

# **Proposal**

Degree Program BIF-SS2025

Course Scientific Writing

## **Interaktive Gebärdensprachübersetzung mit Feedbackschleife**

By: Stefan Rezk

Student number: 2310257056

Advisor: Marius Führer

Wien, 04.12.2024

# Einführung / Hintergrund

## Hintergrund und Kontext

Die Erkennung von Gebärdensprache ist ein zentrales Forschungsfeld zur Verbesserung der barrierefreien Kommunikation und der Integration von gehörlosen Menschen. Bisherige Forschungsarbeiten [1], [3], haben gezeigt, dass tiefe Lernmodelle vielversprechende Ergebnisse in der Erkennung und Übersetzung von Gebärden liefern. Bisherige Modelle konzentrieren sich hauptsächlich auf die Erkennung und Übersetzung von Gebärden in Text, dennoch fehlen oft Methoden, die eine interaktive Rückmeldung ermöglichen [4]. Ein interaktives System, das unmittelbare Rückmeldungen gibt, könnte die Erlernung und Verbesserung der Gebärden unterstützen und dazu beitragen, die Benutzerfreundlichkeit solcher Technologien zu erhöhen. Die Entwicklung eines solchen Systems bietet nicht nur einen praktischen Nutzen, sondern auch wissenschaftliche Erkenntnisse über die Effektivität von Feedback im Lernprozess.

## Problembeschreibung

Derzeit fehlt es an interaktiven Systemen, die eine Rückmeldung zur Qualität der ausgeführten Gebärden geben. Es wurde gezeigt, dass eine fehlende Rückmeldung zu einer verminderten Lernqualität führen kann [5], was die Notwendigkeit interaktiver Ansätze unterstreicht. Modelle zur Gebärdensprachenerkennung sind zwar in der Lage, Gebärden zu erkennen, bieten jedoch keine direkte Information darüber, wie gut die Gebärde ausgeführt wurde. Ein solches Feedback wäre jedoch wichtig, um die korrekte Ausführung zu erlernen und die Konsistenz zu verbessern.

## Relevanz für das Fachgebiet

Das Projekt zielt darauf ab, die bestehenden Erkennungsmodelle, um eine Feedbackschleife zu erweitern, wie vorgeschlagen wurde, dass zusätzliche Informationsquellen und Rückmeldungen die Genauigkeit und Effizienz erhöhen können [2]. Dies könnte die Effizienz des Lernprozesses verbessern und Menschen dabei helfen, die Gebärdensprache schneller und präziser zu erlernen. Die direkte Rückmeldung bietet den Benutzern eine Möglichkeit zur Selbstkorrektur, was zu einer besseren Nutzererfahrung und einer höheren Lernmotivation führt.

# Forschungsfrage

Wie effektiv ist die Verwendung einer Feedbackschleife zur Verbesserung der Erkennungsgenauigkeit und Konsistenz der Gebärdensprachenausführung? Diese Frage baut auf bisherigen Erkenntnissen auf, die die Bedeutung von robusten Rückmeldesystemen zur Verbesserung der Lernprozesse hervorgehoben haben [6].

## Ziel der Bachelorarbeit

Das Ziel der Arbeit ist es, ein interaktives System zur Gebärdensprachenerkennung zu entwickeln, das den Benutzer unmittelbar darüber informiert, ob eine Gebärde korrekt ausgeführt wurde. Hierbei wird ein Prototyp entwickelt, der die Gebärden in Echtzeit erkennt und auf Grundlage der Erfolgsquote über mehrere Wiederholungen eine Rückmeldung zur Qualität der Gebärde gibt.

# Methoden

## Verwendete Methoden

- **Modellentwicklung:** Es wird ein selbsttrainiertes Modell zur Erkennung von Gebärden der amerikanischen Gebärdensprache (ASL) verwendet, das mithilfe des WLASL2000-Datensatzes angepasst und trainiert wird.
- **Benutzeroberfläche:** Eine **einfache Benutzeroberfläche** wird entwickelt, die es den Nutzern ermöglicht, Gebärden vor der Kamera auszuführen und direktes Feedback zu erhalten. Dies kann über eine Webanwendung oder eine Desktop-App geschehen.
- **Rückmeldemechanismus:** Die Rückmeldung basiert auf der **Erfolgsquote** der erkannten Gebärden über mehrere Wiederholungen. Das System zeigt dem Nutzer eine Bewertung der Gebärde in Form einer Erfolgsquote an, z.B. durch ein einfaches Ampelsystem (grün, gelb, rot).

## Begründung der Methodenwahl

- Direkte Anwendung in einem praktischen Umfeld.
- Unmittelbares Feedback durch etablierte Metriken.
- Einfache Benutzeroberfläche zur intuitiven Bedienung.

## Erwartete Ergebnisse

- Ein funktionierender Prototyp, der Gebärden erkennt und unmittelbare Rückmeldung zur Qualität der Gebärde gibt.
- Erkennungsgenauigkeit von über 85%, mit Beeinträchtigung diverser Umgebungsstörfaktoren.
- Erkenntnisse zu potenziellen Verbesserungen der Technologie.

## References

- [1] Yin Xie "American Sign Language Recognition using Deep Learning Models." In Proceedings of the 2023 International Conference on Computer Vision.
- [2] Molchanov, P. "Gesture Recognition using RGB-D Data." IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021.
- [3] Camgoz, N. "Neural Sign Language Translation." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018.
- [4] Koller, O. "Continuous Sign Language Recognition: Towards Large Vocabulary Statistical Recognition Systems Handling Multiple Signers." Computer Vision and Image Understanding, 2019.
- [5] Huang, J. "Video-based Sign Language Recognition without Temporal Segmentation." In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020.
- [6] Rastgoo, R. "Sign Language Recognition: A Deep Survey." Expert Systems with Applications, 2020.