

Odhad nadmořské výšky z obrazu

Jan Vašíček*

Abstrakt

Tato práce se zabývá automatickým odhadem nadmořské výšky kamery z obrazu. Úlohu jsem řešil pomocí konvolučních neuronových sítí, u nichž využívám schopnost učit se nové příznaky na základě trénovacích dat. Trénovací sada obrazů (dataset), která by obsahovala údaje o nadmořské výšce kamery, nebyla k dispozici, a proto bylo nutné vytvořit dataset nový. Schopnosti člověka v dané úloze také nebyly dříve testovány, proto jsem provedl uživatelský experiment s cílem změřit průměrnou kvalitu lidského odhadu nadmořské výšky kamery. Experimentu se zúčastnilo 100 lidí a výsledky ukazují, že průměrná chyba odhadu člověka je 879 m. Automatický systém založený na konvoluční neuronové síti dosahuje lepších výsledků než člověk, neboť průměrná chyba odhadu se pohybuje okolo hodnoty 751 m. Navržený systém může kromě samotného odhadu nadmořské výšky z obrazových dat nalézt uplatnění také ve složitějších úlohách, jako je vizuální geo-lokalizace kamery.

Klíčová slova: Odhad nadmořské výšky — Konvoluční síť — EXIF data

Přiložené materiály: N/A

*xvasic21@stud.fit.vutbr.cz, Faculty of Information Technology, Brno University of Technology

1. Úvod

Lidský mozek dokáže porozumět okolnímu světu v řádu několika stovek milisekund. Během jediného pohledu je schopný z obrazové informace určit kategorii sledovaného objektu/scény [1]. Jak je tomu, ale při odhadu nadmořské výšky z obrazu? Zamysleme se nad fotografiemi na obrázku 1. V jaké nadmořské výšce byly tyto snímky pořízeny? Nadmořská výška kamery se běžně měří přesnou nivelací k střední hladině nejbližšího moře. Dalším způsobem jak změřit nadmořskou výšku je porovnání GPS souřadnic daného bodu s matematicky vypočteným elipsoidem WGS 84 [2]. Často však kromě obrazu nejsou k dispozici žádné doplňující informace. U videa to platí dvojnásob. Pokud bychom byli schopní určit nadmořskou výšku kamery z obrazu, tak to může pomoci při hledání ztracených osob, v automatizovaných meteorologických systémech, geo-lokalizaci obrazu (redukce vyhledávacího prostoru), apod. [3]. V tomto článku se zabývám automatickým odhadem nadmořské výšky kamery pomocí metod strojového učení. K tomuto účelu jsem vytvořil nový dataset a provedl uživatelský test, abych

zjistil, jak si navržené experimenty vedou v porovnání s člověkem.

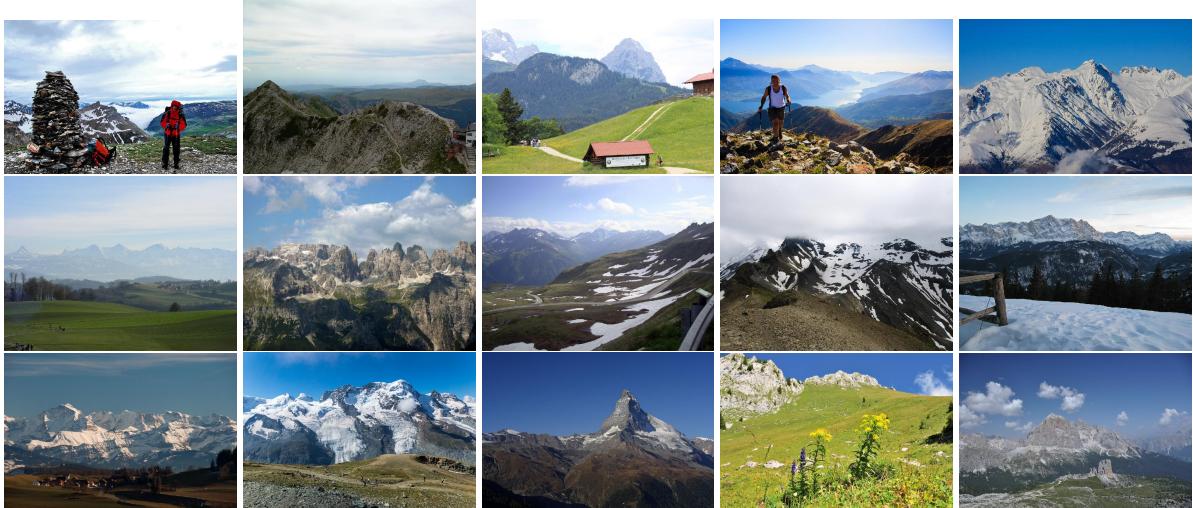
2. Předchozí práce

V současnosti bohužel neexistují dostupné materiály o výzkumech týkajících se odhadu nadmořské výšky kamery z obrazu. V této kapitole jsou uvedeny příklady projektů, které řeší podobnou úlohu.

2.1 IM2GPS

Práce [3] představuje algoritmus určený ke geo-lokalizaci obrazu. Výsledná lokace obrázku je reprezentována jako rozdělení pravděpodobnosti výskytu na zemském povrchu. Pokud je na obrázku známá kulturní památka, kterou systém rozpozná, tak vrací jednu konkrétní pozici. Pokud je na obrázku spíše obecná scéna jako poušt, tak systém vrací vysoké hodnoty pro suchá, písčitá místa.

V této práci je vytvořen dataset, který obsahuje GPS souřadnice ze kterých je možné určit nadmořskou výšku. Problém tohoto datasetu je nedostatek fotek



Obrázek 1. Ukázka obrázků z datasetu.

zachycujících venkovní scény v jednotlivých nadmořských výškách a příliš velký rozsah zeměpisných lokalit.

2.2 Places205

Places 205 [4] je dataset sloužící k trénování klasifikátorů scén. Obsahuje téměř 2,5 milionu obrázků z 205 kategorií. Počet obrázků v jednotlivých kategoriích není stejný a pohybuje se v rozmezí 5 000 - 15 000. Z tohoto datasetu by bylo možné vytvořit vhodný subset pro řešenou úlohu. Dataset Places205 však neobsahuje dodatečné informace k fotkám (GPS souřadnice či nadmořskou výšku). Pro tento výzkum jsem proto vytvořil nový dataset obsahující nadmořskou výšku ke každé fotce (Dále jen *Alpine altitudes*).

3. Dataset Alpine altitudes

Přírodní ráz krajiny se s nadmořskou výškou mění. Nadmořská výška například ovlivňuje teplotu, intenzitu slunečního svitu, či množství srážek. Čím vyšší nadmořská výška, tím nepříznivější bývají klimatické podmínky. Zimy ve vyšších polohách bývají delší, sníh taje později než v nižinách a nástup jara bývá opožděn [5].

Další element, který přispívá k přírodní variabilitě je zeměpisná poloha daného místa. Charakter dvou míst se stejnou nadmořskou výškou ale v různých zeměpisných lokalitách se bude velmi pravděpodobně lišit. Pro tvorbu datasetu jsem tedy vybral pouze oblast pohoří Alp. Tato oblast se dá považovat za dostatečně malou a můžeme říct, že krajinný charakter v jednotlivých nadmořských hladinách bude velmi podobný.

3.1 Vytvoření datasetu

Seznam hor byl vytvořen pomocí metadat ve službě OpenStreetMap [6]. Tento seznam byl následně použit

pro dotazování obrázků na Flickr¹ pomocí hash tagů. Staženy byly veřejné obrázky, které obsahovaly GPS souřadnice. Tyto obrázky nesměly obsahovat hash tagy, které nesouvisejí s venkovními scénami (svatba, oslava, atd). Po stažení byly fotky ještě automaticky upraveny tak, aby neobsahovaly jednobarevné okraje. Takto vytvořená kolekce obsahovala téměř 1,2M fotek. Tento postup bohužel vedl také ke stažení velkého počtu fotek nepoužitelných pro úlohu odhadu nadmořských výšek.

Pro filtraci nevhodných obrázků je možné použít klasifikátor scén, který obrázky z kolekce klasifikuje do různých kategorií. V této práci jsem použil konvoluční sít Places-CNN (viz. kapitola 5.1.1). Výstupem sítě jsou pravděpodobnosti pro jednotlivé kategorie scén². Toho jsem využil při sestavování algoritmu pro filtraci fotek. Vybral jsme kategorie, které souvisí s horami, krajinou a venkovními scénami. Z datasetu byly vyřazeny fotografie, pro které součet pravděpodobností relevantních kategorií byl menší než experimentálně určený práh 0,50. Pomocí klasifikace scén bylo odstraněno 75% fotek (rozložení skóre je zobrazeno v grafu na obrázku 2).

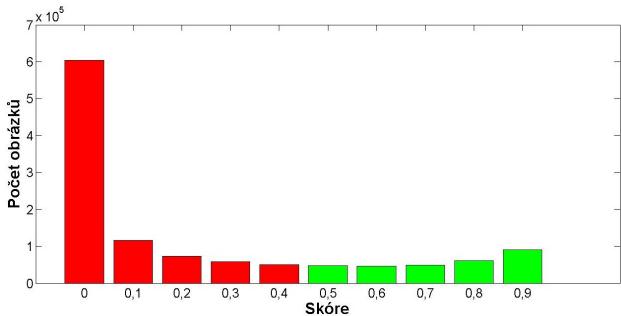
Z datové sady byly také odstraněny duplicitní fotky a fotky jejichž místo pořízení leží mimo Alpy. Místo pořízení bylo určeno pomocí GPS souřadnic uložených v EXIF datech. Výsledný dataset obsahuje 93 733 fotek.

Nadmořská výška fotek se získala pomocí GPS souřadnic z výškových map³ [7].

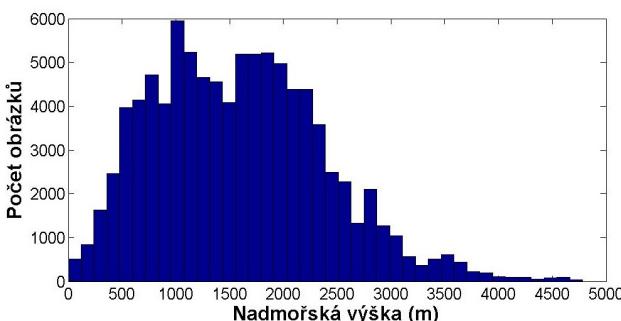
¹www.flickr.com

²lesní cesta, písčité pobřeží, pole, zasněžená hora, moře, jezero, bažina, lyžařské středisko, louka, horská chata, sopka

³Přesnost výškových map se pohybuje kolem hodnoty 30 m.



Obrázek 2. Rozložení skóre fotek. Skóre je získáno z klasifikátoru scén jako suma pravděpodobností pro relevantní kategorie.



Obrázek 3. Rozložení nadmořských výšek v datasetu.



Obrázek 5. Uživatelské rozhraní testu. Vpravo je zobrazen posuvník, kterým uživatel zadával svůj odhad nadmořské výšky.

výšku na vytvořeném datasetu. Kromě toho, že získané výsledky poskytují nové a zajímavé poznatky o schopnostech člověka, slouží také jako reference pro automatické metody, které jsou popsány v následující kapitole.

4.1 Parametry testování

Testování proběhlo pomocí webové aplikace (viz obr. 5). Experimentu se zúčastnilo 100 lidí, kteří odhadovali nadmořskou výšku kamery pro 50 obrázků . Věk účastníků je v rozmezí od 11 do 61. Obrázky byly náhodně vybrány z testovací sady datasetu *Alpine altitudes* tak, aby rovnoměrně pokrývaly rozsah nadmořských výšek od 79 m do 4463 m. (viz obr. 6).

Před spuštěním testu byl každý uživatel poučen o účelu testu a správném způsobu jak test vyplnit (např. uživatel určuje nadmořskou výšku kamery a ne vzdáleného objektu na fotce). Samotný test pak probíhal tak, že se uživateli v náhodném pořadí postupně zobrazovaly obrázky z testovací kolekce. Kromě obrázku nebyla k dispozici žádná doplňující informace. Účastník určoval nadmořskou výšku kamery pomocí posuvníku (viz obr. 5). Tento odhad pak uživatel potvrdil tlačítkem po jehož zmáčknutí se zobrazil další obrázek. Průměrná doba trvání testu byla 10 minut.

3.2 Charakteristiky datasetu

Dataset obsahuje 93 733 obrázků s nadmořskými výškami v rozmezí od 79 m do 4463 m. Histogram nadmořských výšek je zobrazen na obrázku 3. Rozložení grafu není uniformní a odpovídá tomu jak často lidé fotí v různých nadmořských výškách. Množství fotek s nadmořskou výškou větší než 3000 m je malé, což omezuje použitelnost datové sady pro nejvyšší místa Alp. Kromě francouzské části je pokryta většina území Alp (viz obr. 4). Důvodem je nedostupnost fotek z Francie v době vytváření datasetu. Dataset také obsahuje kompletní EXIF data dostupná z Flickru.

4. Odhad nadmořské výšky kamery člověkem

Tato kapitola prezentuje uživatelský test, jehož cílem bylo kvantifikovat schopnost lidí určovat nadmořskou

4.2 Vyhodnocení uživatelského testu

Výsledek uživatelského testu jsem nejdříve analyzoval funkcí ANOVA (analýza rozptylu) [8] abych ověřil, že odhady nadmořské výšky kamery nejsou náhodné a člověk je ji schopný určovat na základě fotografie. K tomuto účelu jsem vytvořil nulovou hypotézu H_0 : *Při odhadu člověka nezáleží na obrázku.* Výsledek analýzy rozptylu je ($F = 165.09^4$, $p < 0.001$), což zna-

⁴F je testová statistika. Vyjadruje poměr mezi rozptylem "mezi" skupinami a rozptylem "uvnitř" skupin. Hodnota testové statistiky výrazně vyšší než 1 umožňuje hypotézu zamítnout.

mená, že člověk je schopný odhadovat nadmořskou výšku kamery pouze na základě obrazové informace.

V tuto chvíli již můžeme zkoumat charakteristiku odhadu člověka. Na obrázku 6 je zobrazena střední hodnota a rozptyl odhadu všech účastníků testu pro každý obrázek. Nadmořské výšky do 1000 m je člověk schopný odhadnout s malou chybou. S rostoucí nadmořskou výškou však chyba roste. Nejvíce je to vidět u nadmořských výšek nad 3000 m, kde člověk odhadu velmi podhodnocuje. Jedním z důvodů může být to, že většina účastníků neměla zkušenosti s pobytom v těchto výškách. Průměrná chyba odhadu člověka je 879,94 m.

5. Automatický odhad nadmořské výšky

Velký vliv na kvalitu výsledků metod strojového učení mají použité příznaky. V tradičním přístupu byly vhodné příznaky navrhovány na základě znalostí a zkušeností odborníků. Pro obecné úlohy klasifikace a vyhledávání fotografií jsou to například GIST [9] a SIFT (Scale-invariant feature transform) [10], které zachycují globální respektive lokální vzhled obrazu. Výběr vhodných příznaků pro konkrétní úlohu je dlouhý a náročný proces, zvláště pokud nejsou dostupné prameny o dřívější práci, tak jako v tomto případě.

Alternativou k ručně navrženým příznakům, která v posledních letech nabírá na popularitě, je hluboké učení, které využívá rostoucího výkonu výpočetní techniky a velkých datových sad k naučení vhodných příznaků přímo z dat.

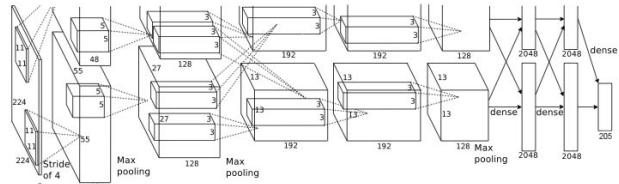
Jednou z úspěšných metod hlubokého učení pro obrazová data jsou hluboké konvoluční sítě, které dosahují v současnosti nejlepších výsledků na úlohách klasifikace obrazu [11], rozpoznávání lidí podle obličeje [12] a detekce objektů [13].

Obecnost konvolučních sítí je vykoupena většími nároky na množství trénovacích dat a velkou výpočetní náročností jejich trénování.

Z důvodu nedostatku dostupných materiálů o dřívějších pracích na odhadu nadmořské výšky kamery z obrazu jsem použil právě konvoluční neuronové sítě. Návrh architektury a výsledky experimentů jsou popsány v následujících kapitolách.

5.1 Architektura sítě

Odhad nadmořské výšky je svým charakterem a složitostí srovnatelný s obecným problémem klasifikace fotografií [11]. Sítě používané v těchto úlohách mají milióny parametrů (viz obr. 7) a vyžadují statisíce trénovacích obrázků. Dataset *Alpine altitudes* není dosažitelně velký pro trénování sítě od začátku s náhodně inicializovanými parametry, a proto jsem zvolil postup



Obrázek 7. Architektura konvoluční sítě Places-CNN.

adaptace příznaků z existující sítě.

V této práci jsem vytvořil 3 regresní modely a zkoumal jejich chování na vytvořeném datasetu. Při návrhu všech modelů sítě jsem vycházel z modelu sítě natrénovaného na datasetu Places205 [4], který je popsán v následující kapitole.

5.1.1 Places-CNN

Places-CNN [4] je konvoluční síť, která v současnosti dosahuje nejlepších výsledků v úloze klasifikace indoor/outdoor scén. Síť je natrénovaná na přibližně 2,5M obrázcích z datasetu Places205. Architektura sítě je shodná s architekturou použitou v Caffe reference network [11]. Vstupem sítě jsou barevné 2D obrázky zmenšené na velikost 256x256 pixelů. Síť je tvořena 8 vrstvami (viz obr. 7). Prvních 5 vrstev jsou vrstvy konvoluční. Zbylé 3 vrstvy jsou plně propojené. Hlavní rozdíl konvolučních a plně propojených vrstev je ten, že plně propojené vrstvy nemají sdílené váhy. Výstup poslední plně propojené vrstvy je vstupem softmax funkce, která produkuje distribuci nad 205 třídami.

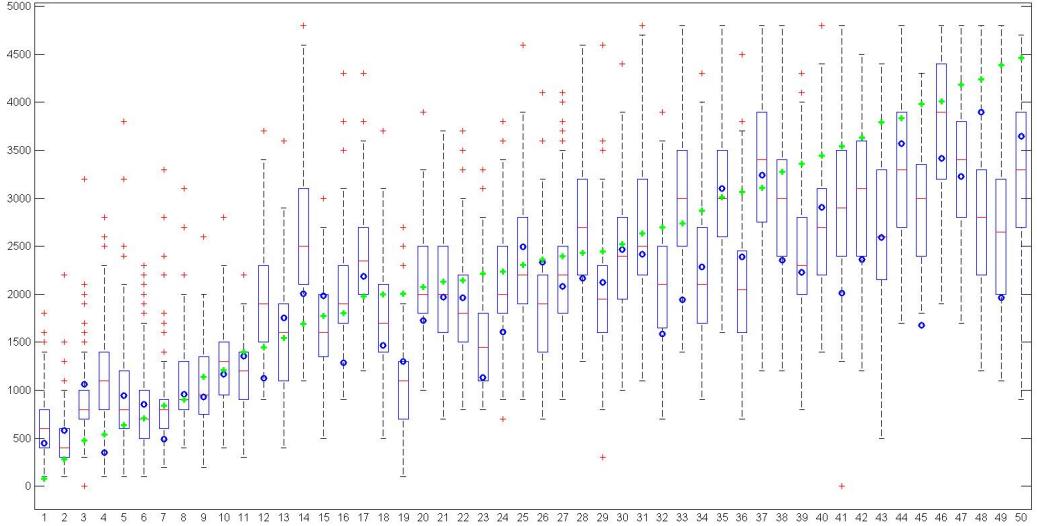
5.1.2 Příznaky Places

Předešlé práce [12] [14] ukazují, že příznaky extrahované z existujících sítí jsou dostatečně obecné a dají informaci i pro podobné úkoly, na které nejsou původně určené. V tomto experimentu jsem zjišťoval jak zafungují příznaky sítě Places-CNN na datasetu *Alpine altitudes*.

Extrahal jsem aktivace první plně propojené vrstvy (4096 neuronů) pro všechny obrázky datasetu. Extrahané aktivace reprezentují globální deskriptor obrazu a slouží jako vstup pro 2-vrstvou neuronovou síť, jejíž parametry jsou na začátku trénování náhodně inicializovány. Výstupní vrstva této sítě je lineární s jedním neuronem, který přímo odhaduje nadmořskou výšku. Síť byla trénována pomocí optimalizační kritéria *mean square error*, které je vhodné pro regresní problémy, kde lze předpokládat normální rozložení chyb.

5.1.3 Adaptované příznaky Places

Lepších výsledků než přímým použitím existující sítě pro extrakci příznaků lze většinou dosáhnout adaptací existující sítě pro konkrétní úlohu — tak zvaným Fine-



Obrázek 6. Odhady člověka a konvoluční sítě na 50 obrázcích z uživatelského testu. Na ose y je vynesena nadmořská výška. Každý "box" odpovídá jednomu obrázku a zobrazuje rozložení odhadů lidí. (zelené křížky - správná hodnota, modrá kolečka - Adaptované příznaky Places, červená čára - střední hodnota odhadu účastníků testu pro daný obrázek).

tunningem [15] [13]. Toho je dosaženo inicializací vah sítě z již existující sítě kromě posledních vrstev, které jsou inicializovány náhodně. Taková síť je pak učena celá pomocí metody Stochastic Gradient Descent. Tento postup jsem použil při trénování druhého modelu (*Adaptované příznaky Places*). Architektura sítě je shodná s architekturou sítě Places-CNN. Parametry jsou inicializovány pomocí parametrů z modelu Places-CNN.

5.1.4 Adaptované příznaky Places + EXIF

Téměř polovina datasetu obsahuje EXIF data, která mohou být použita jako dodatečné informace o obrázku a upřesnit tak odhad nadmořské výšky. Pro tento experiment jsem vytvořil neuronovou síť jejímž vstupem byly aktivace první plně propojené vrstvy z modelu *Adaptované příznaky Places* a zakódovaná EXIF data. Zde uvádím výčet použitých EXIF tagů a způsob jejich kódování:

1. **Čas** - Pozice slunce se během dne mění a to se na snímcích může projevit změnou stínů nebo různou intenzitou světla. Denní dobu kóduji jako 1 z 16 s uniformní velikostí binů.
2. **Datum** - Datum pořízení snímku nás informuje o specifických povětrnostních podmírkách v dané části roku a délce slunečního svitu během dne. Pozici dne během roku kóduji jako 1 z 16 kde biny mají uniformní velikost.
3. **Nastavení fotoaparátu** - Informace o nastavení fotoaparátu (ohnisková vzdálenost, expoziční

čas, ISO) jsem použil k určení množství světla ve scéně v době pořízení snímku. Množství světla ve scéně EV (expoziční hodnota) je vypočítáno jako

$$EV = 2\log_2(F) - \log_2(t) - \log_2\left(\frac{ISO}{100}\right), \quad (1)$$

kde F je clonové číslo, t je expoziční čas a ISO je citlivost snímacího senzoru.

Kromě samotné expoziční hodnoty jsou použity i jednotlivé parametry nastavení. Předpokladem bylo, že i jednotlivé parametry mohou nést informaci důležitou pro odhad nadmořské výšky. Expoziční hodnota je zakódovaná jako 1 z 16 s použitím uniformních binů. Pro kódování ostatních parametrů je použito kódování 1 z N kde velikost binů exponenciálně roste⁵.

4. **Zorné pole** - určuje část scény, která se zobrazí na snímek. To pomáhá při odhadu projekce scény na senzor fotoaparátu. Informaci o zorném poli obsahuje pouze velmi malá část datasetu. Tento údaj jsem musel dopočítat z ostatních metadat pomocí rovnice

$$FOV = 2\text{atan}\left(0.5\frac{SEN}{f}\right)\frac{180}{\pi}, \quad (2)$$

kde FOV je zorné pole, SEN je rozměr senzoru a f je ohnisková vzdálenost.

⁵Stupnice těchto parametrů je logaritmická.

Tabulka 1. Porovnání výsledků všech variant modelů.

Model	RMSE(m)
Příznaky Places	569.63
Ad. příznaky Places	549.82
Ad. příznaky Places + EXIF	510.79

V současnosti neexistuje freeware program s databází obsahující velikost senzorů pro všechny modely fotoaparátu. Velikost senzoru jednotlivých modelů jsem proto manuálně hledal na stránkách výrobců. Statisticky jsem zjistil, že při vytvoření databáze pro 100 modelů je pokryto 40% datasetu. Zorné pole jsem kódoval jako 1 z 16.

5.2 Vyhodnocení modelů

Pro korektní porovnání jednotlivých variant modelů sítí jsem vytvořil subset datasetu *Alpine altitudes* (dále jen *EXIF dataset*), který pro každou fotku obsahuje EXIF data použitá při tvorbě varianty *Adaptované příznaky Places + EXIF*. Trénovací sada obsahuje 34 889 obrázků. Testovací sada má 5046 obrázků. *EXIF dataset* jsem následně použil pro přetrenování modelů *Příznaky Places* a *Adaptované příznaky Places*. Všechny modely tak lze přímo porovnávat, protože byly trénovány/testovány na stejných datech. V následující části popisují chování modelů na testovací sadě *EXIF datasetu*.

5.2.1 Příznaky Places

Příznaky extrahované z konvoluční sítě Places-CNN se ukázaly jako dostatečně obecné. Průměrná chyba odhadu na celé testovací sadě (viz tabulka 1) se málo liší od průměrné chyby modelu *Adaptované příznaky Places* (v uživatelském testu dosahuje lepších výsledků než člověk). Dobré výsledky na této úloze přisuzuji faktu, že příznaky sítě Places-CNN byly trénovány na vysokém počtu různorodých venkovních scén.

5.2.2 Adaptované příznaky Places

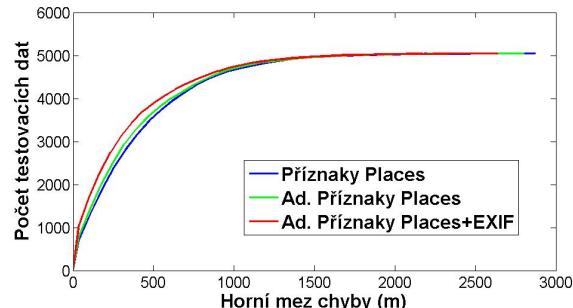
Adaptace konvoluční sítě na *EXIF dataset* vede k mírnému zlepšení výsledků regrese (tabulka 1). Velikost *EXIF datasetu* je malá a přínos metody fine-tuning se v tomto experimentu výrazně neprojevil.

5.2.3 Adaptované příznaky Places + EXIF

Předpoklad, že dodatečné informace k obrázku pomohou při odhadu nadmořské výšky kamery se potvrdil. Konvoluční síť byla schopna využít větší množství informací k zpřesnění odhadu a zmenšení chyby (viz obr. 8, tabulka 1). Provedený experiment zároveň ukazuje vhodnost použitých EXIF dat pro podobné úlohy.

Tabulka 2. Průměrná chyba odhadů člověka a konvoluční sítě na obrázcích z uživatelského testu.

	RMSE(m)
Člověk	879.94
Ad. příznaky Places	751.12

**Obrázek 8.** Graf zobrazující počet testovacích dat (osa y) s chybou do určité hodnoty (osa x).

5.3 Porovnání výkonu člověka a konvolučních sítí

Na základě provedeného uživatelského testu je možné konstatovat, že konvoluční sítě dosahují v této úloze lepších výsledků než člověk (tabulka 2). Jelikož účastníci testu měli k dispozici pouze obrazovou informaci, tak je porovnání provedeno s variantou *Adaptované příznaky Places*⁶. Do grafu na obrázku 6 jsou zaneseny odhady lidí i konvoluční sítě pro obrázky použité v uživatelském testu. Konvoluční sítě mají stejně jako lidé největší problémy u obrázků pořízených ve vyšších nadmořských výškách. V této práci to může být způsobeno malým množstvím fotografií z těchto výšek v datasetu.

6. Závěr

V této práci je ukázáno, že konvoluční neuronové sítě dosahují na úloze odhadu nadmořské výšky kamery z obrazu lepších výsledků než člověk. Odhady nadmořské výšky jsou pouze přibližné a průměrná chyba se pohybuje kolem hodnoty 550 m. Tuto chybu je možné snížit pokud se použijí dodatečná data o obrázku (EXIF data). Na experimentech je dobře prezentována potřeba vhodného datasetu pro řešenou úlohu. Dataset vytvořený v této práci je unikátní sada obrazů z horského prostředí doplněná informací o nadmořské výšce kamery. Tento dataset sice není tak obsáhlý jako obecné datasety používané při trénování konvolučních sítí, ovšem i na něm bylo možné úspěšně provést experimenty, které do budoucna dávají naději na vytvoření

⁶Výrazný rozdíl ve velikosti chyby v tabulkách 1 a 2 je způsoben odlišným rozložením nadmořských výšek v použitých testovacích sadách.

řešení, které by se dalo použít v praxi. Věřím, že při trénování konvoluční sítě na dostatečně velkém datasetu, který bude splňovat požadavky na hustotu a rozmanitost [4], je možné dosáhnout ještě přesnějších odhadů nadmořské výšky kamery z obrazu.

Poděkování

Chtěl bych poděkovat vedoucímu mé diplomové práce Ing. Martinu Čadíkovi, Ph.D. za pomoc a rady při zpracování této práce. Dále bych rád poděkoval Ing. Michalu Hradišovi za odborné rady během návrhu konvolučních sítí. Výzkum vedoucí k těmto výsledkům vznikl v rámci projektu LOCATE 4SGA8694, který je financován z programu SoMoPro II, spolufinancovaného Evropskou unií a Jihomoravským krajem.

Literatura

- [1] Talia Konkle, Timothy F Brady, George A Alvarez, and Aude Oliva. Scene memory is more detailed than you think: the role of categories in visual long-term memory. *Psychological science : a journal of the American Psychological Society / APS*, 21(11):1551–1556, 2010.
- [2] Václav Čada. Úvod do Geodézie.
- [3] James Hays and Alexei A Efros. IM2GPS: estimating geographic information from a single image. In *Proceedings of the {IEEE} Conf. on Computer Vision and Pattern Recognition ({CVPR})*, 2008.
- [4] B Zhou, A Lapedriza, J Xiao, A Torralba, and A Oliva. Learning Deep Features for Scene Recognition using Places Database. *NIPS*, 2014.
- [5] William Burroughs. *Climate: Into the 21st Century*. Cambridge University Press, 2003.
- [6] Kevin Curran, John Crumlish, and Gavin Fisher. OpenStreetMap. *International Journal of Interactive Communication Systems and Technologies*, 2(1):69–78, 2012.
- [7] Lionel Baboud, Martin Čadík, Elmar Eismann, and Hans-Peter Seidel. Automatic Photo-to-terrain Alignment for the Annotation of Mountain Pictures. In *Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition, CVPR '11*, pages 41–48, Washington, DC, USA, 2011. IEEE Computer Society.
- [8] Viv Bewick, Liz Cheek, and Jonathan Ball. Statistics review 9: one-way analysis of variance. *Critical care (London, England)*, 8(2):130–136, 2004.
- [9] Aude Oliva and Antonio Torralba. Modeling the shape of the scene: A holistic representation of the spatial envelope. *International Journal of Computer Vision*, 42(3):145–175, 2001.
- [10] David G Lowe. Distinctive Image Features from Scale-Invariant Keypoints. *Int. J. Comput. Vision*, 60(2):91–110, 2004.
- [11] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. ImageNet Classification with Deep Convolutional Neural Networks. *Advances In Neural Information Processing Systems*, pages 1–9, 2012.
- [12] Yaniv Taigman, Ming Yang, Marc’Aurelio Ranzato, and Lior Wolf. DeepFace: Closing the Gap to Human-Level Performance in Face Verification. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, page 8, 2014.
- [13] Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In *Computer Vision and Pattern Recognition*, 2014.
- [14] Jeff Donahue, Yangqing Jia, Oriol Vinyals, Judy Hoffman, Ning Zhang, Eric Tzeng, and Trevor Darrell. DeCAF: {A} Deep Convolutional Activation Feature for Generic Visual Recognition. *CoRR*, abs/1310.1, 2013.
- [15] Pulkit Agrawal, Ross B Girshick, and Jitendra Malik. Analyzing the Performance of Multilayer Neural Networks for Object Recognition. *CoRR*, abs/1407.1, 2014.