snímek. Celý postup včetně použitých SW je popsán v kapitole 5.1.

Tabulka 1: Parametry krkonošských dat

Data	Původní mozaika	Redukovaná mozaika
Parametry		
Počet pásem	269	54
Prostorové rozlišení	3 cm	9 cm
Velikost v pixelech	4689 x 5541	1563 x 1847
Termíny snímání	13.7., 11.8., 15.9. 2020	

3.3.3 Třídy

Třídy pro klasifikaci byly stanoveny v rámci projektu *Vegetace krkonošské tundry – minulost, současnost a budoucnost* (Kupková, L. et al. (2019)).

Pro vybrané území Luční hory je stanoveno 5 tříd hrubé legendy, do které je roztríděno níže popsaných 9 tříd (Tabulka 2). Mezi nejpočetnější a nejsložitější kategorii ke klasifikaci patří alpské trávníky zapojené/vyfoukávané. Vyfoukávané alpské trávníky charakterizuje Chytrý, M. et al. (2010) jako porosty rozvolněné až zapojené s dominantním zastoupením metličky křivolaké. Tato vytrvalá trsnatá tráva se vyskytuje od nížin po subalpínská pásma na sušších světlých místech. Indikuje kyselé půdy se zhoršeným rozkladem humusu (Botany, 2009). Jejím kvetoucím obdobím je červen až červenec a charakteristickou barvou je stříbřitě bílá s našedlými až svěže zelenými, hladkými a měkkými listy (Příloha 4). Tato travina bývá doplněna tomkou alpskou a ostřicí Bigelowovou. Travinná vegetace je střídána plochami holé půdy, kameny a kamennými moři. Mimo bylin se vyskytují i drobné keřiky, především vřes obecný místy doplněný o subalpínský brusnicový porost.

Oproti tomu zapojené alpínské trávníky mechové patro podle Chytrý, M. et al. (2010) téměř postrádají. Jejich dominantou jsou zapojené nízké trávníky s převládající smilkou tuhou (Příloha 5). Ta je charakteristická především pro Krkonoše a bývá doplněna metličkou křivolakou. Smilka se vyskytuje na severních chladnějších svazích ve středních až vyšších polohách hor. Je charakteristická svou monodominancí s podkladem chudších vlhčích písčitohlinitých až rašelinných půd v blízkosti cest, rašelinišť, vřesovišť a na horských loukách. Její kvetoucí období je od května do června s nafialovělými květy, tvrdými a drsnými stébly a šedozelenými listy (Botany, 2008).

Své zastoupení v tundře a legendě má i metlice trsnatá (Příloha 6). Jedná se o polymorfní druh s rozšířením na téměř celé severní polokouli zejména ve středních až vyšších polohách. Roste na suchých i vlhkých stanovištích, loukách i lesích. Květné období je od června do srpna (Botany, 2012). K tomu jsou všechny výše zmíněné třídy doplněny borovicí kleč. Pro Krkonoše je původní rostlinou s výskytem nad hranicí lesa a kamenitých místech.

Tabulka 2: Legendy

Název	Třídy	Hrubá legenda
metlička křivolaká	af	alpínské trávníky zapojené/vyfoukávané
af + tomka + ostřice	afs	alpínské trávníky zapojené/vyfoukávané
brusnice borůvková, keříčková vegetace	bor	subalpínská brusnicová vegetace
metlice trsnatá	desch	alpínské trávníky zapojené/vyfoukávané
borovice kleč	klec	kosodřevina
smilka tuhá	nard	alpínské trávníky zapojené/vyfoukávané
kamenná moře bez veg., lišejníky	sut	kamenná moře a antropogenní plochy
vřes obecný, keříčková vegetace	vres	alpínská vřesoviště/vyfoukávané alpínské trávníky
kameny, půda, mechy a vegetace	vyfuk	alpínské trávníky zapojené/vyfoukávané

Výzvou v klasifikaci výše popsaných druhů travin je jejich značně podobný spektrální projev. Ten může značně ztížit proces, a především výsledek, klasifikace. Proto je velmi důležitý terénní průzkum a pravidelné pozorování sledované oblasti, aby bylo možné vypočítat rozdíly ve fenologických fázích a s tím související průběh vegetačních sezón. Jak vyplývá z popisu, květné období jednotlivých travin je částečně odlišné od ostatních, což může napomoci jedinečné spektrální odezvě a tím i snazší klasifikaci. Z tohoto pohledu jsou zajímavé i multitemporální klasifikace, které porovnávají delší časový úsek sledovaného území.

4 Metodika

Podstatnou součástí je i představení a rozebrání samotné metodiky Capsule network použité pro klasifikaci dat. K tomuto tématu bylo nalezeno několik článků, z nichž těmi stěžejními jsou tři, které jsou aktuální, věnují se podrobně problematice a ke dvěma z nich jsou otevřené kódy v Python. Jedni z prvních průkopníků tohoto tématu, jejichž kód je volně dostupný, popisují *capsule* jako skupinu neuronů, jejichž vektor aktivity reprezentuje parametr specifického typu entity jako je například objekt (Sabour, S. et al., 2017). Výše zmíněným článkem a jejich metodou se inspirovali i Paoletti, M. et al., (2019a), kteří se snaží vypořádat se s vysokou komplexností HS dat. Daří se jim dosáhnout komplexnosti sítě pomocí jejich navržené metody založené na texturně – prostorové kapslové síti. Vytvářejí rozšíření založené na Hinton's kapslové síti. Výsledky ověřují a porovnávají na pěti běžně používaných data setech. Ve dvou z nich se shodují i s dalším článkem zabývajícím se *Capsule network*. Jeden data set je ze zemědělské oblasti s rozlišením 3.7 m a druhý zachycující bažiny a jejich vegetaci s rozlišením 20 m. Díky tomu je možné tyto dvě metody porovnat a díky částečně podobným legendám uvažovat jako vhodné pro tuto diplomovou práci. Vzhledem k dostupnosti otevřeného kódu bude použita metoda od autorů Zhu, K. et al. (2019). Kapslové sítě považují za novou éru hlubokého učení s jejich výhodami v řízených klasifikacích. Výhodou metody je i fakt, že používá vektory (místo skalárních hodnot) k reprezentaci příznaků, což obohacuje jejich schopnost prezentace. Klíčovou částí sítě je modifikace kapslové sítě zvaná *Conv-Capsule*. Jejím přínosem je využití lokálních spojení a sdílených přeměněných matic, což vede ke snížení trénovaných parametrů, které zmírňují problém přeučení při omezeném množství trénovacích vzorků

Následující kapitola představuje samotnou metodu, která je použita pro tuto práci. Je zasazena do rámce hlubokého učení, porovnána s CNN a jsou vyzdvihnuty její výhody. Jsou popsány oba přístupy využití CapsNets od autorů Paoletti a Zhu. Snahou je vystihnout a popsat jejich charakteristické znaky a odlišnosti.

4.1 Capsule network

Vybraná metoda Capsule Network spadá do kategorie hlubokého učení, které se vyznačuje schopností řešit komplexní problémy spojené s obrazovou segmentací, rozpoznáváním objektů a dalšími. Síť umělých neuronů osahují vrstvy a neurony, které jsou vzájemně propojené synapsemi s váhami. Nastavení vah probíhá skrz zpětnou propagaci (*backpropagation*), což umožňuje síti se učit. Síla neuronů je kontrolována aktivační funkcí (například *sigmoid function*, *ReLU*, *Soft-Max Activation function* a *Hyperbolic Tangent function*) (Nielson, 2019).

Předchůdcem CapsNets jsou CNN, které jsou charakteristické svými několika vrstvami - konvoluční, pooling a plně propojenou. V konvoluční fázi je využito $n \times m$ velkého filtru (*kernel*) k automatickému získání příznaků ze vstupního snímku. Tento filtr je menší než originální snímek, po kterém se pohybuje s daným krokem (*stride value*) a vytváří tak *feature map* (výstup jednoho filtru použitého na předchozí vrstvu). Čím je hodnota kroku větší, tím menší je výsledný výstup. Pokud krok přesáhne hodnotu dvě, dochází ke ztrátě některých prvků snímku. K získání co největšího množství prvků je využito několika unikátních filtrů. Pooling vrstva neboli redukční slouží k zaručení, že CNN rozpozná stejné objekty různých podob a taky ke snížení paměťových nároků modelu. V různých metodách využívají několik typů (max, min, average a sum pooling). Poslední částí je ANN složené ze vstupní, skryté a výstupní vrstvy.

CNN se potýkají s řadou problémů jako například, že vyžadují velké množství trénovacích dat, jejichž získávání je nejen časově náročné. Dalším záporem je neschopnost CNN rozpoznat texturu, deformace a postoj (pose) v obrazu (Sabour et al., 2017). Jak bylo výše zmíněno, v pooling fázi dochází ke ztrátě některých prvků. Aby vynahradili tuto ztrátu, vyžadují velké množství trénovacích dat a mají také delší trénovací čas než CapsNets. Jsou také náchylnější k některým nežádoucím pixelovým poruchám vedoucím k nesprávné klasifikaci. Jak tvrdí Patrick, M., et. al (2019) tyto potíže vedou k představení Capsule Network, která by jim měla do jisté míry předcházet. Ačkoliv se jeví jako nová senzace, potýká se s několika problémy. Tou podstatnou je částečná neznalost architektury a vnitřních procesů, které vedou k nemožnosti zcela využít potenciálu metody. Patrick, M., et. al (2019) proto poskytují komplexní přehled současných moderních architektur, metodik a nástrojů stávajícího využití Capsule Network. Pro ostatní výzkumníky poskytují motivaci do dalšího objevování CapsNet ve formě shrnutí dosavadních úspěchů, neúspěchů a příležitostí. Velký rozdíl mezi CNN a CapsNet je především ve využití prostorových vztahů a orientaci objektů. CapsNet rozpozná obličej pouze pokud budou oči nahoře a nos mezi nimi, kdežto CNN za obličej uzná cokoliv, co bude mít nos a oči nezávisle na jejich vzájemné poloze.

Metodu Capsule Network navrhli Hinton a jeho kolegové (Hinton, G., et al., 2011). Po představení v roce 2017, s dosažením dokonalejšího výsledku nad datasetem MNIST, se využití metody výrazně zvýšilo. CapsNets mají několik verzí, které lze použít. Tou nejpopulárnější používanou využívá, podle Patrick, M., et al. (2019), algoritmus zvaný *routing by agreement*. Hlavním znakem je nahrazení pooling vrstvy v CNN a jako výstup vektory místo skalárních v konvolučních sítích. Hlavním principem CapsNets, jak vyplývá z výše uvedeného, je následující. Základem jsou vektory, které jsou neurony sítě, jak přijímány, tak i vytvářeny. Díky tomu se mohou z obrazu učit příznaky, jejichž zpracování závisí na typu

použité verze Capsule Network, včetně deformací a podmínek prohlížení (*viewing conditions*). Každá kapsle metody je tvořena skupinou neuronů, kde výstup každého z těchto neuronů reprezentuje různé vlastnosti jednoho příznaku. Výhodou toho je, že jsou nejprve rozpoznány jednotlivé části entity, čímž je rozpoznána celá. Výstup z kapsle je složen ze sady vektorových hodnot zvaných *instantiation parameters* (doložení příkladem) a pravděpodobnosti, že je přítomen kapslový příznak, jehož úkolem je zajistit invarianci sítě. Instantiation parameters reprezentují rovnocennost sítě (*equivariance of the network*), která zajišťuje schopnost sítě rozpoznávat texturu, deformace a pozice (*pose*). Tato schopnost je velmi důležitá, protože zaručuje, že prostorové umístění daného příznaku bude bráno v úvahu. Například zajistí, že v obrazu obličeje nebude brát v úvahu pouze přítomnost oka, ale také jeho pozici v rámci obličeje. Toto je kýženou vlastností CapsNets. Zajímavou vlastností je i neměnnost (invariance), která bez ohledu na jakékoliv transformace vstupů, umožňuje modelu zůstat neměnný (Patrick, M., et al., 2019).

Existují tři hlavní metody využití Capsule Network: *auto - encoders*, *vector capsules based on dynamic routing* a *matrix capsules based on expectation - maximization routing*. Ve vybraných metodách je použita dynamic routing between capsules, která je podrobněji popsána v následující kapitole.

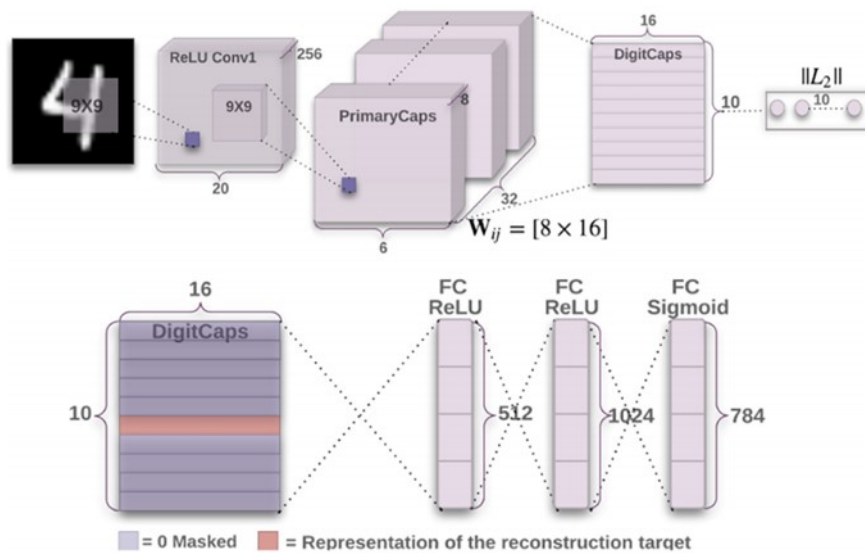


Fig. 5. Structure of the capsule network in Sabour et al. (2017).

Obrázek 4: Základní schéma sítě Capsule Network

Příklad je znázorněn na základním prvku, kterým jsou číslice. Na nich je i založen původní testovaný dataset MNIST. Na vstupu je snímek o určité velikosti. Následuje konvoluční část s filtrem o dané velikosti. Poté následuje samotná kapslová část s jejími kroky.

4.2 Konkrétní Capsule Network

V předešlé kapitole je metoda Capsule Network představena obecně. V této části je snaha podrobněji ukázat přístup *dynamic routing between capsules*, který je využit v obou zvolených metodách. Cílem je pokusit se vyzdvihnout jejich hlavní znaky, případné klady, a především vzájemné odlišnosti v přístupu.

V návaznosti na informace o obecné struktuře Capsule Network navazuje Sabour, S., et al. (2017) s charakteristikou přístupu *dynamic routing between capsules*. Je popsána jako síť tvořená konvoluční, primary capsule (PC) a class capsule vrstvou. Konvoluční vrstva extrahuje příznaky ze snímku a výstup je vstupem pro vrstvu primary capsule. PC je první vrstva, která je následována nespecifikovaným množstvím capsule vrstev, až po poslední, která se nazývá Class capsule. Každá kapsle ve vrstvě jedna má vektor aktivity, aby mohla kódovat (encode) prostorovou informaci ve formě instantiation parameters. Zjistit, jak moc primární kapsle přispívá do class capsule indikuje predikční vektor vrstvy PC. Výstupní vektor kapsle nižší úrovně je vstupem do všech kapslí následující vrstvy. Kapsle této vrstvy získává aktivity vektor společně se související maticí hmotnosti. Na zvýšení či snížení coupling koeficientu (je určen iteračním dynamickým routing procesem) má vliv míra shody mezi kapslemi. Pokud je shoda vysoká jsou dvě kapsle příslušné jedna druhé. Shoda je určena vektorem predikce a coupling koeficientem a získává predikci jednotlivé PC pro class capsule. K získání kandidátů pro squashing funkci slouží vážený součet všech výše zmíněných individuálních PC predikcí pro class capsule. Squashing funkce zajišťuje, že délka výstupu z kapsle leží mezi 0 a 1, stejně jako pravděpodobnost. Shahroudnejad et al., (2018) tvrdí, že během každé iterace jsou coupling coefficients aktualizovány nalezením tečkového (dot) produktu predikčního vektoru. Konkrétně, vektorové hodnoty přidružené ke každé kapsli mohou být zobrazeny jako segment dvou čísel.

Algorithm 1 (*Dynamic Routing Algorithm (Sabour et al., 2017).*).

1. **procedure** ROUTING (\hat{u}_{ji}, r, l)
 2. **for** all capsule i in layer l and capsule j in layer $(l + 1)$: $b_{ij} \leftarrow 0$.
 3. **for** r iterations **do**
 4. **for** all capsule i in layer l : $c_i \leftarrow \text{softmax}(b_i)$
▷ *softmax computes* c_{ij}
 5. **for** all capsule j in layer $(l + 1)$: $s_j \leftarrow \sum_i c_{ij} \hat{u}_{ji}$
 6. **for** all capsule j in layer $(l + 1)$: $v_j \leftarrow \text{squash}(s_j)$ ▷ *squash computes* v_j
 7. **for** all capsule i in layer l and capsule j in layer $(l + 1)$: $b_{ij} \leftarrow b_{ij} + \hat{u}_{ji} \cdot v_j$
 8. **return** v_j
-

Obrázek 5: Dynamic routing algoritmus

Obě zvolené metody využívají Capsule Network s přístupem Dynamic routing between capsules. I tak zůstává pro autory prostor zvolit jiné úhly pohledu a odlišné kroky pro dosažení cíle. V následující části je snaha vyzdvihnout tyto hlavní rozdíly a případně popsat jejich klady či zápory.

Paoletti et al. (2019)	Zhu et al. (2019)
<ul style="list-style-type: none"> • ENCODER <ul style="list-style-type: none"> • <u>Konvoluční vrstva</u> <ul style="list-style-type: none"> • subsequent capsule vrstva • batch normalization + ReLU • <u>PC vrstva</u> <ul style="list-style-type: none"> • inverse rendering proces • squashing funkce • <u>Dense capsule vrstva</u> • 2 a 3 routing by agreement • DECODER <ul style="list-style-type: none"> • <u>Plně propojené vrstvy</u> <ul style="list-style-type: none"> • Dense capsule vrstva • loss function 	<ul style="list-style-type: none"> • 1D <ul style="list-style-type: none"> • <u>Vstupní vrstva</u> <ul style="list-style-type: none"> • PCA, LeakyReLU • <u>Konvoluční vrstvy</u> <ul style="list-style-type: none"> • LeakyReLU, Batch Normalization • <u>Translation vrstva</u> <ul style="list-style-type: none"> • LeakyReLU, Batch Normalization • <u>1D Conv-Capsule</u> <ul style="list-style-type: none"> • Dynamic routing + coupling koef. • <u>Class capsule vrstva</u> = plně propojená <ul style="list-style-type: none"> • Dynamic routing + margin loss • 3D <ul style="list-style-type: none"> • <u>Vstupní vrstva</u> • <u>Konvoluční vrstvy</u> + ReLU a BN • <u>Consecutive vrstvy</u> + ReLU a BN

Graf 7: Porovnání přístupu Paoletti a Zhu

Základním rozdílem, který je patrný i z Graf 7, je hlavní struktura metod. Autorka Paoletti a její kolegové (2019a) navrhují architekturu, která se skládá ze dvou hlavních částí. První z nich je encoder síť, která se skládá ze tří vrstev, které mají za úkol extrahovat příznaky a provést klasifikaci. Následuje decoder síť, která má zlepšit výsledek klasifikace pomocí rekonstrukce dat. Oproti tomu Zhu et al. (2019) využívají dvě verze Conv-Capsule sítě. První z nich je 1D-Conv-Capsule síť, která využívá PCA, CNN, Capsule síť a pouze spektrální příznaky. Druhá část zvaná 3D Conv-Capsule síť využívá obou typů příznaků – spektrální i prostorové. Obě části se skládají ze šesti vrstev, které budou popsány níže. Důležitým faktem je, že Zhu et al. (2019) nevyužívají plně propojené kapslové vrstvy. Ty získávají velké množství trénovacích parametrů, což může způsobit problém přeučení, zejména při nedostatku trénovacích dat. Proto využívají lokálních spojení a sdílené transformační matice, díky čemuž mohou výrazně snížit počet trénovacích parametrů a tím zmírnit problém přeučení (overfitting).

Pokud jsou porovnány první části obou metod, rozdílem je počet použitých vrstev, jejich druhy a typy použitých funkcí. Paoletti, M., et al. (2019a) se na začátku modelu snaží získat relevantní příznaky z originálních HS snímků, které pomohou klasifikaci získat co nejpřesnější a nejužitečnější informace. První vrstvy jsou klasické konvoluční, které přijímají *patches* získané z originálních HS datových krychlí (*data cubes*) jako vstupní příznak. Jejich snahou je uspořádat tyto příznaky, aby byly vstupem do *subsequent capsule* vrstvy s využitím konvolučního filtru dané velikosti následovaného *batch normalization* a použitím ReLU aktivační funkce. Tato vrstva připravuje data k získání vektoru aktivity nejvyšších kapslových vrstev. Následující vrstva zvaná *primary capsule (PC)* je složena z množství konvolučních kapslí, které jsou složeny z daného počtu konvolučních neuronů. Pro dosažení výstupního vektoru jsou konvoluční neurony aplikovány na region o daném objemu. Výstupní vektor poskytuje strukturu, která je mnohem přizpůsobivější při ukládání dalších detailů o prostorových příznacích jako jsou orientace, pozice (*pose*) nebo velikost. Poslední část druhé vrstvy je *inverse rendering* proces, jehož cílem je odhalit objekt a extrahovat vektor aktivity. Po získání výstupu z PC vrstvy a použití *nonlinear squashing* funkce následuje *dense capsule* vrstva. Ta je složena z X třídních kapslí, kde X představuje počet tříd land coveru v originálních HS. Pro každou třídu je obdrženo vektor aktivity. A zde je místo pro implementaci speciálního mechanismu mezi vrstvami 2 a 3, kterému se říká *routing by agreement*. Jeho cílem je spojit tyto dvě vrstvy a navrhnout lepší trénovací proces ve srovnání s tradičními pooling metodami. Velmi podrobný popis tohoto mechanismu je popsán autory Paoletti, M., et al. (2019a) a celé schéma sítě na Obrázek 6.

Oproti tomu Zhu, K., et al. (2019) nejprve používají jednodušší část sítě zvanou 1D Conv_Capsule. Jejím principem je extrahovat pouze spektrální příznaky pro HS klasifikaci. První fáze jejich metody sestává ze šesti vrstev a několika funkcí. Oproti předchozí metodě Zhu, K., et al. (2019) využívají dynamic routing algorithm již v první vrstvě v kombinaci s hlavní ideou CNN – local connections and shared transform matrices. Vstupem do první vrstvy, zvané vstupní, jsou vstupní data, jejichž dimensionalita je zmenšena pomocí PCA. Tato redukce vede k menšímu množství trénovacích parametrů vstupujících do sítě. Každá kapsle této vrstvy je spojena pouze s kapslemi v rámci jejího místního receptivního pole v poslední vrstvě kapsle (*local receptive field in the last capsule layer*). Druhá a třetí vrstva jsou tradiční konvoluční vrstvy vyskytující se v CNN a také v metodě autorů Paoletti, M., et al (2019a). Oproti nim však využívají Zhu, K., et al. (2019) *leaky rectified linear unit (LeakyReLU)* a *batch normalization*. Čtvrtá vrstva je první capsule vrstvou, zvaná také *transition*. Jejím úkolem je překládat konvoluční jednotky na kapsle. Stejně jako ve vrstvě dvě a tři je i v této konvoluční

operace následována *batch normalization* a *LeakyReLU* aktivační funkcí. Předposlední vrstva je 1D Conv-Capsule, která používá místní spojení a sdílené *transform matrices*. Využívá také dynamic routing algoritmus pro iterativní aktualizaci coupling koeficientu. Následně je možné získat výstup kapsle této vrstvy. Poslední je plně propojená capsule vrstva zvaná *Class capsule*. Každá kapsle této vrstvy reprezentuje jednu třídu. Pravděpodobnost, že pixel náleží do jedné třídy je vyjádřena délkou výstupního vektoru každé kapsle. I v této vrstvě je využit dynamic routing algoritmus a jako ztrátová funkce je použita *margin ztráta*. Schéma této části metody je znázorněno na Obrázek 7.

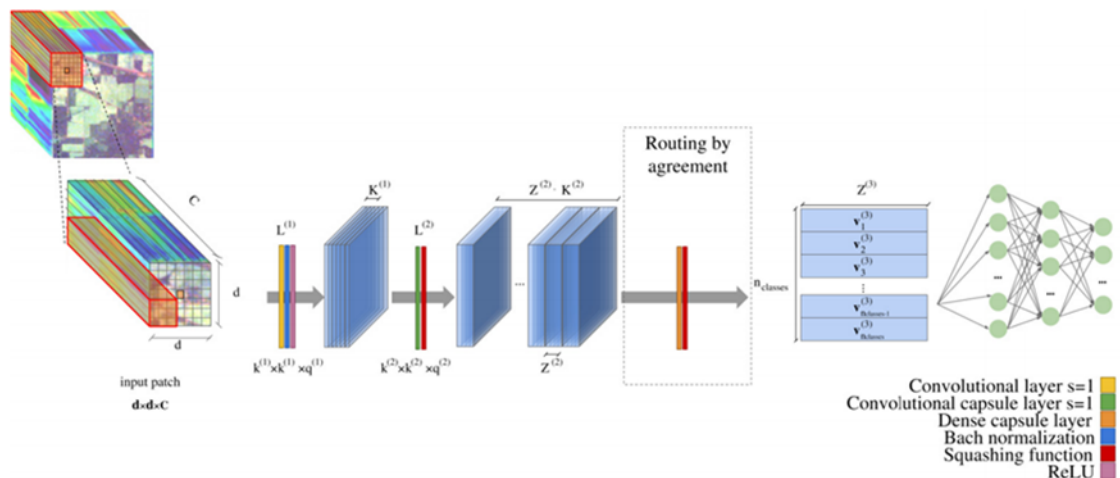


Fig. 2. Proposed neural network architecture. The neural model is composed by an encoder network (in blue) and a decoder network (in green).

Obrázek 6: Schéma Capsule Network Paoletti, M., et al. (2019a)

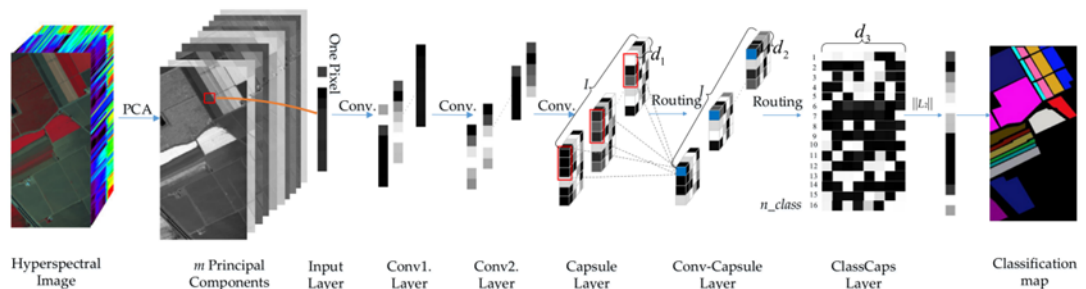


Figure 2. The framework of the 1D-Conv-Capsule network for hyperspectral image (hsi) classification.

Obrázek 7: Schéma 1D Conv-Capsule sítě Zhu, K., et al. (2019)

Druhá část sítě autorů Paoletti, M., et al. (2019a) má za cíl zlepšit jemně laděný proces využití parametrů v navrhované metodě. Decoder síť je složena z několika plně propojených vrstev, které využívají výstupní vektor aktivity dense capsule vrstvy k rekonstrukci vstupního snímku. Tím podněcují kapsle ke kódování nejvíce relevantních *instantiation* parametrů vstupních dat. Nakonec je navržený model optimalizován ztrátovou funkcí. Jak je zmíněno výše, plně propojené vrstvy jsou použity i v poslední vrstvě 1D Capsule sítě autorů Zhu, K., et al. (2019). Zde by se daly nalézt podobnosti obou metod. K redukci dimensionality

vstupů do druhé části metody autorů Zhu, K., et al. (2019) slouží technika EMAP (*Extended Multi-Attributes Profile*) stejně tak jako ke snížení množství trénovacích parametrů. 3D Conv-Capsule síť využívá ReLU jako aktivační funkci a Batch normalization ke zmírnění problému přechodu a k podpoření přesnosti klasifikace. Stejně jako 1D část se i ta 3D skládá ze šesti vrstev. První je vstupní vrstva, následují dvě konvoluční, a nakonec tři *consecutive* (následující) capsule vrstvy. Konvoluční vrstvy slouží jako lokální detektor příznaků. Následně je použita *transition* vrstva (capsule vrstva), stejně jako ta v 1D síti. V posledních dvou capsule vrstvách je použit dynamic routing algoritmus k výpočtu výstupů v Conv-Capsule a ClassCaps vrstvách. V porovnání s 1D částí se vstup do sítě mění z 1D spektrální informace na 3D spektrálně prostorovou a z 1D konvoluční operace na 2D. Jsou použity i ReLU jako aktivační funkce a Batch normalization je použita ke zmírnění problému přechodu a k podpoření přesnosti klasifikace.

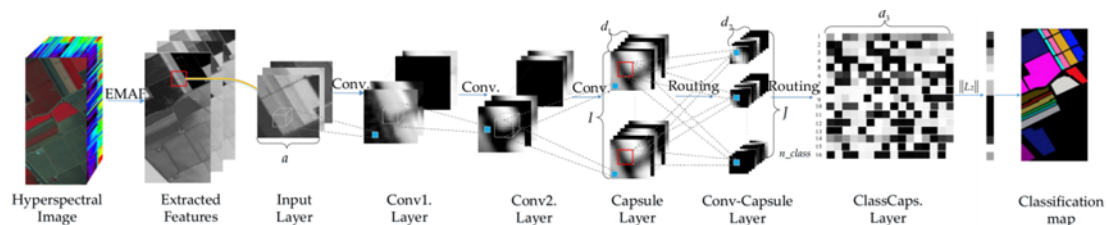
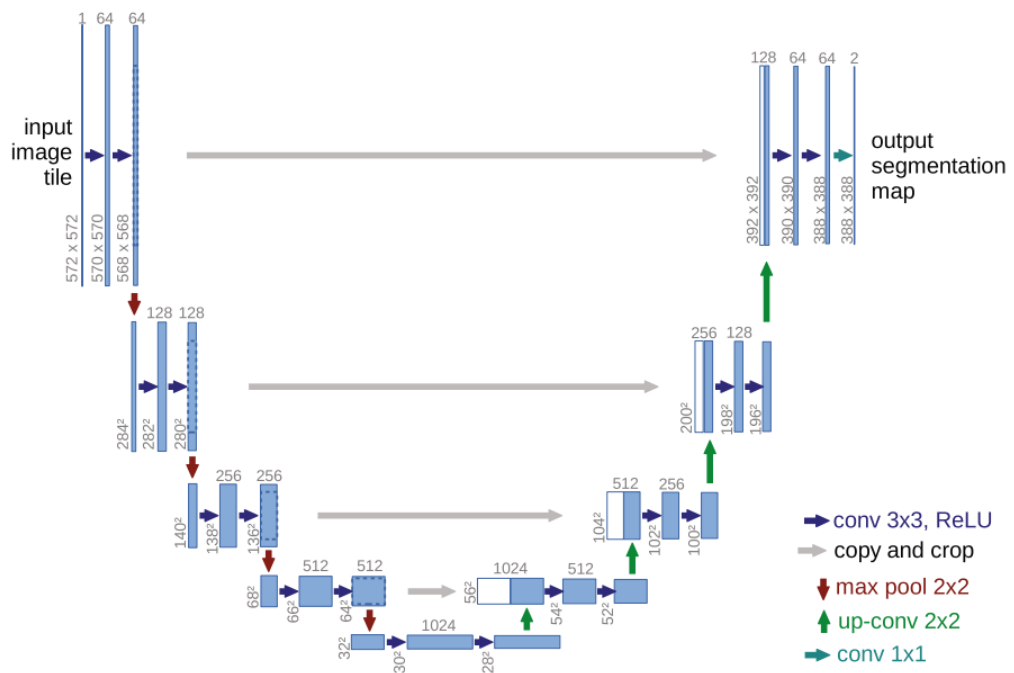


Figure 4. The framework of the 3D-Conv-Capsule network for HSI classification.

Obrázek 8: Schéma 3D Conv-Capsule sítě Zhu, K., et al. (2019)

4.3 U-Net

Architektura sítě U-Net (Ronneberger, O., et al., 2015) sestává ze dvou částí, respektive ze dvou stran. Levá půlka, zvaná uzavírací (*contracting*), je v podstatě typickou architekturou konvoluční sítě. Je složena z opakované aplikace dvou 3x3 konvolucí následovaných ReLU a 2x2 max pooling operací s krokem 2 pro převzorkování. Při každém z těchto kroků je zdvojnásoben počet příznakových kanálů. Pravá strana zvaná expanzivní sestává z převzorkování feature map následovaného 2x2 konvolucí, které dělí dvěma počet příznakových kanálů. Celkový počet konvolučních vrstev v síti je 23. Aby autoři zajistili bezešvý výstup segmentace je nutné zajistit, aby vstupní snímek měl všechny strany stejných rozměrů. Samotný trénink sítě využívá *stochastic gradient descent* Caffe implementaci. S cílem využít maximum GPU, autoři preferují využití větších dlaždic nad využitím větší batch size.



Obrázek 9: Architektura U-Net

5 Praktická část

Obecné kroky předzpracování vstupních dat jsou popsány v sekcích 3.3.1 a 3.3.2. V této části jsou podrobně popsány další nezbytné úpravy pro to, aby byla data kompatibilní pro vstup do vybraných architektur. Níže jsou vyjmenovány použité softwary a funkce pro dosažení cílených úprav. Dále jsou popsány kroky, jak je pracováno s vybranými architekturami. Použitým programovacím jazykem je Python. Bylo vyzkoušeno několik *IDE (integrated development environment* = integrované vývojové prostředí) pro Python. Jedná se o uživatelsky přívětivý software, ve kterém lze kódy zobrazovat, upravovat a i tvořit. Každá z architektur využívá jiné knihovny, které jsou popsány včetně jejich hlavních odlišností.

5.1 Příprava dat

Po získání dronem jsou surová data zpracována pomocí radiometrických a geometrických korekcí včetně georeferencování (sekce 3.3.1 Předzpracování dat). Následně jsou data upravena s ohledem na jejich spektrální a prostorové rozlišení. I přese všechny tyto změny je nutné udělat ještě několik kroků pro získání správných vstupů do vybraných architektur. Úpravy se částečně liší podle toho, do jaké architektury data následně vstupují. První z nich je síť Capsule Network od autorů Paoletti, M., et al. (2019a) a druhou je metoda U-Net použitá v bakalářské práci Dvořák, J. (2020). K přípravám jsou využity následující softwary: ENVI 5.1, ArcMap 10.8, QGIS Desktop 3.12.1 a MATLAB R2020a. V nich jsou zpracovány všechny typy vstupních dat, která jsou k dispozici. Z původního formátu *tif* a *shp* je následujícími úkony dosaženo formátu *mat* a vytvoření dlaždic.

Prvním krokem je zobrazení dat v SW ENVI, kde jsou data zobrazena a jsou zkontrolovány jejich parametry. Následně jsou snímky otevřeny v softwaru ArcMap. Vstupem jsou dva typy dat. Prvním je HS snímek (s redukováným rozlišením, viz sekce 3.3.2.2), druhým typem jsou vektorová data. Jedním z vektorových souborů jsou trénovací data (dále TR data) ve formě polygonů představujících zkoumané třídy, které jsou podrobně popsány v kapitole 3.3.3. Druhým souborem je opět vektor, představující konkrétní vymezení území (dále jen plocha), podle něhož byl oříznut HS snímek a uložen do formátu *tif*. Posledním vektorovým souborem je sada validačních bodů, které slouží pro ověření přesnosti provedené klasifikace. TR data byla převedena na rastr s označenými třídami od 1 do 10 a NoData jako 0. Toho je dosaženo následujícími kroky. Nejprve se do atributové tabulky TR dat přidá sloupec s názvem Třídy, do kterého se přiřadí čísla jednotlivým třídám. Poté je použita funkce Polygon to Raster s polem hodnot (value field) nastaveným na třídy a velikostí buňky 0,09. To odpovídá požadované hraně pixelu 9 cm. Protože je potřeba, aby celé území bylo vyplněné, a to hodnotami nula za NoData

nebo hodnotami 1–10 za třídy je nutné ještě použít funkci Reclassify. Ta umožní ve vytvořeném rasteru přenastavit hodnoty pro jednotlivé třídy tak, aby odpovídaly výše zmíněným hodnotám. Následuje export do formátu tif. Je třeba dbát na to, aby mřížky HS a TR rastrů sobě odpovídaly.

Následující úprava rozměrů souvisí s cílem rozdělit data na dlaždice o velikosti 256 x 256 pixelů. Musí tak být splněna podmínka, že rozměry snímků jsou dělitelné beze zbytku číslem 256. Tento krok je zajištěn v SW ENVI pomocí funkce Resize. Zde je možné nastavit pro HS i TR data jejich potřebné rozměry. Redukovaná data na velikost 1024 x 1024 pixelů a pro originální 3072 x 3072 pixelů a vše následně vyexportovat do formátu tif.

V této formě jsou data otevřena v SW QGIS, kde je funkce Retile (znovu nadlaždicovat). Ta rozdělí připravené snímky (HS snímek a TR data) na dlaždice o velikosti 256 x 256 pixelů a s překryvem 128 pixelů. Tento krok je vyžadován na vstupu do obou vybraných architektur. V případě Capsule Network je posledním krokem přípravy převedení dat do formátu mat, což je možné v MATLAB. Příkazem `A = imread('navez_snimku.tif');` se dlaždice načtou a příkaz `save('novy_navez_snimku.mat','A');` uloží data do nového formátu.

Posledním krokem zůstává použití funkce Stack layers (SW ENVI) pro vytvoření multitemporálního kompozitu. Zvolené snímky z několika období jsou spojeny do jednoho, který bude možné pustit do sítě.

5.2 Práce s kódy

Jak vyplývá z výše uvedených informací, pro tuto práci jsou použity tři přístupy. Jejich autoři zvolili odlišné postupy pro dosažení výsledků (Kapitola 4.2). Stejně tak každý z nich se rozhodl pro jiné sdílení a otevřenost kódů. Z toho důvodu je i rozdílné řešení vzniklých problémů. V následující sekci je popsán přístup k vybraným architekturám, jejich otevření, spuštění a v krátkosti nastíněny chyby a problémy, ke kterým se dospělo. Programovacím jazykem je Python.

Vzhledem k dostupnosti kódu na platformě GitHub (odkaz viz Zdroje – Internet) byl nejprve použit kód od autorů Zhu, K., et al. (2019) využívající *Deep Convolutional Capsule Network*. Použité datové sady jsou přístupné z webové stránky viz Zdroje – Internet – Datové sady. Mimo možnost stažení dat je zde i jejich stručný popis včetně počtu tříd, spektrálních pásem a případně i provedených úpravách. Pro spuštění kódů bylo nakonec využito prostředí *Anaconda Navigator*. Pro správnou funkčnost architektury byla zásadní instalace knihoven TensorFlow s rozšířením Keras. K dispozici jsou tři skripty. První částí je soubor *data_prepare.py*, kde, jak již z názvu vyplývá, se data načítají včetně jejich parametrů a připravují se pro vstup do hlavní části. V části *load_data* lze změnit data set, který se načte. Následuje *capsulelayers_keras.py*, tj. část, která má za úkol připravit vrstvy tvořící síť – Conv-

Capsule, Class Capsule a PrimaryCap. Tyto vrstvy jsou podrobně popsány v kapitole 4.2 Vybrané metody. Poslední a zároveň hlavní částí architektury je *CapsuleNet_Keras.py*, do které jsou načteny dvě předchozí části, a tvoří hlavní část samotné sítě. Nastavuje se v ní cesta k data setům a je možné v ní upravit i některé parametry jako například počet tříd a epoch, batch size a počet opakování. Autoři neopomenuli ani zajistit, po úspěšném proběhnutí, vypsání hodnot jako je OA, AA, kappa, výpočetní čas a další.

Po odstranění množství problémů spojených s verzí jednotlivých knihoven a jejich kompatibilitou, úpravy dat a drobných úprav kódu došlo po spuštění k chybě související s proměnnou *training* a hodnotou, kterou nabývá. Na ní pak následně kód padá. Vzhledem, k tomu, že se nepodařilo navázat kontakt s autory sítě, bylo od tohoto řešení upuštěno.

Druhá použitá architektura založená na *Capsule network* publikovaná Paoletti, M., et al. (2019a) využívá knihovnu PyTorch. Zdrojové kódy byly na základě e-mailové komunikace s Dr. Paoletti po delší době zpřístupněny na GitHub (<https://github.com/mhaut/CapsNet-HSI>). Architektura se skládá ze čtyř částí, kde první z nich *auxil.py* načítá a připravuje vstupní data. Druhou částí je *hyper_pytorch.py*, která ještě pracuje se vstupními daty a připravuje je pro další zpracování. Další *model.py* s tvorbou samotné sítě a jejích vrstev jako jsou *Dense Capsule*, *Primary Capsule* a *Capsule Net*. Poslední část architektury je *main.py*, který bere výstupy ze všech předešlých a tvoří finální část sítě. Podstatnou součástí je i odstavec kódu o využití CUDA, který má důležitou roli ve funkčnosti celého modelu. Taky je zde možné upravovat, která vstupní data budou použita včetně některých parametrů jako jsou batch size, počet epoch a počet opakování. Množství ostatních vlastností je bráno přímo ze vstupních dat.

Po několika neúspěšných pokusech o rozběhnutí sítě pod OS Windows (i přes úspěšnou instalaci požadované podpory Nvidia CUDA a spuštění virtuálního Linuxového prostředí), bylo po konzultaci s autory přistoupeno k instalaci v OS Linux na PC s grafickou kartou umožňující výpočty s podporou GPU. Ve vývojovém prostředí PyCharm byl po zmenšení velikosti parametru batch size úspěšně spuštěn výpočet na volně dostupném data setu Indian Pines. Následným pokusem bylo spustit data z tundry, u kterých se však narazilo na chybu s jejich načtením, kterou se nepodařilo vyřešit – konkrétně *RuntimeError: index 7 is out of bounds for dimension 1 with size 2*, která pravděpodobně souvisí s nastavením některých z parametrů vstupních dat. Kapitola výsledky tak pouze obsahuje testy na zmíněné volně dostupné datové sadě.

Zdrojové kódy architektury U-Net jsou převzaty od studenta Jakuba Dvořáka (Dvořák, J., 2020), který zvolenou metodu testoval v bakalářské práci na MS snímcích. Cílem je metodu otestovat na podrobnějších HS datech a výsledky porovnat s výsledky jiných klasifikací nad

daty z krkonošské tundry. Architektura sestává ze dvou částí. Jak bylo zmiňováno výše, vstupem je snímek rozdělený na dlaždice 256 x 256 pixelů. V Anaconda Navigator je otevřeno webové interaktivní prostředí *Jupyter notebook* a v něm, po nastavení správné cesty do adresáře, první část architektury *U_Net_Radova2021_CZ.ipynb*. Zde jsou pak postupně spouštěny jednotlivé kroky kódu. V některých je nutné nastavit například cestu do adresáře s připravenými dlaždicemi, název výstupu, a především také volitelné parametry sítě. V prvním kroku si síť načte data a vypíše počet dlaždic a pásem. Dále načítá snímky do stanoveného formátu, načítá si další údaje o datech, využívá knihovnu TensorFlow, připravuje U-Net a předpřipravuje trénování dat. Důležitým krokem je nastavení vstupních parametrů, které jsou volitelné, jako je počet epoch, počet zpracovávaných tříd, váhy pro jednotlivé třídy, které jim přiřadí důležitost, a počet cross validací. Je možné změnit i hodnotu batch size, jejíž hodnota je často omezena velikostí GPU. Následně probíhá samotné trénování. Výsledný natrénovaný model se načte a je vstupem do samotné klasifikace. Následně je možné zjistit řadu hodnot o přesnosti výsledku jako například počet pixelů zařazených do tříd po klasifikaci, OA, průměrné F1 skóre, chybovou matici, a především výsledný snímek. Ten je v této fázi stále v podobě dlaždic. K jejich spojení slouží druhá část *Georeference_results_gdal*. Zde jsou výsledky spojeny do jediného rastru a vyexportovány do cílové složky ve formátu tif.

6 Výsledky

Výsledků bylo dosaženo v několika krocích. Nejprve byl sestaven přehledný rámec metod CNN, z něhož byla následně vytipována vhodná metoda pro klasifikaci HS snímků s vysokým prostorovým rozlišením a tříd s malou spektrální odlišností. Zvolenou metodou je Capsule Network. I tato metoda zahrnuje více přístupů, jak dosáhnout požadovaného cíle. Pro účely této práce tak byly zvoleny dvě různé architektury s cílem následně porovnat jejich výsledky a přínosy. Kvůli problémům popsaným v kapitole 5.2 byl pouze studován vliv známých vstupních parametrů (počet epoch, batch size) sítě publikované Paoletti, M. et al. (2019a) na výkon této sítě a přesnost klasifikace na datové sadě Indian Pines (IP), jak je patrné z Tabulky 3.

Tabulka 3: Trénovaná síť IP s různými parametry

Počet epoch	Batch_size	Final train loss	Přesnost	Výpočetní čas [m:s]
20	20	0.0844	0.9796	11:31
50	20	0.0673	0.9896	26:25
100	20	0.0550	0.9890	52:27
20	60	0.0897	0.9784	10:57
50	60	0.0664	0.9924	25:32
100	60	0.0553	0.9925	51:08
20	97	0.0964	0.9759	10:04
50	97	0.0698	0.9924	25:55
100	97	0.0589	0.9947	50:53

Jak vyplývá z výše uvedené tabulky, počet *epoch* a *batch size* mají vliv na časovou náročnost výpočtu a na dosaženou přesnost. Tyto proměnné se nazývají hyperparametry sítě a jsou několika málo proměnnými, se kterými je možné ve vybrané Capsule Network architektuře pracovat. Parametr *batch size* udává počet dopředných propagací před upravením vah uvnitř sítě. Jinak řečeno, stanovuje počet záznamů vstupních dat, které jsou zároveň využity pro jednu dopřednou propagaci (Matyáš, M., 2020). Váhy jsou upravovány na základě ztrát záznamů v jedné *batch*. Epochy se v sítích obvykle využívají ve velkém množství (desítky až tisíce), zároveň však, se zohledněním *batch size*, záleží na jejich počtu vzhledem k využití množství procesní paměti. Epochou je nazývána situace, kdy jsou využity všechny záznamy vstupních dat pro naučení sítě.

Úspěšné spuštění sítě Capsule Network (Paoletti, M., et al., 2019a) proběhlo na PC s operačním systémem Ubuntu 20.04.3 LTS typu 64bitovém, s grafickou kartou NVIDIA GP107M o paměti 15,5 GB a procesorem Intel Core i7.

Dalším krokem bylo otestování klasifikace HS dat sítí U-Net. Vstupem do sítě tak byly snímky s TR daty a původní HS snímky (viz sekce 5.1), oboje rozděleny na dlaždice o velikosti 256 x 256 pixelů. Po úvodním představení sítě se přikročilo k testům směřujícím k vyladění hyperparametrů (počet epoch, cross validace, váhy tříd apod.) tak, aby bylo dosaženo co nejpřesnější klasifikace. Vzhledem k tomu, že některé třídy byly v trénovacích datech zastoupeny v malém množství, prvním krokem bylo zvýšení hodnoty cross validace (*args.crossval_nfold*). Tento krok zajistil, aby se do klasifikace dostaly i právě zmíněné málo zastoupené třídy. A naopak pomohl i zajistit, aby hodně zastoupené třídy nebyly nadhodnoceny. Tomu napomohla i změna vah pro jednotlivé třídy. To znamená, že třídám s menším zastoupením se přiřadí vyšší váha, a naopak těm hodně zastoupeným se přiřadí váha nižší.

Testování nastavení parametrů probíhalo na datech z krkonošské tundry nasnímaných v srpnu roku 2020 s 9¹ třídami. Hodnoty, které byly měněny se týkaly parametrů počet epoch, váhy tříd, cross validace a batch size. U té se však brzy narazilo na limit dostupné paměti, takže nakonec konečná hodnota zůstala na čísle 2. Pro váhy tříd byly otestovány tři verze, přičemž musí být splněna podmínka, že součet všech vah je roven jedné. První rozložením bylo pro každou třídu váha 0,1. To se ukázalo jako nevhodné vzhledem k rozdílnému zastoupení tříd. Proto bylo nastaveno jiné rozložení tříd, viz Tabulka 4.

Tabulka 4: Nastavení vah tříd

	Třída 0	Tř 1	Tř 2	Tř 3	Tř 4	Tř 5	Tř 6	Tř 7	Tř 8	Tř 9	Tř 10
	NoData	af	afs	bor	desch	klec	lis	nard	sut	vres	vyfuk
Třídy	3098319	3756	2488	17583	1616	4824	352	12630	40026	23856	5814
Váhy 1	0.0	0.125	0.125	0.05	0.125	0.125	0.175	0.075	0.025	0.05	0.125
Váhy 2	0.0	0.1	0.1	0.075	0.15	0.1	0.25	0.075	0.025	0.025	0.1

¹ Při prvních testech bylo použito tříd 10 (navíc třída lis)

Výše zmíněné dvě verze rozložení vah byly odvozeny na základě získaných výsledků z trénování. Prvním ukazatelem bylo rozdělení tříd po trénování, které ukázalo, zda jsou zastoupeny všechny třídy a s jakým množstvím pixelů. Na základě znalosti původního zastoupení tříd, tak bylo možné odhadnout, zda je nové zastoupení relevantní. Druhým ukazatelem byla chybová matice, která ukazuje, jaké třídy jsou zaměněny a na základě toho lze upravit váhy pro jednotlivé třídy. Příkladem je Obrázek 12 s chybovou maticí za období srpen z trénování pokusu 5. Zde si lze všimnout, že největší chyby ve správném zařazení pixelů jsou mezi třídami 2; 4; 7; 10 a 6, která ale pro další trénování byla vynechána.

[[3746	8	0	0	0	2	0	0	0	0	0]
[71	2365	36	0	0	0	0	2	13	1]	
[0	0	17583	0	0	0	0	0	0	0]	
[0	52	0	1091	0	0	66	407	0	0]	
[0	0	0	0	4824	0	0	0	0	0]	
[1	17	5	0	0	215	0	61	25	28]	
[46	816	5	362	0	0	11401	0	0	0]	
[0	0	0	0	0	79	0	39926	0	21]	
[3	37	13	0	0	0	0	100	23703	0]	
[6	4	32	150	12	6	0	331	0	5273]	

Obrázek 10: Chybová matice po trénování pokus 5

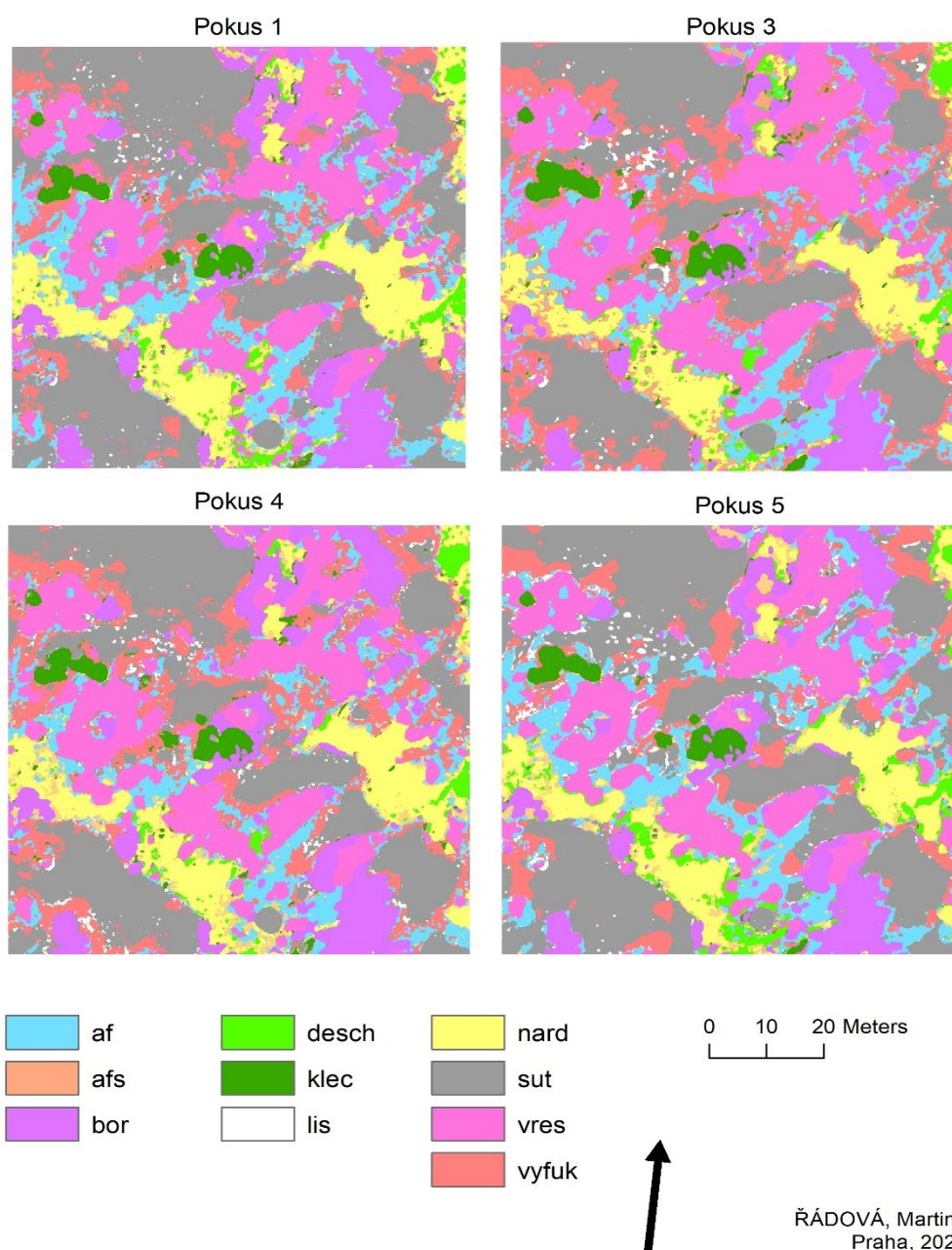
Počet epoch byl upravován na základě grafu získaného z trénování sítě, ze kterého bylo možné vypočítat, zda zvyšování počtu epoch má ještě smysl nebo ne. Z výsledků byly získány i celková přesnost a F1 skóre za každou cross validaci. Na základě těchto hodnot mohl být vybrán nejpřesnější model, který byl následně vstupem do klasifikace. Mimo to byly získány i celková přesnost a F1 skóre za celou trénovanou síť.

V prvních pokusech bylo použito rozdělení vah 2 dle Tabulky 4, potom bylo rozdělení změněno na váhy 1. Společně s těmito hodnotami byly, jak je zmíněno výše, měněny parametry počet epoch a cross validace. Vše je shrnuto v následující Tabulka 5 společně s časem, který trénování zabralo, a výslednou dosaženou OA. Na základě těchto výsledků byly vybrány parametry, které byly následně použity pro finální trénování a klasifikaci. K výběru a zhodnocení napomohlo i vizuální zhodnocení doposud získaných trénovacích snímků (Obrázek 13). V úvahu se bralo rozložení tříd, jejich promíchanost a hladkost celého snímku. Na základě všech těchto výsledků byl zvoleny hodnoty z pokusu 5.

Tabulka 5: Ladění parametrů pro ostrou klasifikaci

	Počet epoch	Cross validace	Čas	Váhy	OA [%]
Pokus 1	100	6	51 min 56 s	2	96.51
Pokus 2	150	8	1h 49 min	2	97.44
Pokus 3	125	5	52 min 17 s	2	96.38
Pokus 4	150	6	1h 17 min	2	95.93
Pokus 5	150	8	1h 51 min	1	97.50

KLASIFIKACE HS SNÍMKŮ KRKONOŠKÉ TUNDRY METODOU U-NET ladění parametrů



Obrázek 11: Výsledné snímky trénovacích klasifikací

Po všech výše zmiňovaných testech parametrů bylo možné přistoupit k samotnému finálnímu trénování a klasifikaci. Parametry byly zvoleny viz Tabulka 6. Celkový počet tříd je 10, protože je nutné zahrnout i třídu 0 - NoData.

Tabulka 6: Parametry pro vstup do klasifikace

Počet epoch	Batch size	Počet tříd	Počet cross validací
150	2	10	8

Třída	NoData	af	afs	bor	desch	klec	nard	sut	vres	vyfuk
Hodnota váhy	0.0	0.15	0.175	0.05	0.175	0.1	0.1	0.05	0.05	0.15

Klasifikace byly provedeny v termínech srpen a září v roce 2020 (viz Tabulka 1) a z nich byl následně, jejich spojením, vytvořen multitemporální kompozit. Úspěšně proběhly všechny klasifikace kromě první za červenec. Červencový termín byl i po pokusech o opravy vynechán z důvodu velmi nízkých výsledných přesností. Chyba byla pravděpodobně v původních datech, které do klasifikace vstupovaly.

Z vyklasifikovaných snímků byly vytvořeny finální mapy, které jsou přiloženy v přílohách 7; 8 a 9. Výsledky byly porovnávány s výsledky projektu Vegetace krkonošské tundry – minulost současnost a budoucnost, konkrétně její částí zaměřenou na vývoj metod pro monitoring změn vegetace tundry za využití MS, HS a LIDAR senzorů nesených UAV. V obou případech se jednalo o snímky s 54 spektrálními pásmy a prostorovým rozlišením 9 cm. V této práci byly úspěšně zpracovány snímky za měsíce srpen a září roku 2020 a multitemporální kompozit z těchto dvou měsíců. V projektu využili snímky z měsíců červen až září, zpracované metodami ML, RF, SVM a OBIA. Kompozity vytvořili za měsíce červenec až září. Co se týkalo parametrů, pro každou z metod byly na vstupu jiné, proto není zcela možné je porovnávat. Pro ověření přesnosti jednotlivých tříd a celé klasifikace sloužila sada validačních dat. Tato sada sestávala z množství v terénu získaných bodů, které se vztahují ke konkrétnímu místu a zastupují danou konkrétní třídu. Ověření proběhlo v SW ArcMap, kde byly pomocí funkce *Extract values to points*, přiřazeny hodnoty z klasifikovaného rastru do vrstvy validačních dat. Z těchto dat byla následně vytvořena chybová matice. Z ní byly odvozeny OA pro celý snímek, PA (producer's accuracy) a UA (user's accuracy) pro jednotlivé třídy. Všechny tyto hodnoty jsou přehledně sepsány v Tabulce 9. Stejně hodnocení přesnosti byl použito i v rámci projektu pro výsledky jednotlivých metod. Vše je doplněno chybovými maticemi za jednotlivé měsíce.

V této diplomové práci za zvolené území z výsledků samotné klasifikace vyšla nejlepší hodnota OA 97.82% pro multitemporální kompozit. Nezaostaly ani přesnosti za srpen a září, které byly jen o přibližně 1% horší. Ověření přesností bylo realizováno pomocí chybových matic, které vytvořeny v SW ArcMAP z validačních bodů a vyklasifikovaných snímků. Tyto tabulky převážně potvrdily přesnost výstupů. Nejlepší zaznamenaná hodnota vyšla opět za multitemporální kompozit s OA 93.62%. Opět nezaostaly ani výsledky pro jednotlivé termíny, které vyšly 92.75% pro srpen a 93.1% pro září. V rámci projektu vyšly nejlepší hodnoty pro HS data s klasifikační metodou ML.

Důležitým ukazatelem, vyčteným z chybových matic byly i uživatelská (user's accuracc - UA) a zpracovatelská přesnost (producer's accuraccy). Tyto hodnoty vyšly nejlépe za všechna období pro třídy kleč a suť s průměrnou přesností přes 99%. Těsně za nimi s vysokými přesnostmi byly i třídy bor a nard. Tyto třídy většinou ani nebyly zaměněny za žádné jiné třídy, především kleč, a ani jiné třídy nebyly zaměňovány za ně. Pouze v případě vyfuk, který byl ve všech sledovaných obdobích zaměněn za suť. V ostatních případech se jednalo o záměnu pouze v jednotkách pixelů. Naopak nejhůře vyšly třídy afs, desch a af jejichž přesnosti se pohybovaly v rozmezí 55–84% pro UA a 53–88% pro PA. Jejich nízké přesnosti by mohly být způsobeny jednak jejich nízkým zastoupením v trénovacích datech a také tím, že se jedná převážně o třídy travin, které mají velmi podobné spektrální vlastnosti. I z toho důvodu je zajímavé tyto třídy sledovat v průběhu různých období, kdy se mění jejich fenologické fáze a s tím související případná změna ve spektrální odrazivosti. Jelikož pozorování probíhá v období několika let, bude zajímavé zjistit, zda bude možné vypočítat jisté vzorce, které by se ve výsledcích objevovaly. Výše zmíněné třídy, které vyšly nejhůř, jsou, na základě chybových matic, nejčastěji zaměňovány mezi sebou. To znamená af za afs a opačně. Dále jsou taky za af ve značné míře zaměněny i třídy vřes a vyfuk.

Všechny klasifikace byly realizovány na počítači s OS Windows 10 Enterprise verze 10.0.19042. Procesor Intel® Core™ i7-10700KF CPU 3.80GHz s RAM pamětí 32GB a na základu 64 bitovém.

Tabulka 7: Výsledky klasifikací

	Čas	OA trénovací	Mean F1	Třídy	OA matice	UA	PA
Srpen	1h 48 min	96.85 %	90.96 %	af	92.75%	79.35%	53.68%
				afs		55.69%	62.86%
				bor		97.27%	91.71%
				desch		77.42%	63.16%
				klec		92.23%	100%
				nard		92.17%	98.48%
				sut		99.91%	97.71%
				vres		89.18%	94.71%
				vyfuk		79.17%	93.96%
Září	1h 49 min	96.89%	92.57%	af	93.1%	84.88%	49.32%
				afs		55.69%	69.84%
				bor		99.07%	92.49%
				desch		72.73%	88.89%
				klec		100%	100%
				nard		94.93%	97.04%
				sut		99.91%	96.69%
				vres		88.41%	99.28%
				vyfuk		74.89%	92.39%
MTK	1h 52 min	97.82%	93.12%	af	93.62%	86.90%	54.07%
				afs		58.67%	75.86%
				bor		97.86%	92.75%
				desch		68.57%	82.76%
				klec		97.94%	100%
				nard		97.37%	94.18%
				sut		100%	96.69%
				vres		90.35%	98.80%
				vyfuk		74.67%	96.61%

Všechny klasifikace byly realizovány na počítači s OS Windows 10 Enterprise verze 10.0.19042. Procesor Intel® Core™ i7-10700KF CPU 3.80GHz s RAM pamětí 32GB a na základu 64 bitovém.

Tabulka 8: Chybová matice srpen

Třída	1	2	3	4	5	6	7	8	9	Celkem	UA
1	73	4	0	0	0	0	0	9	6	92	79.35%
2	25	44	3	1	0	0	0	5	1	79	55.69%
3	1	0	321	1	0	0	0	7	0	330	97.27%
4	1	2	0	24	0	4	0	0	0	31	77.42%
5	0	0	8	0	95	0	0	0	0	103	92.23%
6	0	9	0	12	0	259	0	0	1	281	92.17%
7	0	0	0	0	0	0	1108	0	1	1109	99.91%
8	23	7	18	0	0	0	0	412	2	462	89.18%
9	13	4	0	0	0	0	26	2	171	216	79.17%
Celkem	136	70	350	38	95	263	1134	435	182	2703	
PA	53.68%	62.86%	91.71%	63.16%	100%	98.48%	97.70%	94.71%	93.96%	OA	92.75%

Tabulka 9: Chybová matice září

Třída	1	2	3	4	5	6	7	8	9	Celkem	UA
1	73	6	0	1	0	0	0	0	6	86	84.88%
2	24	44	8	0	0	3	0	0	0	79	55.69%
3	0	1	320	0	0	0	1	1	0	323	99.07%
4	0	2	0	24	0	5	0	0	2	33	72.73%
5	0	0	0	0	95	0	0	0	0	95	100%
6	4	8	0	2	0	262	0	0	0	276	94.93%
7	0	0	0	0	0	0	1111	1	0	1112	99.91%
8	30	0	18	0	0	0	0	412	6	466	88.41%
9	17	2	0	0	0	0	37	1	170	227	74.89%
Celkem	148	63	346	27	95	270	1149	415	184	2697	
PA	49.32%	69.84%	92.49%	88.89%	100%	97.04%	96.69%	99.28%	92.39	OA	93.1%

Tabulka 10: Chybová matice multitemporální kompozit

Třída	1	2	3	4	5	6	7	8	9	Celkem	UA
1	73	4				1	1		5	84	86.90%
2	22	44	5			4				75	58.67%
3	1	2	320					4		327	97.86%
4		4		24		7				35	68.57%
5			2		95					97	97.94%
6		3		4		259				266	97.37%
7							1111			1111	100%
8	22		18	1		2		412	1	456	90.35%
9	17	1				2	37	1	171	229	74.67%
Celkem	135	58	345	29	95	275	1149	417	177	2680	
PA	54.07%	75.86%	92.75%	82.76%	100%	94.18%	96.69%	98.8%	96.61	OA	93.62%

7 Diskuse

Z výzkumu v rámci rešerše práce byla vytipována metoda Capsule Network s implementací v programovacím jazyce Python. Použity byly dva přístupy této metody – od autorů Paoletti, M., et al. (2019a) *Capsule Networks for Hyperspectral Image Classification* a od autorů Zhu, K., et al. (2019) *Deep Convolutional Capsule Network for Hyperspectral Image Spectral and Spectral-Spatial Classification*. V první řadě bylo úkolem použít jejich trénovací data a samotnou síť na nich úspěšně spustit. V průběhu tohoto procesu, u obou přístupů, se narazilo na množství chyb. Autoři v popise zpřístupněných kódů explicitně neuvádějí verze použitých knihoven a operačních systémů, požadavků na HW, vstupní parametry apod., jejichž postupné hledání způsobilo prodlevu v řešení vlastní aplikace metod na cílová data. Předcházet těmto problémům by mohla pomoci užší spolupráce s autory metody, kteří by mohli poskytnout podrobnější vzhled a instrukce pro práci s danou architekturou. Řešení bylo ztíženo o fakt, že znalost nejen samotného programovacího jazyku Python, ale i prostředí s ním spojenými, byla pouze základní. Cíly do budoucna tak zůstává navázání hlubší spolupráce především s autory sítě Paoletti et al. (2019a) a dotažení zpracovávané architektury do zdárného konce. Tím je míněno především správně předpřipravit snímky z krkonošské tundry pro vybraná území a časová období, pro vstup do sítě a získat klasifikované snímky společně s výstupními hodnotami v podobě hodnocení přesnosti trénování, průběhu ztrátové funkce v závislosti na počtu epoch apod. Samozřejmostí by byla jejich validace a porovnání s doposud dosaženými výsledky nejen v rámci projektu, ale i samotných autorů.

Především zajímavým by mohlo být hlubší porovnání architektur Capsule Network a U-Net, jejíž použití pro klasifikaci HS snímků krkonošské tundry bylo důležitou součástí práce. Metoda byla již dříve testována nad MS daty z krkonošské tundry, a proto bylo zajímavé porovnání se zpracováním HS dat. V první řadě bylo nutné vyladit některé parametry, ale nakonec došlo k úspěšnému trénování a klasifikaci dat. Z dosažených výsledků (podrobně viz kapitola 6) vyplývá, že metoda U-Net byla přínosem pro klasifikaci HS dat s vysokým prostorovým rozlišením. Dosaženými přesnostmi překonala tradičně používané metody jako jsou SVM, ML, RF a OBIA, které byly používány v rámci projektu v Krkonoších. Jejich nejlepší výsledek dosáhl OA 88.33% pro multitemporální snímek července až září. Nezaostaly ani kompozity pro červen až září s OA 87.53% a kompozit pro červenec a srpen s OA 87.33%. Srovnatelných výsledků dosáhly i multitemporální kompozity klasifikované metodu RF. Nejlepších hodnot v projektu nad monotemporálními snímky dosáhla metoda RF. Dosažené přesnosti pro všechny termíny se pohybovaly mezi 83 a 84 %. Hodnot nad 80% dosáhly i metody ML a OBIA pro termíny červenec až září. Nejhůře dopadla metoda SVM, jejíž přesnosti

pro všechna období nepřesáhly 79% OA. Porovnány byly i výsledky zařazení pixelů do jednotlivých tříd. Dosažené závěry této DP potvrzují i výsledky projektu, kde píšou, že nejlépe vyšly třídy kleč, suť, bor, nard a vřes. Nejhůře zařazené třídy pro ně vyšly af, afs a vyfuk. Liší se zařazení u třídy desch, která ve výsledcích této práce byla zařazena jako horší, protože nedosáhla příliš vysokých přesností (68-88% UA a PA). V projektu byla zařazena mezi dobře klasifikované, přičemž za všechna období a za všechny metody nepřesáhla přesnost (UA, PA) 83.7%. Její nejhorší hodnota byla 8.14% PA pro červenec s metodou OBIA.

Vzhledem k úspěchům, kterých metoda U-Net dosáhla, zejména v porovnání s výsledky projektu, by bylo zajímavé ji otestovat i na ostatních testovaných územích a pro všechna období. Možnými kroky do budoucna by mohlo být i rozsáhlejší testování parametrů, které by bylo zajímavé dovést až do "limitních" hodnot, které by bylo možné testovat, ale zároveň sít nepřeučit. Tento proces bývá časově náročnější, proto nebyl pro tuto práci zcela realizovatelný. Úkolem by byl i pokus o zpracování původních nepřevzorkovaných HS dat (269 pásem, 3cm). Pokus o jejich klasifikaci v této DP byl, avšak trval desítky hodin a z časových možností tak nebyla příležitost otestovat vstupní parametry, a výsledek vyšel nepřesně. Problém se vyskytl i s přípravou těchto dat kvůli jejich velikosti.

Bylo zajímavé porovnat výsledky této práce i s jinými publikacemi věnujícími se klasifikaci HS snímků metodami CNN. Prvním porovnávaným článkem byla bakalářská práce Jakuba Dvořáka (Deep learning for tree line ecotone mapping from remote sensing data) (Dvořák, J., 2020), která využívala metodu U-Net pro klasifikaci MS snímků z krkonošské tundry. Klasifikovaným cílem byly tři kategorie borovice kleč, smrk ztepilý a pozadí. Klasifikované snímky byly s rozlišením 0,2m a před vstupem do sítě rozděleny na dlaždice o velikosti 512 x 512 pixelů bez překrytí. Upravovanými vstupními parametry byly learning rate, počet epoch a počet cross validací. Na jejich základě byly prezentovány výsledky, které dosáhly OA 57%, 66% a 73%, podle poměru trénovacích a validačních dat (zleva nejméně trénovacích). V této DP byl zvolen poměr 80% trénovacích a 20% validačních dat, což odpovídá rozložení v citované bakalářské práci výsledku 73% OA. V případě této práce vyšly OA přes 90%. Což opět ukazuje na to, že využití metody U-Net pro klasifikaci HS snímků je značným přínosem s dobrými dosaženými přesnostmi.

Díky těmto přesnostem bylo možné si troufnout porovnat zvolenou metodu i s výsledky od autorů Paoletti, M., et al (2019a) a Zhu, K., et al. (2019) a jejich Capsule Networks. Oba tyto autoři testovali metodu na několika různých datových sadách s různým počtem trénovacích vzorků a porovnávali ji s dalšími klasifikačními metodami CNN. Pro porovnání v této práci budou vzata v úvahu pouze metoda Capsule Network s 300 trénovacími vzorky. Autorem Zhu

a jeho kolegy bylo dosaženo nejlepší OA přes 99% pro datasey Salinas a Kennedy Space Centre, které jsou oblastmi zemědělskými s klasifikací zaměřenou na vegetaci. O něco hůř vyšel dataset Houston s OA 94%. I autorka Paoletti se svými kolegy dosáhla vysokých přesností přes 99% OA s datasey Indian Pines, University of Pavia a SV, zaměřenými na zemědělské oblasti. S tím souvisí i vysoká přesnost zařazení pixelů do jednotlivých tříd. Ačkoliv metoda U-Net nedosáhla tak vysokých přesností, pořád její výsledky stačí na to, aby mohla být s výše uvedenými metodami srovnávána. Do budoucna tak bude jistě velmi zajímavé a přínosné zprovoznění metody Capsule Network nad daty z krkonošské tundry a porovnání s výsledky metody U-Net.

8 Závěr

Přínosem pro tuto diplomovou práci byla možnost podílet se na projektu Vegetace krkonošské tundry – minulost, přítomnost a budoucnost, probíhajícího od roku 2019, díky kterému byla získána i zpracovávaná dat. Díky tomu byla příležitost s těmito daty pracovat od začátku, vidět, jak se snímaly dronem, přes georeferencování až po práci s jejich původní i zmenšenou finální verzí.

V rešeršní části práce se podařilo podat přehled metod CNN pro zpracování HS snímků. Ze sestaveného přehledného rámce byla vytipována metoda, svou skladbou vhodná pro účely této práce. Zvolená architektura Capsule Network byla získána ve dvou verzích přístupu. Snahou bylo oba tyto přístupy prostudovat a úspěšně spustit na trénovacích datech a následně na zpracovaných snímcích z krkonošské tundry. Vzhledem k limitům operačního systému, HW a nekompatibility některých knihoven nebylo zcela dosaženo stanoveného cíle. Po mnoha pokusech se nakonec podařilo spustit architekturu od autorů Paoletti, M., et al. (2019a) nad jejich trénovacím datasetem IP. Snahou bylo provést stejný proces i nad zmenšenými daty z krkonošské tundry, avšak nepodařilo se překonat překážky s kompatibilitou a nastavením správných vstupních parametrů.

Nejen z toho důvodu byla využita metoda U-Net, která se podařila úspěšně spustit, včetně nastavení a otestování vstupních parametrů, a získat požadované výsledky. Splněn byl i úkol týkající se porovnání výsledků této DP s výstupy z projektu. Dosaženo bylo poměrně vysokých celkových přesností, pro všechna období a multitemporální kompozit nad 92% OA. Z čehož vyplývá, že zvolená U-Net metoda představuje zlepšení oproti používaným metodám v projektu. V něm nejlepší přesnost dosahovala pouze 88% OA. Zlepšení bylo dosaženo i v zařazení jednotlivých pixelů do tříd, které ve dvou třídách v rámci UA a PA dosáhlo přesnosti přes 99%. Jednalo se o třídy suť a kleč. Nejhorší zařazení se pohybovalo v přesnostech okolo 50%. Třídy se špatným zařazením byly af, afs a desch se záměnou za třídy af, afs, vres a vyfuk.

V průběhu práce byly postupně zodpovězeny stanovené otázky v úvodu. V první řadě bylo snahou vyhnout se zásahům do kódu architektury. Tento krok vyžaduje detailní znalost sítě, včetně jejího pozadí (matematické) a především také hlubší znalost jazyka Python a jeho principů. Proto do kódu Capsule Network zasahováno nebylo. Metoda U-Net musela být pro její úspěšné spuštění částečně modifikována. Důležitá byla úprava nastavení vstupních parametrů. Bylo provedeno několik trénování pro vyladění parametrů jako byly váhy pro jednotlivé třídy, počet epoch, cross validací a další. S takto vybranými parametry bylo následně spuštěny finální klasifikace. V případě sítě Capsule Network nebyly odhaleny všechny

parametry, které by bylo možné upravit, respektive byly pravděpodobně skryty v síti a načítány rovnou ze snímků. Ve spojitosti se vstupními snímky se u metody Capsule Network vyskytlo několik chyb, které byla snaha řešit. Snahou bylo spustit tuto architekturu nad daty z krkonošské tundry. I přes úpravu velikosti, správného překrytí a formátu vstupních snímků se metodu Capsule Network na snímcích z tundry spustit nepodařilo. Nepodařilo se připravit vstupní data do takové míry a podrobnosti, aby mohla do sítě úspěšně vstoupit. Na těchto krocích je v plánu ještě zpracovat. Oproti tomu metoda U-Net byla s HS snímky z tundry úspěšnější. Před samotným vstupem byla data připravena a rozdlaždicována po 256 x 256 pixelech. Metoda tak úspěšně rozklasifikovala převzorkovaná data (54 pásem, 9cm) s vysokými přesnostmi, jak je zmíněno výše. Byla snaha zpracovat i původní data (269 pásem, 3cm), která se podařilo spustit, avšak jejich zpracování zabralo množství času, a ještě by bylo třeba doladit samotné vstupy.

Důležitou otázkou ke zodpovězení zůstalo, zda byly zvolené metody přínosem. Z dosažených výsledků vyplývá, že metoda U-Net byla zlepšením oproti tradičně používaným metodám strojového učení. Dosáhla nejen celkově vyšších přesností, ale také zlepšení zařazení pixelů do jednotlivých tříd. I její relativně rychlý (pro menší datasets) výpočetní čas je kladným rysem. Metodu Capsule Network bude nutné v tomto ohledu ještě otestovat a zhodnotit.

I přes některé nepřekonané problémy, které zůstaly úkoly pro další práci a bádání, je tato práce přínosem. Byl na ní odveden velký kus práce s věnovaným množstvím času a snahy. Bylo získáno opravdu velké množství zkušeností a nových znalostí. Přínosem byly konzultace s odborníky (doc. Bayer, pan Omrt a další) a z nich nabyté nové informace včetně praktického uplatnění. Důležitá byla i možnost uplatnit znalosti z oblasti GIS a schopnost využít více SW pro dosažení požadovaného cíle. Cílem do budoucna je i nadále využívat získané znalosti z výzkumu v rámci řešerše a využít informace z načtené literatury pro další bádání včetně zužitkování proniknutí do programovacího jazyku Python.

9 Zdroje

9.1 Publikace

- [1] Bratsch, S., Epstein, H., Buchhorn, M., Walker, D. (2016). Differentiating among Four Arctic Tundra Plant Communities at Ivotuk, Alaska Using Field Spectroscopy. MDPI – Remote Sensing
- [2] Červená, L., Lysák, J., Potůčková, M., Kupková, L. (2020). Zkušenosti se zpracováním hyperspektrálních dat pořízených UAV. GIS Ostrava
- [3] Davidson, S., Santos, M., Sloan, V., Watts, J., Phoenix, G., Oechel, W., Zona, D. (2016). Mapping Arctic Tundra Vegetation Communities Using Field Spectroscopy and Multispectral Satellite Data in North Alaska, USA. MDPI – Remote Sensing.
- [4] Dvořák, J. (2020). Diplomová práce – Deep learning for tree line ecotone mapping from remote sensing data. Univerzita Karlova [vid srpen 2021]
- [5] Matyáš, M., (2020). Diplomová práce – Klasifikace zástavby pro účely kartografické generalizace státního mapového díla. Univerzita Karlova [vid 20.6.2021]
- [6] Fu, G., Liu, Ch., Zhou, R., Sun, T., Zhang, Q. (2017). Classification for High Resolution Remote Sensing Imagery Using a Fully Convolutional Network. MDPI – Remote Sensing.
- [7] Ghamisi, P., Maggiori, E., Li, S., Souza, R., et al. (2018). Frontiers in Spectral-Spatial Classification of Hyperspectral Images. HAL archives-ouvertes.
- [8] Ghamisi, P., Souza, R., Benediktsson, J., Rittner, L., Lotufo, R., Zhu, X. (2016). Hyperspectral Data Classification Using Extended Extinction Profiles. Ieee Geoscience And Remote Sensing Letters, Vol. 13, No. 11.
- [9] Halounová, L., Pavelka, K. (2009). Dálkový průzkum Země. ČVUT Praha, str. 169–171
- [10] Hinton, G., Krizhevsky, A., Wang, S. (2011). Transforming auto-encoders in Artificial Neural Networks and Machine Learning. Springer.
- [11] Hu, W., Huang, Y., Wei, L., Zhang, F., Li, H., (2015). Deep Convolutional Neural Networks for Hyperspectral Image Classification. Hindawi.
- [12] Hu, Y., Zhang, Q., Zhang, Y., Yan, H. (2018). A Deep Convolution Neural Network Method for Land Cover Mapping – A Case Study of Qinhuangdao, China. MDPI – Remote Sensing.
- [13] Chen, Y., Jiang, H., Li, Ch., Jia, X., Ghamisi, P. (2016). Deep Feature Extraction and Classification of Hyperspectral Images Based on Convolutional Neural Networks. Ieee Transactions On Geoscience And Remote Sensing.
- [14] Chen, Y., Li, J., Xiao, H., Jin, X., Yan, S., Feng, J. (2017). Dual Path Networks
- [15] Chollet, F. (2018): Deep learning with Python. Manning Publications Co, Shelter Island, New York.
- [16] Chytrý, M., Kučera, T., Kočí, M., Grulich, V., Lustyk, P. (2010). Katalog biotopů České republiky. AOPK ČR
- [17] Jarocińska, A., Kacprzyk, M., Marcinkowska-Ochtyra, A., Zagajewski, B., Meuleman, K. (2016). The application of APEX images in the assessment of the state of non-forest vegetation in the Karkonosze Mountains. Miscellanea Geographica – Regional Studies On Development
- [18] Jedrych, M., Zagajewski, B., Marcinkowska, A. (2017). Application of Sentinel-2 and EnMAP new satellite data to the mapping of alpine vegetation of the Karkonosze Mountains. Polish Cartographical Review Vol. 49, no. 3
- [19] Jeník, J., Štursa, J. (2003): Vegetation of the Giant Mountains, Central Europe. In: NAGY, L., GRABHERR,

- G., KÖRNER, CH., THOMPSON, D. B. A. (eds.), *Alpine Biodiversity in Europe* (Ecology Studies Vol. 167), New York: Springer, s. 47–51.
- [20] Kattenborn, T., Eichel, J., Fassnacht, F. (2019). Convolutional Neural Networks enable efficient, accurate and fine-grained segmentation of plant species and communities from high-resolution UAV imagery. *Scientific reports*.
- [21] Kolář, J. (1990). *Dálkový průzkum Země*. Str 126-127.
- [22] Krizhevsky, A., Sutskever, I., Hinton, G. (2017). ImageNet Classification with Deep Convolutional Neural Networks. *Communications of the ACM*.
- [23] Kupková, L., Červená, L., Potůčková, M., Lysák, J., Šašková, M., Šrollerů, A., Klinerová, T., Bobek, P., Müllerová, J. (2020). Project: Vegetation of the tundra in the Krkonoše Mts. – past, present and future. Part: Development of methods for monitoring of the Krkonoše Mts. tundra vegetation changes using multispectral, hyperspectral and LIDAR sensors from UAV (2019 –2023). *Annual Report*.
- [24] Kupková, L., Červená, L., Suchá, R., Jakešová, L., Zagajewski, B., Březina, S., Albrechtová, J. (2017). Classification of Tundra Vegetation in the Krkonoše Mts. National Park Using APEX, AISA Dual and Sentinel-2A Data. *European Journal of Remote Sensing*.
- [25] Kupková, L., Müllerová, J., Červená, L., Roubalová, M. (2019). Metodika klasifikace multispektrálních a hyperspektrálních dat pořízených v roce 2019.
- [26] Langford, Z., Kumar, J., Hoffman, F. (2017). Convolutional Neural Network Approach for Mapping Arctic Vegetation using Multi-Sensor Remote Sensing Fusion. *IEEE International Conference on Data Mining Workshops*.
- [27] Langford, Z., Kumar, J., Hoffman, F., Breen, A., Iversen, C. (2018). Arctic Vegetation Mapping Using Unsupervised Training Datasets and Convolutional Neural Networks. *MDPI – Remote Sensing*.
- [28] Lee, H., Kwon, H. (2017). Going Deeper with Contextual CNN for Hyperspectral Image Classification. *IEEE*
- [29] Li, S., Song, W., Fang, L., Chen, Y., Ghamisi, P., Benediktsson, J. (2019). Deep Learning for Hyperspectral Image Classification: An Overview. *Ieee Transactions On Geoscience And Remote Sensing*, Vol. 57, No. 9.
- [30] Li, Y., Zhang, H., Shen, Q. (2017). Spectral–Spatial Classification of Hyperspectral Imagery with 3D Convolutional Neural Network. *MDPI – Remote Sensing*.
- [31] Li, Z., Huang, L., He, J. (2019). A Multiscale Deep Middle-level Feature Fusion Network for Hyperspectral Classification. *MDPI – Remote Sensing*.
- [32] Liu, B., Yu, X., Zhang, P., Yu, A., Fu, Q., Wei, X. (2018). Supervised deep feature extraction for hyperspectral image classification. *IEEE Transactions On Geoscience And Remote Sensing*, Vol. 56, No. 4
- [33] Liu, P., Choo, K., Wang, L., Huang, F., (2016). SVM or deep learning? A comparative study on remote sensing image classification. *CrossMark – METHODOLOGIES AND APPLICATION*.
- [34] Luo, Y., Zou, J., Yao, C., Zhao, X., Li, T., Bai, G. (2018). Hsi-cnn A novel convolution neural network for hyperspectral image. *ICALIP*
- [35] Ma, X., Fu, A., Wang, J., Wang, H., Yin, B. (2018). Hyperspectral Image Classification Based on Deep Deconvolution Network With Skip Architecture. *Ieee Transactions On Geoscience And Remote Sensing*, Vol. 56, No. 8
- [36] Matyáš, M., (2020). Diplomová práce – Klasifikace zástavby pro účely kartografické generalizace státního mapového díla. Univerzita Karlova [vid 20.6.2021]
- [37] Marcinkowska, A., Zagajewski, B., Ochtyra, A., Jarocińska, A., Raczko, E., Kupková, L., Stych, P., Meuleman, K. (2014). Mapping vegetation communities of the Karkonosze National Park using APEX

hyperspectral data and Support Vector Machine. MISCELLANEA GEOGRAPHICA – REGIONAL STUDIES ON DEVELOPMENT.

- [38] Marcinkowska, A., Zagajewski, B., Ochtyra, A., Jarocińska, A., Wojtuń, B., Rogass, C., Mielke, C., Lavender, S. (2017). Subalpine and alpine vegetation classification based on hyperspectral APEX and simulated EnMAP images. *International Journal of Remote Sensing*
- [39] Marcinkowska, A., Zagajewski B., Raczko, E., Ochtyra, A., Jarocińska, A. (2018). Classification of High-Mountain Vegetation Communities within a Diverse Giant Mountains Ecosystem Using Airborne APEX Hyperspectral Imagery. *MDPI – Remote Sensing*
- [40] Mikheeva, A., Tutubalina, O., Zimin, M., GolubevaE. (2017). A Subpixel Classification of Multispectral Satellite Imagery for Interpretation of Tundra-Taiga Ecotone Vegetation. *Methods and means of processing and interpretation of space information.*
- [41] Mou, L., Ghamisi, P., Zhu, X. (2017). Unsupervised Spectral-Spatial Feature Learning via Deep Residual Conv-Deconv Network for Hyperspectral Image Classification. *IEEE Transactions on Geoscience and Remote Sensing.*
- [42] Nielson M. (2019). *Neural Networks and Deep Learning.*
- [43] Paoletti, M., Haut, J., Fernandez-Beltran, R., Plaza, J., Plaza, A., Li, J., Pla, F. (2019a). Capsule Networks for Hyperspectral Image Classification. *Ieee Transactions On Geoscience And Remote Sensing, Vol. 57, No. 4*
- [44] Paoletti, M., Haut, J., Fernandez-Beltran, R., Plaza, J., Plaza, A., Li, J., Pla, F. (2019b). Deep Pyramidal Residual Networks for Spectral–Spatial Hyperspectral Image Classification. *Ieee Transactions On Geoscience And Remote Sensing, Vol. 57, No. 2.*
- [45] Paoletti, M., Haut, J., Plaza, J. (2019c). Deep learning classifiers for hyperspectral imaging: A review. *ISPRS Journal of Photogrammetry and Remote Sensing 158.*
- [46] Patrick, M., Adekoya, A., Mighty, A., Edward, B. (2019). Capsule Networks – A survey. *Journal of King Saud University*
- [47] Pouyanfar, S., Sadiq, S., Yan, Y., Tian, H., Tao, Y., Reyes, M., Shyu, M., Chen, S., Iyengar, S. (2018). A Survey on Deep Learning Algorithms, Techniques and Applications. *ACM Computing Surveys, Vol. 51, No. 5, Article 92.*
- [48] Rezaee, M., Mahdianpari, M., Zhang, Y., Salehi, B. (2018). Deep Convolutional Neural Network for Complex Wetland Classification Using Optical Remote Sensing Imagery. *Ieee Journal Of Selected Topics In Applied Earth Observations And Remote Sensing, Vol. 11, No. 9.*
- [49] Ronneberger, O. (2015)
- [50] Sabour, S. et al. (2017). Dynamic Routing Between Capsules. *NIPS 2017*
- [51] Scott, G.J., England, R., Starms, A., Marcum, A., Davis, H. (2017). Training deep convolutional neural networks for land - cover classification of high-resolution imagery. *IEEE Geosci. Remote Sens.*
- [52] Shahroudnejad, A., Mohammadi, A., Plataniotis, K. (2018). Improved Explainability of Capsule Networks: Relevance Path by Agreement.
- [53] Shrestha, A., Mahmood, A. (2019). Review of Deep Learning Algorithms. *IEEE Access*
- [54] Song, W., Li, S., Fang, L., Lu, T. (2018). Hyperspectral Image Classification With Deep Feature Fusion Network. *Ieee Transactions On Geoscience And Remote Sensing, Vol. 56, No. 6.*
- [55] Soukupová, L., Sekyra, J., Jeník, J., Kociánová, M. (1995): Arctic-alpine tundra in the Krkonoše, the Sudetes. *Opera corcontica, 32.*

- [56] Springenberg, J., Dosovitskiy, A., Brox, T., Riedmiller, M. (2015). Striving for simplicity: The all Convolutional net. Department of Computer Science, University of Freiburg.
- [57] Suchá, R., Jakešová, L., Kupková, L., Červená, L. (2016). Classification of Vegetation above the Tree Line in the Krkonoše Mountains National Park Using Remote Sensing Multispectral Data. *AUC Geographica*
- [58] Štursa, J. (2013). Arktoalpínská tundra Krkonoš. *Živy Av ČR*.
- [59] Štursa, J., Wild, J. (2014) Kleč a smilka – klíčové hráči vývoje alpínského bezlesí Krkonoš. *Opera Corcontica* 51
- [60] Vali, A., Comai, S., Matteucci, M. (2020). Deep Learning for Land Use and Land Cover Classification Based on Hyperspectral and Multispectral Earth Observation Data: A Review. *MDPI – Remote Sensing*
- [61] Wagner, F., Sanchez, A., Tarabalka, Y., Lotte, R., Ferreira, M., Aidar, M., Gloor, E., Philips, O., Aragao, L. (2019). Using the U-net convolutional network to map forest types and disturbance in the Atlantic rainforest with very high resolution images. *White Rose*
- [62] Wang, W., Dou, S., Wang, S. (2019). Alternately Updated Spectral–Spatial Convolution Network for the Classification of Hyperspectral Images. *MDPI – Remote Sensing*.
- [63] Wu, H., Prasad, S. (2017). Convolutional Recurrent Neural Networks for Hyperspectral Data Classification. *MDPI – Remote Sensing*.
- [64] Yu, S., Jia, S., Xu, Ch. (2017). Convolutional neural networks for hyperspectral image classification. *Neurocomputing* 219, 88–98
- [65] Zagajewski, B., Krówczyńska, M., Kozłowska, A., Sobczak, M., Wrzesien, M. (2014). Mapping high mountain vegetation using hyperspectral data. *ResearchGate*.
- [66] Zhang, H., Li, Y., Zhang, Y., Shen, Q. (2017). Spectral-spatial classification of hyperspectral imagery using a dual-channel convolutional neural network. *Remote Sensing Letters*.
- [67] Zhao, Y., Yuan, Y., Song, M., Ding, Y., Lin, F., Zhang, D. (2019): Use of Unmanned Aerial Vehicle Imagery and Deep Learning UNet to Extract Rice Lodging. *Sensors*, 18, 19, 3859.
- [68] Zhong, Z., Li, J., Luo, Z., Chapman, M. (2017). Spectral–Spatial Residual Network for Hyperspectral Image Classification – A 3-D Deep Learning Framework. *Ieee Transactions On Geoscience And Remote Sensing*
- [69] Zhu, J., Hu, J., Jia, S., Jia, X., Li, Q. (2018). Multiple 3-D Feature Fusion Framework for Hyperspectral Image Classification. *Ieee Transactions On Geoscience And Remote Sensing*, Vol. 56, No. 4.
- [70] Zhu, K. et al. (2019). Deep Convolutional Capsule Network for Hyperspectral Image Spectral and Spectral-Spatial Classification. *MDPI – Remote Sensing*
- [71] Zhu, L., Chen, Y., Ghamisi, P., Benediktsson, J. (2018). Generative Adversarial Networks for Hyperspectral Image Classification. *IEEE Transactions on Geoscience and Remote Sensing*.
- [72] Zhu, X., Tuia, D., Mou, L., Xia, G., Zhang, L., Xu, F., Fraundorfer, F. (2017). Deep Learning in Remote Sensing - A comprehensive review and list of resources. *IEEE Geoscience and remote sensing magazine*.

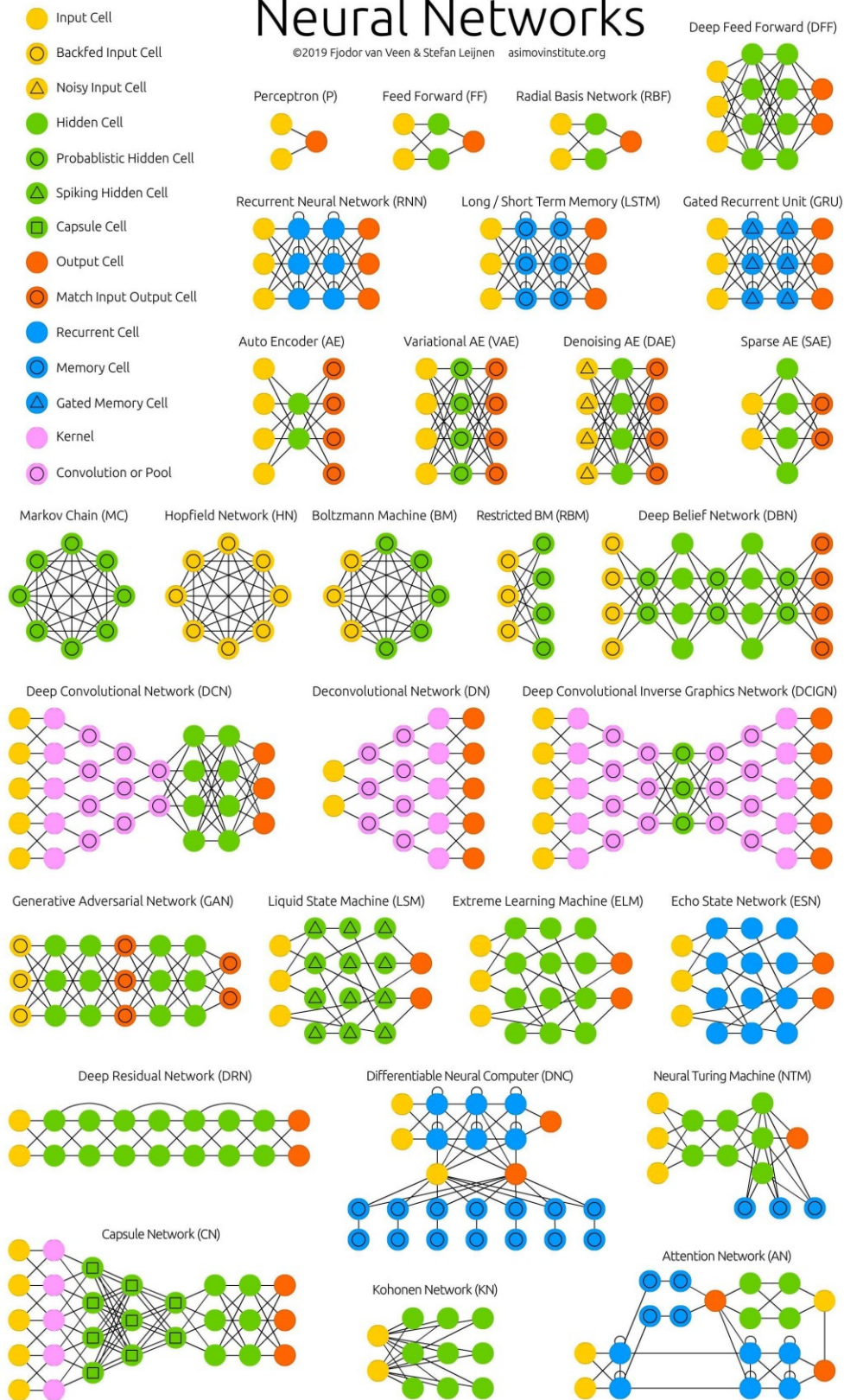
9.2 Internet

- [1] Botany.cz (2009). *Avenella Flexuosa*. Dostupné z: <https://botany.cz/cs/avenella-flexuosa/> [vid 28.6.2021]
- [2] Botany.cz (2003). *Pinus Muga Turra*. Dostupné z: <https://botany.cz/cs/pinus-mugo/> [vid 1.7.2021]
- [3] Botany.cz (2012). *Deschampsia Cespitosa*. Dostupné z: <https://botany.cz/cs/deschampsia-cespitosa/> [vid 19.7.2021]

- [4] Botany.cz (2008). Nardus Stricta. Dostupné z: <https://botany.cz/cs/nardus-stricta/> [vid 15.7.2021]
- [5] Wiki – ML (2020). Machine Learning. Dostupné z: https://en.wikipedia.org/wiki/Machine_learning [vid 13.4.2021]
- [6] Červená, L. (2016). Neřízená klasifikace. UK Praha. Str. 1. Dostupné z: <https://www.natur.cuni.cz/geografie/geoinformatika-kartografie/ke-stazeni/projekty/moderni-geoinformacni-metody-ve-vyuce-gis-kartografie-a-dpz/nerizena-klasifikace/> [vid 20.5.2021]
- [7] Datové sady (2021). Hyperspectral Remote Sensing Scenes. Dostupné z: http://www.ehu.eus/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes#Salinas [vid 15.4. 2021]
- [8] GitHub (2019). Conv-Caps-HSI-Classification. Dostupné z: <https://github.com/YushiChen/Conv-Caps-HSI-Classification> [vid 8.5.2021]
- [9] JavaTpoint (2021). Artificial Neural Network in TensorFlow. Dostupné z: <https://www.javatpoint.com/artificial-neural-network-in-tensorflow> [vid 30.6.2021]
- [10] KRNAP (2016). Vrcholy Krkonošské tundry. Dostupné z: https://www.krnep.cz/data/Files/downloads/krnap-vrcholy-krkonoske-tundry-web_148244429651.694.pdf [vid 7.5.2021]
- [11] KRNAP (2018). Krkonošská arкто-alpínská tundra. Dostupné z: <https://www.krnep.cz/krkonoska-arkto-alpinska-tundra/>. [vid 15.6.2021]
- [12] KRNAP (2015) Krkonoše známé i neznámé. https://www.krnep.cz/data/Files/downloads/krkonose-zname-nezname-web-2_146305354935.438.pdf [12.5.2021]
- [13] Nebeske (2017). Jirka Jiroušek – Luční hora. Dostupné z: <https://www.nebeske.cz/tag/lucni-hora/>
- [14] Slovník. Machine Learning Glossary. Dostupné z: <https://developers.google.com/machine-learning/glossary#p>
- [15] (UNet, 2019) <https://towardsdatascience.com/u-net-b229b32b4a71> [vid 2.9.2021]

A mostly complete chart of Neural Networks

©2019 Fjodor van Veen & Stefan Leijnen asimovinstitute.org



Příloha 2: Přehled NN. Zdroj: <https://www.asimovinstitute.org/neural-network-zoo/>

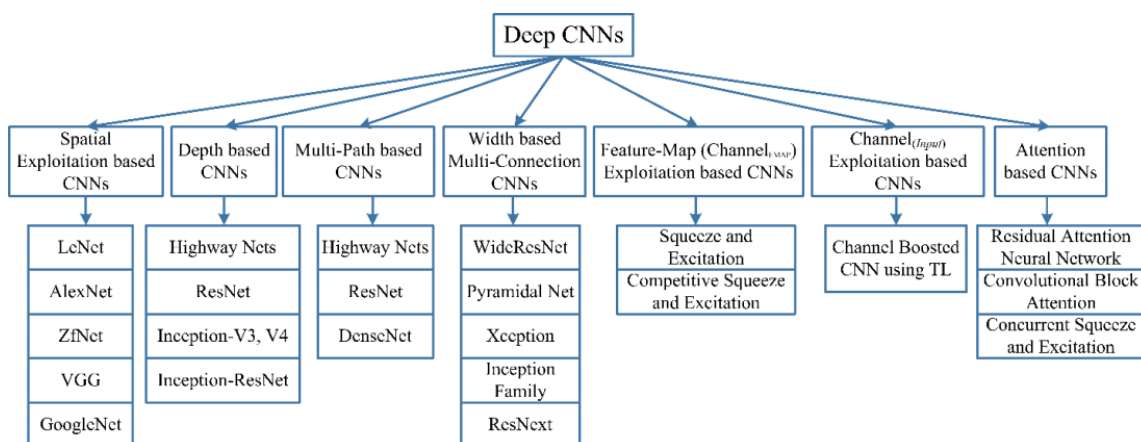


Fig. 4 Taxonomy of deep CNN architectures showing seven different categories.

Příloha 3: Taxonomy DCNN. Zdroj: <https://arxiv.org/ftp/arxiv/papers/1901/1901.06032.pdf>



Příloha 4: Snímky metličky křivolaké (vlevo polovina června) (vpravo srpen)

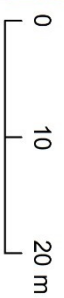
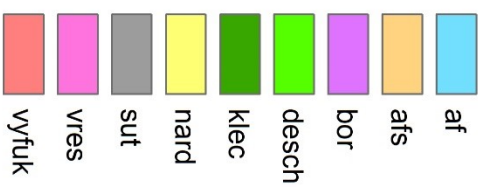
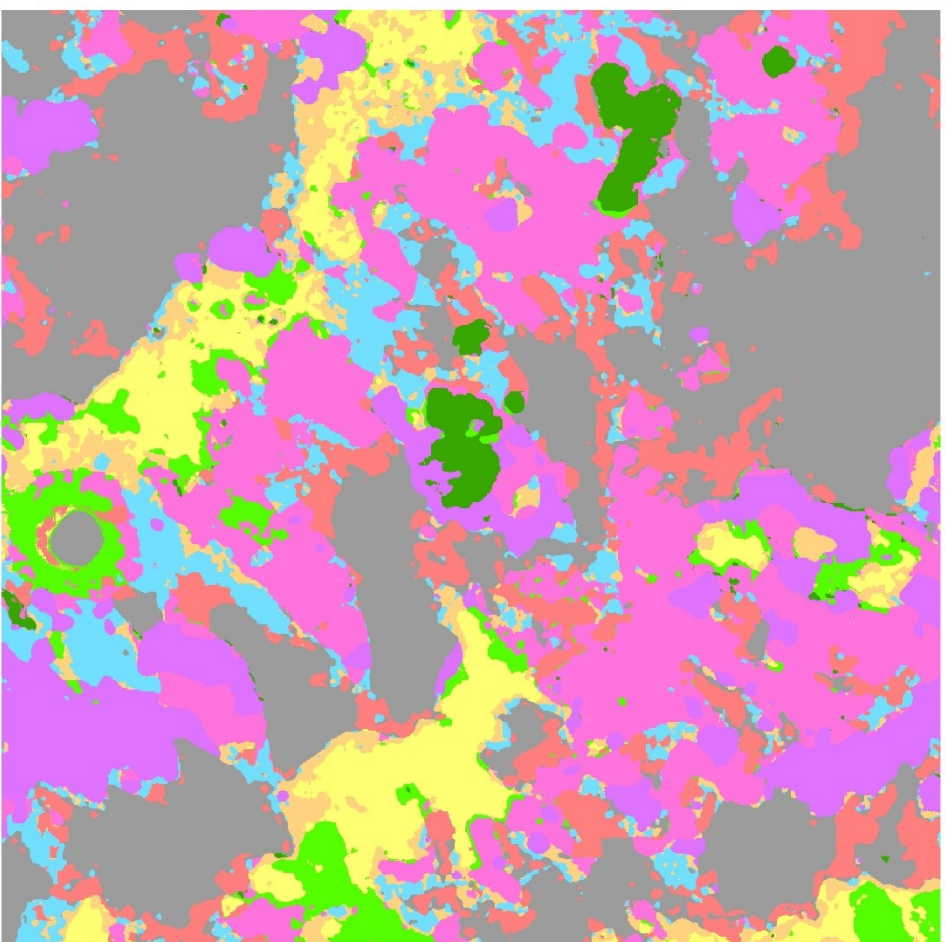


Příloha 5: Snímky smilky tuhé (vlevo druhá polovina července) (vpravo konec srpna)



Příloha 6: Snímky metlice trsnaté (vlevo konec června) (vpravo začátek srpna)

KLASIFIKACE HS SNÍMKŮ KRKONOŠSKÉ TUNDRY METODOU U-NET, SRPEN 2020

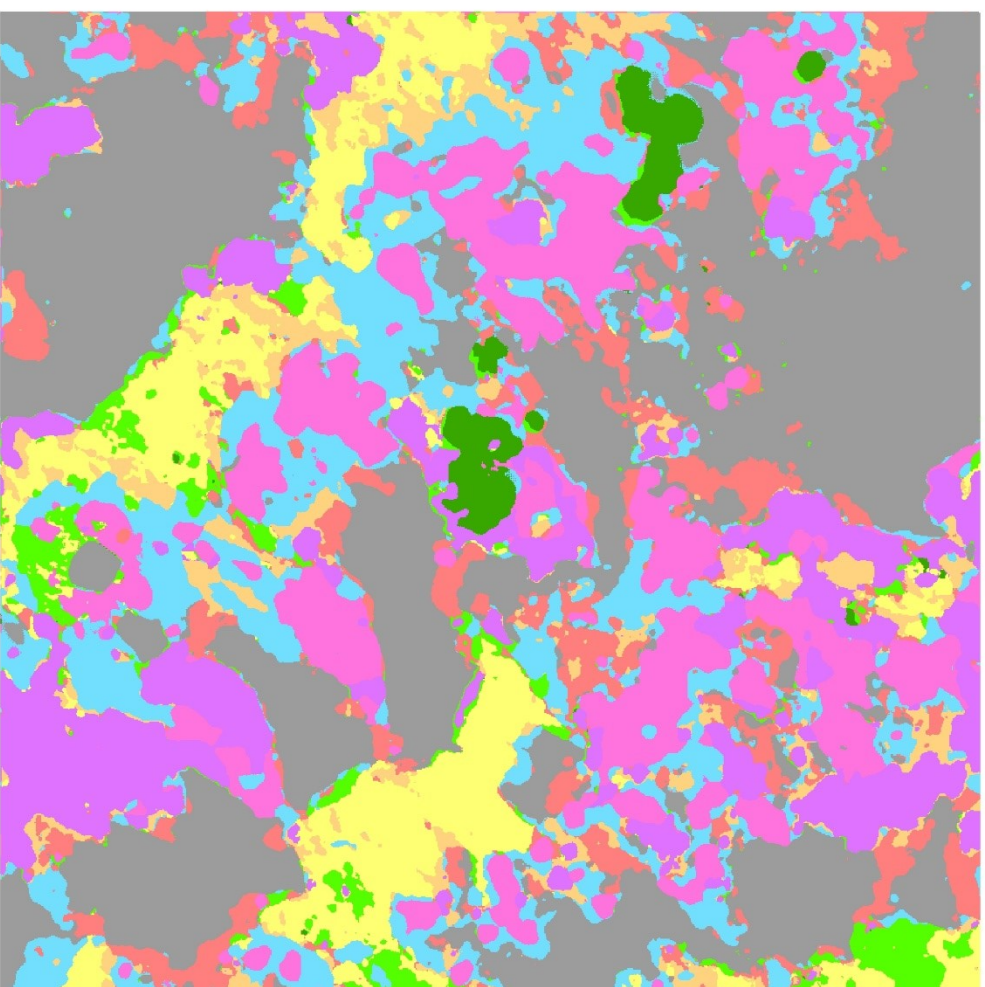


RÁDOVÁ, Martina
Praha, 2021



Příloha 7: Výsledný snímek klasifikace za období srpen

KLASIFIKACE HS SNÍMKŮ KRKONOŠSKÉ TUNDRY METODOU U-NET, ZÁŘÍ 2020

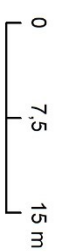
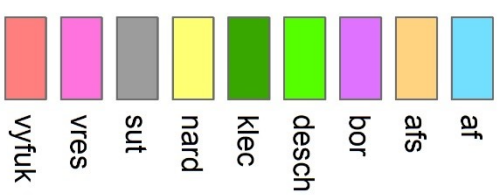
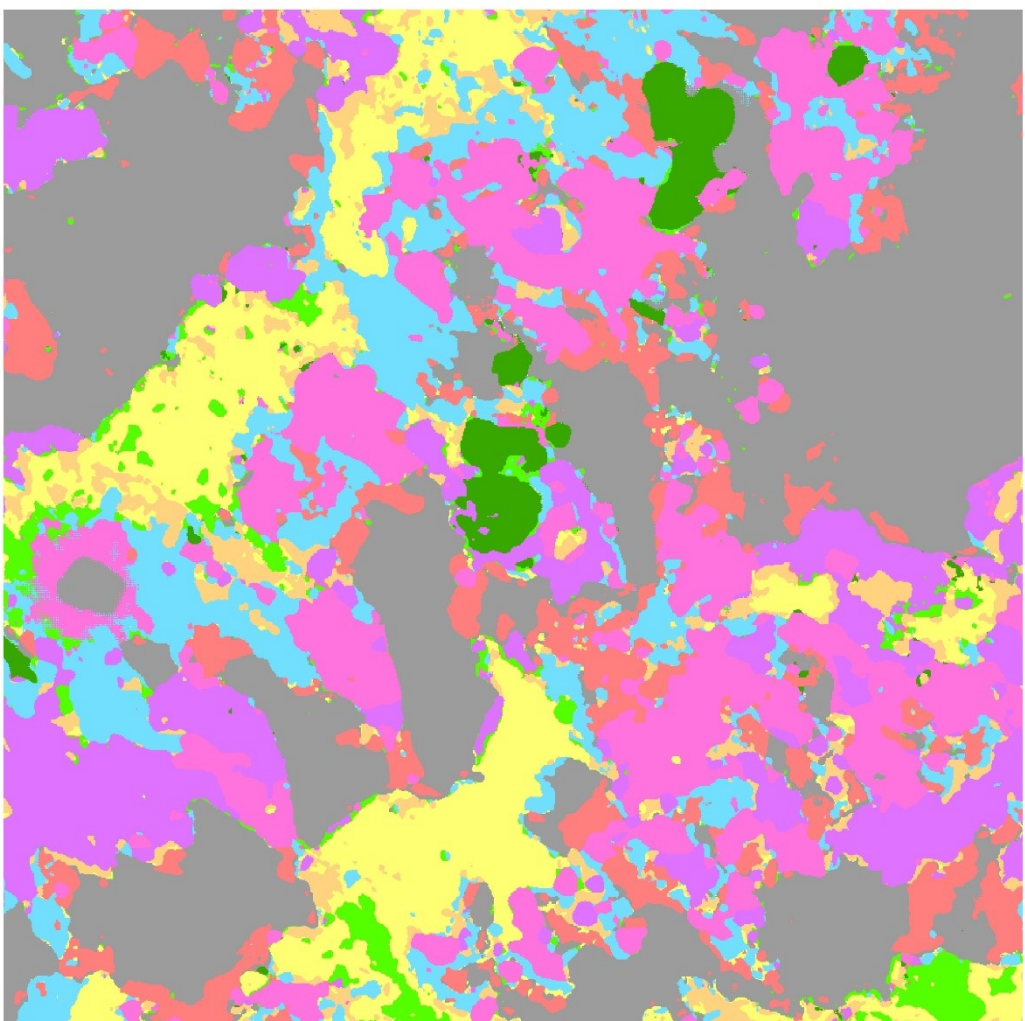


ŘÁDOVÁ, Martina
Praha, 2021



Příloha 8: Výsledný snímek klasifikace za období září

KLASIFIKACE HS SNÍMKŮ KRKONOŠSKÉ TUNDRY METODOU U-NET MULTITEMPORÁLNÍ KOMPOZIT



ŘÁDOVÁ, Martina
Praha, 2021



Příloha 9: Výsledný snímek klasifikace multitemporální kompozit