

UNIVERSITETI I PRISHTINËS “HASAN PRISHTINA”

FAKULTETI I SHKENCAVE MATEMATIKE-NATYRORE

DEPARTAMENTI I MATEMATIKËS

PROGRAMI SHKENCË KOMPJUTERIKE



Përdorimi i IA në krijimin e modelit për interpretimin e gjuhës së shenjave në Shqip

(Punim Diplome-Master)

Mentori:

Prof. asoc. Dr. Ermir Rogova

Kandidati:

Bsc. Egzon Mjeku

Prishtinë

Qershor 2025

Abstrakti

Kjo punë fokusohet në ndërtimin e një modeli inteligjent për njohjen automatike të gjuhës së shenjave shqipe, duke përdorur teknologji të avancuara të përpunimit të të dhënave vizuale dhe të mësimin të thellë. Qëllimi kryesor është zhvillimi i një sistemi që mund të interpretojë shenja të ndryshme të gjuhës së shenjave në kohë reale, duke ndihmuar në përmirësimin e komunikimit për personat me vështirësi në dëgjim dhe të folur. Arkitektura e propozuar bazohet në përdorimin e Mediapipe Hands për detektimin e 21 pikave kyçe në secilën dorë, të cilat shërbejnë si input për një rrjet të avancuar Long Short-Term Memory (LSTM), i cili është projektuar për të kapur varësitë kohore midis lëvizjeve të dorës përgjatë një sekuence të regjistruar.

Një aspekt qendror i kësaj pune ka qenë ndërtimi i një databaze të personalizuar, duke qenë se nuk ekziston një dataset i gatshëm për gjuhën e shenjave shqiptare. Të dhënat janë mbledhur manualisht duke përdorur kamerën e laptopit dhe librarinë Mediapipe për të regjistruar të dhënat e nevojshme. Për secilën shenjë janë regjistruar 50 sekuenca unike, secila prej tyre me 30 kuadro (frames), duke përfshirë alfabetin dhe numrat nga 0 deri 9. Të dhënat e mbledhura janë përpunuar përmes teknikave të normalizimit dhe filtrimit të zhurmës, për t'u siguruar që secila sekuenca të përfaqësojë me saktësi lëvizjen specifike të shenjës.

Modeli LSTM është ndërtuar dhe trajnuar duke përdorur platformën TensorFlow, me një arkitekturë e cila përfshin disa shtresa LSTM dhe Dense, të kombinuara me teknikën e dropout për të shmangur mbipërshtatjen (overfitting). Procesi i trajnimit është mbështetur në ndarjen e të dhënave në sete trajnimi dhe testimi, me qëllim vlerësimin e performancës së modelit në të dhëna të panjohura. Metrikat e përdorura për vlerësim përfshijnë saktësinë (accuracy), precizionin, recall-in dhe F1-score, duke ofruar një pamje të qartë të kapacitetit të përgjithshëm të modelit. Rezultatet eksperimentale kanë treguar një shkallë të kënaqshme të saktësisë, duke dëshmuar që qasja e propozuar është premtuese për zbatime reale në sisteme ndihmëse të bazuara në inteligjencë artificiale.

Fjalët kyçe: LSTM, Gjuha e shenjave kosovare, njohja e gjuhës së shenjave, ndërtimi i datasetit, hand landmark, MediaPipe, Tensor Flow

Abstract

This thesis focuses on the development of an intelligent system for the automatic recognition of Albanian Sign Language, using advanced technologies in visual data processing and deep learning. The main objective is to build a system capable of interpreting various hand signs in real time, thereby improving communication for individuals with hearing and speech impairments. The proposed architecture utilizes Mediapipe Hands to extract 21 keypoints from each hand, which serve as the input features for a deep Long Short-Term Memory (LSTM) network designed to capture the temporal dependencies across hand movement sequences.

A central component of this research is the construction of a custom dataset, as no publicly available dataset currently exists for Albanian Sign Language. The dataset was manually recorded using a laptop webcam and the Mediapipe library, with 50 unique sequences recorded for each sign, and each sequence consisting of 30 frames. The signs include the alphabet, and numbers from 0 to 9. The collected data was preprocessed through normalization and noise filtering techniques to ensure that each sequence accurately reflects the intended hand gesture.

The LSTM model was developed and trained using the TensorFlow framework. The architecture includes multiple LSTM and Dense layers, combined with dropout layers to prevent overfitting. The training process involved splitting the data into training and testing sets, enabling the evaluation of the model's performance on unseen samples. Evaluation metrics such as accuracy, precision, recall, and F1-score were employed to assess the classification performance comprehensively. Experimental results demonstrate that the proposed system achieves a satisfactory level of accuracy and generalization, suggesting strong potential for real-world implementation in AI-based assistive technologies.

Keywords: LSTM, Kosovar Sign Language, sign language recognition, dataset construction, hand landmark, MediaPipe, Tensor Flow