

Analýza dokumentů hudební notace systémem YOLO

Specifikace ročníkového projektu

Vojtěch Dvořák



Kontext

Optické rozpoznávání notopisů (Optical Music Recognition, OMR) je úloha, jejíž cílem je z fotky nebo skenu hudební notace automaticky rozpoznat hudební obsah dokumentu pro účely dalšího zpracování (vyhledávání, tisk, editace, muzikologická analýza). Pro zpracování existujících archiválií v knihovnách je ovšem nutné nejprve rozlišit, zda daná stránka obsahuje hudební notaci, případně jaký druh notace (moderní, menzurální, chorální, tištěnou, psanou). V současnou chvíli nejlepší rozpoznávací modely pracují na úrovni jednotlivých řádků notace, proto je také nutné během analýzy dokumentu identifikovat jednotlivé řádky, aby mohla být hudebnina dobře zpracována rozpoznávacími modely.

Cíle projektu

Cílem projektu je vytvořit analyzátor stránek s hudební notací, který bude sloužit jako první fáze při automatickém rozpoznávání stránek hudební notace. Analýza bude koncipovaná jako image object detection task, kde hlavním cílem bude natrénovat (přetrénovat) state of the art model YOLOv8. Těžiště práce leží v přípravě trénovacích dat, jelikož bude nutné harmonizovat řadu existujících datasetů do jednotného formátu s případným začleněním syntetických dat. Zároveň bude nutné rozhodnout, jaké parametry hudební notace budeme analyzovat - je třeba najít kompromis mezi tím, co potřebujeme umět a tím, co jsme schopni se z dostupných datasetů naučit.

Použité technologie

Projekt bude realizovaný jako sada transformačních a orchestračních skriptů v Pythonu, organizovaných pomocí python modulů do čitelné a modulární struktury. Transformace dat bude využívat knihovny numpy, opencv, svgelements, cocotools, muscima. Interakce s YOLO modelem probíhá přes knihovnu ultralytics. Pro syntézu dat lze použít software MuseScore nebo knihovnu mashcima.