# Jelnyelv fordítása, hallássérült jelelők támogatására, póz approximációs módszerek, szekvencia feldolgozásra alkalmas mesterséges intelligencia algoritmusok és nagy nyelvi modellekre épülő eredmény korrekció felhasználásával

A nyelvfeldolgozás terén elért sikerek rengetek kaput nyitottak ki előttünk. A hangalapú asszisztensek, alap esetben is hasznos, de legtöbbször forradalmi funkciókat hoznak életünkbe. A mesterséges intelligencián alapuló „voice to text” modellek manapság gyakorlatilag tetszőleges nyelvről képesek felismerni szöveget, valamint a közelmúltban nyelvfeldolgozás területén elért eredményeknek köszönhetően pedig nincsenek határok, a felhasználók, és fejlesztők előtt. Illetve előbbi állítás sajnos csak egy nagy ferdítéssel igaz. Ugyanis világszerte körülbelül 1,5 milliárd ember hallássérült, és több mint 70 millió ember használja a jelnyelvet, mint elsődleges kommunikációs formát. Számukra „saját nyelvükön” ezek a funkciók nem, vagy csak korlátozottan érhetőek el. A jelnyelv egy speciális fajtája az ujjbetűzés. Gyakran használják nevek, címek, telefonszámok, valamint olyan fogalmak közvetítése során, melyekre nincs bevett gesztus. Egy tapasztalt jelelő képes közel kétszer olyan sebesen betűzni, mint egy virtuális billentyűzeten pötyögő egyén, nem beszélve arról ha minden jelet alkalmazva kommunikál. Így égető egy számukra is kényelmesen használható interfész kialakítása.

Még jelentősebb problémát vet fel a kommunikáció kérdése. Manapság a nyelvek közötti gépi fordítás széles körben és formátumban elérhető az internet kapcsolattal rendelkezőknek. Régen túl vagyunk már az egyszerű szótár alapú fordítókon, a különböző transzformer, és nagy nyelvi modellek térhódításával gyakorlatilag tetszőleges nyelvek között megoldható a kommunikáció. Ez alól kivétel azonban a jelnyelv. Nem létezik olyan megoldás a piacon, ami segíti a jelelő kommunikációját a jelelni nem tudó személy felé. A terület előrehaladásán nem segít, hogy a jelnyelv csakúgy mint a verbális párjai, nem standardizált. Szinte minden nemzetnek saját jelnyelve van, nem beszélve a helyi sajátosságokról, és dialektusokról. Tovább nehezíti a feladatot, hogy a rendelkezésre álló, nagy méretű, és jó minőségű adatbázisok száma eltörpül a klasszikus nyelvfeldolgozásban megszokottaktól.

Munkámmal a széles tömegek számára elérhető jelnyelv fordító rendszerek fejlesztéséhez járulok hozzá. Megvizsgálom többek között a kesztyűvel történő felismerés, valamint hagyományos képfeldogozás eredményeit, továbbá részletesen foglalkozom a póz approximáción alapuló algoritmusokkal. A nyílt kérdés utóbbival kapcsolatban, hogy bár rendkívül jó arányban tömöríti a képi adatot, ami lehetővé tenné nagy mennyiségű, egységes adatbázis konstruálását, de vajon tart-e ott a technológia, hogy megbízható módon kódoljon minden fordításhoz szükséges információt?

Az amerikai jelnyelven belül külön-külön vizsgálom az ujjbetűzés, és általános jelbeszéd lehetőségeit, kitérve a pillanatképből, valamint mozgásszekvenciából dolgozó megoldásokra. A teljesség igénye nélkül összehasonlítom a feladatra adaptált konvolúciós, rekurrens, LSTM, Transzformer háló architektúrákat. A jelelés nyelvtani adottságai következtében, hiányoznak a segédigék, valamint gyakran más a szórend a hagyományos angolhoz képest. A primitív fordítás eredményeit ezért kontextus függő módon transzformálom generatív nagy nyelvi modellekkel, valamint az ujjbetűzés esetében a megbízhatóság növelésére is felhasználom őket.

A dolgozatomban bemutatom az adatgyűjtés és a tervezés lépéseit, az alkalmazott mesterséges intelligencia algoritmusok részleteit, valamint elemzem a fejlesztés során felmerült tervezői döntéseket, és alternatíváikat. Ezen felül bemutatásra kerül az általam készített teljes megoldás, mely nagymértékben segíthet a hallássérültekkel való kommunikáció során.