

1 Dataset - Kinetics400

Pro základní otestování metod *action recognition* byl zvolen *Kinetics 400* [3] dataset. *Kinetics 400* byl vybrán, protože je často využíván pro testování metod rozpoznávání činností, kterými se zabývá i tato diplomová práce a bude pak snadné porovnat naše rozpoznávací systémy s jinými.

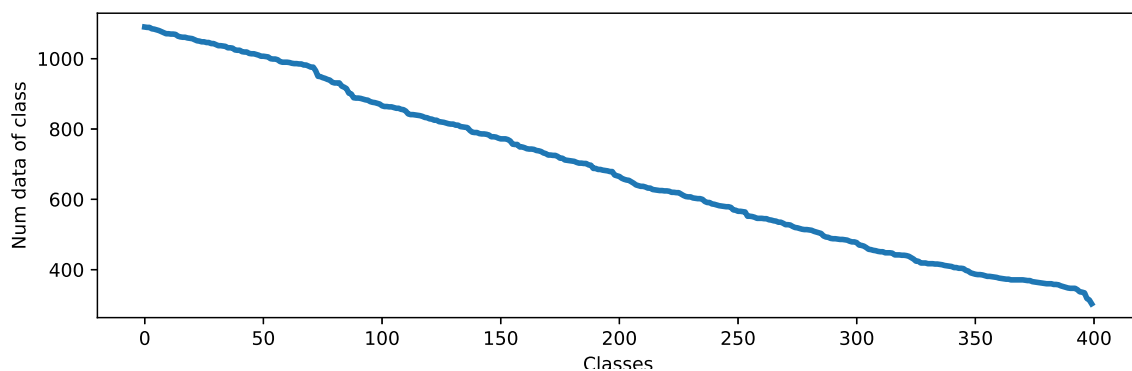
Kinetics 400 je kolekce videí s lidskou činností. Tyto videa jsou reprezentovány pomocí URL odkazy na videa internetového serveru *YouTube.com*. U každého URL odkazu je zaznamenaný počáteční a koncový čas označené činnosti, přičemž každé video trvá zhruba 10 vteřin. Činnosti jsou zaměřeny na člověka a pokrývají širokou škálu tříd včetně interakcí mezi člověkem a objektem, jako je hraní na hudební nástroje a také interakcí mezi lidmi, jako je podání ruky. V tomto datasetu je celkem 400 tříd. Každá třída obsahuje minimálně 400 videoklipů. Všechny tyto informace o URL odkazu, času ve kterém se vyskytuje činnost a korektně pojmenovaná činnost, jsou uloženy ve formátu *.json* nebo *.csv*.

Na základě dostupného datasetu *Kinetics 400* proběhla základní analýza dat. Celkový počet videí je zaznamenan v Tabulce 1. Stažený dataset byl už rozdělen do třech množin dat, trénovací, validační a testovací.

	Počet videí
Trénovací data	219 782
Validační data	18 035
Testovací data	35 357
Celkový počet dat	273 174

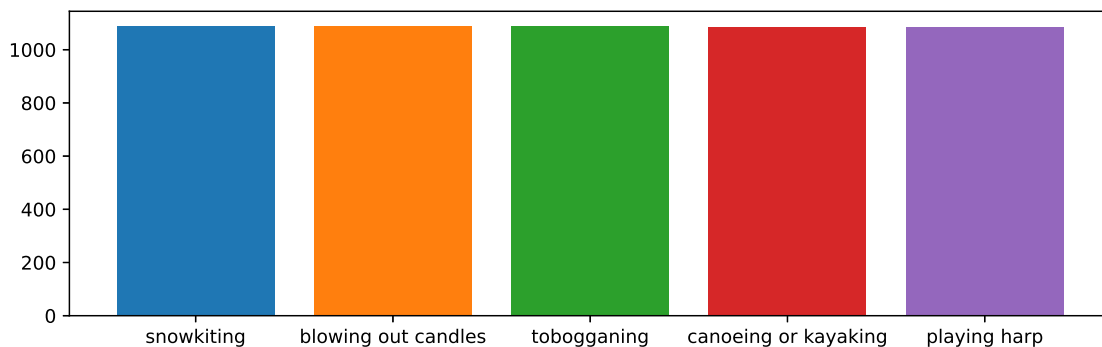
Tabulka 1: Počet dat

Při analýze datasetu je důležité zjistit jak jsou data v jednotlivých třídách rozprostřeny. Distribuce dat mezi třídami je zobrazena na Obrázku 1.



Obrázek 1: Distribuce tříd

Z Obrázku 1 je patrné, že zde není žádná třída, která by výrazně v počtu dat převyšovala ostatní třídy. Pro každou třídu je také dostatečné množství videí. Na Obrázku 2 je zobrazený graf s pěti nejčetnějšími třídami.



Obrázek 2: Ukázka nejčtetnějších tříd

Pro stažení *Kinetics 400* videí byl vytvořený script s využitím knihoven *youtube-dl* a *ffmpeg*. Tento způsob extrakce videí datasetu je ale příliš zdlouhavý. Při praktickém otestování rychlosti stahování *360p* videí bylo zjištěno, že stažení jednoho průměrného videa zabere $t_{mean} = 15.57s$. Z tohoto údaje vyplývá, že stažení celého datasetu o 273174 videí potrvá při průměrném stahování 49.2 dne. Podobných výsledků dosáhl také Mark Gituma(2021) ve svém článku *Downloading The Kinetics Dataset For Human Action Recognition in Deep Learning* [2]. V tomto článku popisuje stahování datasetu *Kinetics 700* je zde uvedeno, že bylo dosaženo průměrného stažení jednoho videa na hodnotu $t_{mean} = 23.52s$ a stažení celého *Kinetics 700* by trvala 176.1 dní.

Alternativně byl dataset získán s využitím *Academics Torrents* [1][4]. Tento projekt poskytuje ukládání, ale i stahování velkých data pro výzkum a vědecké účely. *Academics Torrents* je platforma využívající *peer-to-peer* nástroj *BitTorrent*, díky tomu je umožněné skupině členů takzvané "seedovat" a díky tomu sdílet data ostatním účastníkům.



Obrázek 3: Ukázka - snowkiting



Obrázek 4: Ukázka - tobogganing



Obrázek 5: Ukázka - blowing out candles



Obrázek 6: Ukázka - canoeing or kayaking

Odkazy

- [1] Joseph Paul Cohen a Henry Z. Lo. “Academic Torrents: A Community-Maintained Distributed Repository”. In: *Annual Conference of the Extreme Science and Engineering Discovery Environment*. 2014. DOI: 10.1145/2616498.2616528. URL: <http://doi.acm.org/10.1145/2616498.2616528>.
- [2] Mark Gituma. *Downloading the kinetics dataset for Human Action Recognition in deep learning*. 2021. URL: <https://towardsdatascience.com/downloading-the-kinetics-dataset-for-human-action-recognition-in-deep-learning-500c3d50f776>.
- [3] Will Kay et al. “The Kinetics Human Action Video Dataset”. In: *CoRR* abs/1705.06950 (2017). arXiv: 1705.06950. URL: <http://arxiv.org/abs/1705.06950>.
- [4] Henry Z. Lo a Joseph Paul Cohen. “Academic Torrents: Scalable Data Distribution”. In: *Neural Information Processing Systems Challenges in Machine Learning (CiML) workshop*. 2016. URL: <http://arxiv.org/abs/1603.04395>.