

Program na podporu výuky prstové abecedy

Autor: Mgr. Filip Martiník Vedoucí: doc. RNDr. Miroslav Kolařík, Ph.D.

Univerzita Palackého v Olomouci
Přírodovědecká fakulta
Katedra informatiky

Představení

V práci byla vytvořena inovativní webová aplikace, která umožňuje procvičování prstových abeced a základních znaků znakového jazyka. Zaměřuje se zejména na jednoruční českou prstovou abecedu. V různých režimech nabízí uživateli výběr, jak se chce zmíněné abecedy učit nebo procvičovat.

Animace prstové abecedy

Prstová abeceda je v aplikaci zobrazovaná ve 3D animaci. To umožňuje pohodlné zobrazení reprezentace znaků z jakéhokoliv úhlu v libovolné rychlosti přehrávání s možností zastavení.

Pro každou abecedu je vytvořena jedna výchozí pozice kostry, ze které vycházejí animace znaků. Prolnáním konce animace jednoho znaku se začátkem animace dalšího znaku vzniká plynulý přechod.

Řídící algoritmus aplikace přehrává jednotlivé znaky vstupu s vhodnými prodlevami mezi nimi.

Aplikace obsahuje animace pro jednoruční a dvouruční českou prstovou abecedu včetně diakritiky a také pár vybraných znaků českého znakového jazyka.

Prstová abeceda

- zobrazení písmen abeced pomocí pozice rukou a prstů
- rozdělení na jednoruční, dvouruční
- každý znakový jazyk má svou vlastní
- použití: vlastní jména, termíny, slova bez stanoveného znaku

Rozpoznání znaku z videa

Pro každý zaznamenaný snímek z videa:

1. Rozpoznání 21 souřadnic ruky
MediaPipe Hand Landmarker
2. Zpracování a normalizace souřadnic
3. Klasifikace souřadnic
Vytvořený Gradient Boosted Trees model

Dataset a trénování modelu

- 125 snímků pro každý znak a každou ruku
- 2 osoby, dataset byl vytvořen v rámci práce
- Augmentace dat pomocí horizontálního zrcadlení

Gradient Boosted Trees model poskytuje nejpřesnější výsledky z vícero prozkoumaných metod.
- Přesně a rychle klasifikuje zobrazený znak

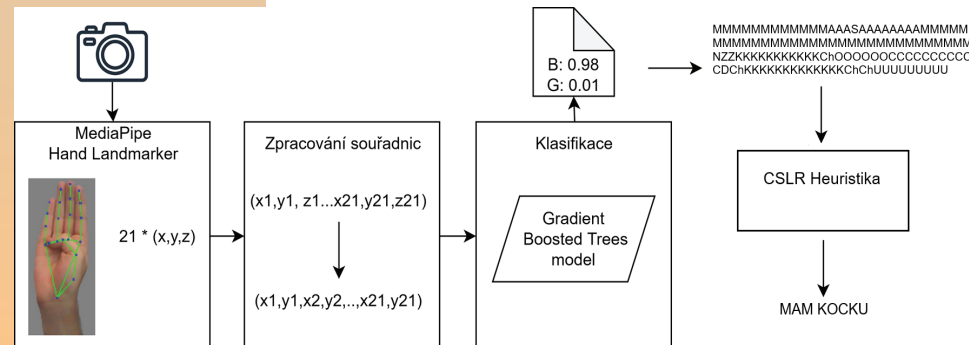
Rozpoznání souvislého znakování

Vytvořený klasifikátor postupně zaznamenává všechna rozpoznání.

Nad touto posloupností je provedena heuristika pro transformaci na skutečně znakovanou posloupnost.

Technologie

Blender, Three.js, SvelteKit, Tailwind, PWA, TensorFlow, MediaPipe, Python, Ollama



Propojení s LLM

Nedokonalostí souvislého rozpoznávání je výskyt chyb v textu v důsledku nesprávné klasifikace znaku. Tento problém lze řešit jazykovými modely, které umí pracovat s kontextem okolo nesprávně rozpoznávaného slova a tím porozumět i slovu s chybou.

Režim konverzace

Režim konverzace využívá zmíněné vlastnosti a umožňuje uživateli komfortně komunikovat s jazykovým modelem prostřednictvím prstové abecedy.

Uživatel zadá vstup pomocí souvislého rozpoznávání, který je následně odeslán jazykovému modelu, jehož odpověď je interpretována animací prstové abecedy.

Tímto vzniká interaktivní prostředí pro procvičování znakování a odezírání prstové abecedy s možností vést zajímavé konverzace s umělou inteligencí.

Ukázka aplikace

