

# Posudek diplomové práce

Matematicko-fyzikální fakulta Univerzity Karlovy

**Autor práce** Bc. Jiří Mayer  
**Název práce** Semi-supervised learning in Optical Music Recognition  
**Rok odevzdání** 2022  
**Studijní program** Informatika    **Studijní obor** Softwarové a datové inženýrství

**Autor posudku** doc. RNDr. Pavel Pecina, Ph.D    **Role** vedoucí  
**Pracoviště** Ústav formální a aplikované lingvistiky

## Text posudku:

**Úvod** Diplomová práce Jiřího Mayera se zabývá automatickým rozpoznáváním ručně psaných notových zápisů. Tato úloha je podobná rozpoznávání ručně psaného textu, je ale mnohem složitější, a to nejen kvůli velké variabilitě stylu, ale také kvůli nelineárnosti zápisu (nejenom horizontální, ale i vertikální umístění symbolů má význam). Zásadním problémem je také omezená dostupnost trénovacích dat, která jsou nutná pro nasazení moderních metod strojového učení. Nedostatek anotovaných trénovacích dat lze obecně řešit buď automatickým generováním tzv. syntetických trénovacích dat nebo využitím neanotovaných dat pro tzv. *semi-supervised* učení. Zatímco první možnost diplomant zkoumal již v bakalářské práci, druhou možností se zabývá v předkládané diplomové práci. Její téma je vysoce aktuální. Problematice rozpoznávání ručně psaných notových zápisů se z uvedených důvodů nevěnuje velká pozornost, i když by její řešení mělo praktické využití (zpracovávání archivních dokumentů, ale i komponování nové hudby).

**Obsah** Práce je strukturována do 7 kapitol. Po obsáhlém úvodu (kapitola 1) se autor věnuje shrnutí existující literatury (kapitola 2), popisu studované úlohy včetně diskuse dřívějších a aktuálních přístupů (kapitola 3) a dále problematice využití neanotovaných dat pro *semi-supervised* učení neuronových sítí (kapitola 4). Hlavní přínos práce je potom obsažen v kapitole 5, kde je popsána použitá architektura neuronové sítě a také data, s kterými autor pracoval, a v kapitole 6, které popisuje provedené experimenty a jejich výsledky. Práci uzavírá kapitola 7, která shrnuje dosažené výsledky a představuje možné směry dalšího výzkumu. Text je opatřen bohatou bibliografií.

Autor velmi pečlivě nastudoval problematiku rozpoznávání notových zápisů a soustředil se na úlohu tzv. segmentace, pro kterou se dříve ukázala vhodná architektura U-Net. Tu pozměnil tak, aby mohla zpracovávat i neanotovaná data a konkrétně řešit tzv. *denoising*, kdy se ta samá síť (až na poslední vrstvu) učí rekonstruovat obrázky not, jejichž části byly maskovány (vymazány). Návrh této architektury je jedním z hlavních přínosů práce, ale neméně cenná je také pečlivá

analýza možností pro trénování (včetně stanovení hodnot hyperparametrů, volby ztrátové funkce, aktivačních funkcí, řešení problémů se *skip-connections*, *drop-outu* a skladby trénovacích dat atd.) a návrh celého trénovacího postupu, který vedl k úspěšnému natrénování. Výsledná síť sice nedokázala překonat tzv. *supervised baseline* (tedy síť natrénovanou jen na anotovaných datech), ale podrobná analýza ukázala, že její trénování je stabilnější a použití neanotovaných trénovacích dat vede k regularizaci (předchází přetrénování). Navrženou architekturu lze v principu použít pro úlohu segmentace všech hudebních symbolů, i když se autor soustředil jen na notové hlavičky. Rozšíření na další symboly je jen technická záležitost, které by ovšem vyžadovala další nemalé úsilí. Navíc by tento krok měl být proveden, až poté, co je nalezeno řešení, které dokáže více těžit z použití neanotovaných dat a poskytuje lepší výsledky než *supervised* přístup.

**Hodnocení** Po formální stránce je práce zcela v pořádku. Je psaná anglicky, bez chyb, vhodně strukturovaná, psaná srozumitelně, čtivě, opatřena přehlednými obrázky. Po obsahové stránce je práce na vynikající úrovni. Lze ji hodnotit jako výzkumně experimentální a má všechny náležitosti vědecké práce. Text práce ukazuje, že autor postupoval precizně a řešil správné problémy odpovídajícím způsobem. Studovaná problematika je ovšem mnohem bohatší a tak práci považuji za vynikající základ pro hlubší prozkoumání této oblasti a studium dalších způsobů využití neanotovaných dat pro rozpoznávání notových zápisů, kterému se autor bude věnovat během postgraduálního studia na MFF.

**Závěr** Předkládaná práce Bc. Jiřího Mayera splňuje požadavky kladené na diplomovou práci na MFF UK.

**Práci doporučuji k obhajobě.**

**Práci nenavrhuji na zvláštní ocenění.**

V Praze dne 31. 5. 2022

Podpis: