

Posudek diplomové práce

Matematicko-fyzikální fakulta Univerzity Karlovy

Autor práce Patrik Veselý
Název práce Similarity Models for Content-based Video Retrieval
Rok odevzdání 2023
Studijní program Informatika **Studijní obor** Softwarové a datové inženýrství

Autor posudku Ladislav Peška **Role** Vedoucí
Pracoviště KSI

Text posudku:

Zadání práce obsahovalo několik dílčích cílů, které dohromady měly umožnit tvorbu vylepšených modelů pro podobnostní vyhledávání ve videu. Student měl za úkol sestavit dataset zachycující vnímání podobnosti reálnými osobami, nad ním vyhodnotit existující metody pro podobnostní vyhledávání (povětšinou založené na feature extraction pomocí předtrénovaných konvolučních sítí) a následně pak prověřit možnosti fine-tuningu daných metod.

Všechny popsání úkoly byly studentem splněny na velmi vysoké úrovni. Student nejprve sestavil webovou aplikaci pro sběr uživatelských dat (porovnání který z dvojice kandidátů je podobnější zdrojovému obrázku) a provedl sběr s pomocí reálných osob. Sběr probíhal ve dvou fázích s mírně odlišným zadáním. V první fázi bylo získáno cca 4K porovnání od 38 účastníků, ve druhé pak 17K porovnání od 84 účastníků. Celkově se jedná o datasety dostatečně rozsáhlé pro splnění následných úkolů. Jako velmi zdařilé pak hodnotím zapojení gamifikačních prvků do aplikace pro sběr dat, ale i návrh algoritmu pro výběr testovaných trojic, nebo poměrně kreativní řešení problému čištění dat. Z hlediska budoucího výzkumu je poměrně zásadní i vytvoření snadno použitelného a rozšiřitelného frameworku pro feature extraction.

Co se týká analýzy výkonu jednotlivých feature extraktorů, tak práce přinesla potvrzení dominance DL-based metod nad jednoduššími color-based nebo texture based extraktory, nicméně autor přišel i s některými závažnějšími zjištěními, například že větší modely ze stejné třídy (např. EfficientNetB0-B7) ne vždy dosahují lepších výsledků. Zrovna tak i fakt že nejvíce konzistentní s "human judgment" je poměrně starý (z roku 2019) W2VV++ model je vcelku překvapivý. Hlavní výsledky práce založené na první fázi sběru dat již byly publikovány na konferenci MMM 2023 (CORE B), zatímco rozšiřující výsledky vzešlé z rozsáhlejšího datasetu v současné době připravujeme k publikaci. Práce nechala i několik zajímavých otázek pro future work, například problematiku polarizujících trojic a z ní vyplývající prostor pro personalizaci.

V poslední části práce se pak student zabýval možnostmi zlepšení konzistence feature extraktorů s human judgments pomocí fine-tuningu. Student vytvořil jednoduchý fine-tuning framework a ve výsledcích ukázal, že i u nejlepšího pre-trained extraktoru je prostor pro zlepšení pomocí fine-tuningu.

Celkově hodnotím práci jako ukázkový příklad vědecky zaměřené diplomové práce. V práci jsou zastoupeny všechny typické kroky základního výzkumu od sběru dat, tvorby ad-hoc softwarových řešení, analýzy dat, aplikace machine-learning algoritmů atd. Práce přinesla řadu vědecky zajímavých výsledků, z nichž některé již byly publikovány a další k publikaci připravujeme. Ze stylistického pohledu by bylo možné v některých případech volit jasnější či

výstižnější slovní obraty a použití britské angličtiny je v dnešní době již poněkud rušivé, nicméně i přes popsané drobné výhrady doporučuji práci k obhajobě.

Práci doporučuji k obhajobě.

Práci nenavrhuji na zvláštní ocenění.

Pokud práci navrhuje na zvláštní ocenění (cena děkana apod.), prosím uveďte zde stručné zdůvodnění (vzniklé publikace, významnost tématu, inovativnost práce apod.).

Datum

Podpis