

Posudek bakalářské práce

Matematicko-fyzikální fakulta Univerzity Karlovy

Autor práce Jonáš Havelka
Název práce Generating Music Symbols Using Neural Networks
Rok odevzdání 2024
Studijní program Informatika
Specializace [Click here to enter text.](#)

Autor posudku Jan Hajič
Pracoviště ÚFAL

Role Oponent

Prosím vyplňte hodnocení křížkem u každého kritéria. Hodnocení *OK* označuje práci, která kritérium vhodným způsobem splňuje. Hodnocení *lepší* a *horší* označují splnění nad a pod rámec obvyklý pro bakalářskou práci, hodnocení *nevyhovuje* označuje práci, která by neměla být obhájena. Hodnocení v případě potřeby doplňte komentářem. Komentář prosím doplňte všude, kde je hodnocení jiné než *OK*.

K celé práci	lepší	OK	horší	nevyhovuje
Obtížnost zadání	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Splnění zadání	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Rozsah práce ... <i>textová i implementační část, zohlednění náročnosti</i>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<p>Komentář Práce popisuje systém pro generování syntetických znaků ručně psané hudební notace pomocí hlubokého učení. Tvorba syntetických dat je pro obor automatického rozpoznávání notace klíčová, neboť tvorba anotovaných dat je velmi nákladná a pro modely hlubokého učení, které se pro samotné rozpoznávání používají, je třeba trénovacích dat velké množství. V rámci syntézy notace je syntéza jednotlivých symbolů jeden z potřebných kroků, který zapadá do syntézy celých stránek. (Předtím je třeba vygenerovat zapisovanou hudební sémantiku, následně vybrat způsob, kterým danou sémantiku zapsat, a vygenerovat rozložení symbolů na stránce. Poté se generují jednotlivé symboly a doplňují do vznikajícího obrázku. Na to navazuje syntéza inkoustu, pozadí, a deformací papíru.) Autor práce se postaral, aby byl jeho generátor symbolů do takového procesu zapojen (systém Mashcima) a vyhodnocuje své výsledky velmi korektně v tomto kontextu. Ukazuje, že syntéza symbolů pomocí zvolených generativních neuronových sítí (variační, adversariální a kategorické autoenkodéry) pomáhá vytvářet syntetické obrázky hudební notace, které pomáhají trénovat lepší systémy. Jedná se o dobře navrženou, dobře odvedenou a přes nedlouhý text poměrně rozsáhlou experimentální práci, kterou jednoznačně doporučuji k obhajobě; kvalitu výsledků dokládá ostatně i to, že již byly prezentovány i na mezinárodním workshopu o rozpoznávání notace WoRMS 2023.</p>				

Textová část práce

lepší OK horší nevyhovuje

Formální úprava ... <i>jazyková úroveň, typografická úroveň, citace</i>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Struktura textu ... <i>kontext, cíle, analýza, návrh, vyhodnocení, úroveň detailu</i>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Analýza	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vývojová dokumentace	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Uživatelská dokumentace	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Komentář Text sděluje, co je třeba, avšak často suboptimálním způsobem. Nejdůležitějším nedostatkem pro čitelnost práce je nedostatečný popis systému Mashcima. Tento systém pro syntézu obrázků notopisu je klíčový pro měření úspěšnosti implementované metody, avšak sekce 2.4, která jej uvádí, neobsahuje žádné příklady jeho výstupů, a koncept “attachemnt points”, který je předmětem takřka třístránkové sekce 3.1 a důvod pro používání řady metod “klasického” zpracování obrazu v obhajované práci, není vysvětlen. Stejně tak je potom špatně srozumitelná část 3.4 Generating output for Mashcima. Bylo by ideální workflow systému Mashcima ilustrovat v sekci 2.4 obrázkem, většina těchto nejasností by tak zmizela.

Práci by také prospělo více obrázků, především přehled celého systému. Nebylo by také špatné poskytnout příklady kompletních vygenerovaných obrázků po zapojení generátoru symbolů do systému Mashcima.

Občas problematický tok informací a jejich hierarchie. Například:

— V úvodu: “Thus we can divide it into two steps – generating symbol images and joining them together. Mashcima does the second step. This thesis does the first one.” Tyto informace jsou zásadní vymezení práce — bylo by lepší je nějak vizuálně odlišit.

(Například lze udělat z oněch dvou kroků seznams, krok 1 vyznačit tučně, a opatřit rovnou komentářem “1. (...) This step is the subject of this thesis.” a následně “2. (...) This is handled by the existing mashcima system.”)

— V části 2.4 je “(with denoted attachment points)” uvedeno v závorce, ačkoliv se zde takto mchodem uvádí předpoklad pro použití systému Mashcima, který následně vyžaduje celou sekci (3.1).

— Kapitola 4: Experiments by před částí 4.1 zasluhovala odstavec s přehledem provedených pokusů: je jich hodně, snadno se v nich ztrácí.

—Table 4.1: Nadpisy “Experiment 1” atd. by bylo vhodnější nahradit informativnějšími názvy: “all Primus”, “all synthetic”, “Mix 1/2”, “Mix all”.

Text je místy kostrbatý, nicméně neobratnosti jen výjimečně zasahují do terminologie či konvencí odborného stylu. (“derive”-> “differentiate”, “real space”-> “real-valued space”).

Bibliografie je adekvátně rozsáhlá, formátování by mohlo být konzistentnější (především u položek z repozitáře arxiv.org), avšak stále jsou všechny položky jednoznačně dohledatelné. Citace jsou korektní.

Zároveň má však text i řadu dobrých vlastností. Struktura textu je přehledná. Použité modely jsou popsány dobře. Příklady vygenerovaných symbolů, pravděpodobně čtenářsky nejdůležitější informace úspěšnosti generátoru, jsou dostačující. I přes uvedené nedostatky text není nepřesný či zavádějící a uvádí k experimentální části práce potřebné detaily pro replikaci. Relativně menší rozsah textu je více než vyvážen množstvím pokusů a svědčí spíše o chvályhodné tendenci autora nepopisovat provedenou práci zbytečně košatě. Celkově tak text považuji za vyhovující.

Implementační část práce

lepší OK horší nevyhovuje

Kvalita návrhu ... architektura, struktury a algoritmy, použité technologie	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Kvalita zpracování ... jmenné konvence, formátování, komentáře, testování	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Stabilita implementace	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Komentář Zdrojový kód jakožto implementace zmiňovaných modelů a experimentů stačí. Je vhodně strukturovaný, obsahuje vše potřebné pro instalaci, README je v pořádku (ačkoliv by možná bylo vhodnější jako příklad pro trénování něco jiného než “onecat”, což trénuje jeden model pro každou kategorii notačního symbolu, a tak uvedený příklad trvá poměrně dlouho). Implementace modelů je založená na knihovně tensorflow, kterou využívá vhodně. Poskytnutým kódem lze popisované pokusy replikovat. K modelům jsou dokonce implementované i (omezené) testy.				

Celkové hodnocení Výborně (spíše horší)

Práci navrhuji na zvláštní ocenění Ne

Datum 29. ledna 2024**Podpis**