Friedrich-Alexander-Universität
Erlangen-Nürnberg

MASTER THESIS

To students within
Medical Engineering / Medizintechnik,
Computer Science/ Informatik / Artificial Intelligence,
Data Science, Computational engineering

Abteilung für
Phoniatrie und Pädaudiologie
Prof. Dr.-Ing. Michael Döllinger

Telefon: 09131 85-33814
Fax: 09131 85-32687
Michael.doellinger@uk-erlangen.de
Waldstr. 1, 91054 Erlangen
CICERO-Gebäude: Raumerstr. 1a

03.08.2026

Title: Development of a Neural Network-Based Preprocessing Method for Artifact Reduction in High-Speed Vocal Fold Recordings

Background: High-speed recordings of oscillating vocal folds are used to investigate vocal fold dynamics and to analyse the complex processes involved during phonation. A key outcome of these recordings is the extraction of the glottal area waveform (GAW), which describes the temporal variation of the glottal opening and provides important information for the analysis of vocal fold vibration. The GAW is commonly obtained using grayscale-based image segmentation, which assumes a strong intensity contrast between the dark glottal opening and the surrounding vocal fold tissue.
In experimental studies, laser illumination has been employed to visualize airflow and particle motion above the glottis. While this enables the simultaneous investigation of aerodynamic parameters, the introduced laser sheet produces bright illumination artifacts in the recorded images. Since the intensity of these artifacts varies throughout the oscillation cycle as the glottis opens and closes, standard threshold- based segmentation methods fail or produce significant errors in the extracted GAW. Consequently, the quality of the subsequent quantitative analysis strongly depends on the ability to suppress these illumination artifacts while preserving the anatomical structures of the vocal folds.
Recent advances in deep learning and image restoration have demonstrated considerable potential for removing structured image artifacts. Neural network-based preprocessing methods are capable of learning complex image transformations and therefore offer a promising alternative to conventional image processing techniques. This thesis aims to investigate such a preprocessing approach for suppressing laser-induced illumination artifacts while preserving the anatomical information required for accurate glottal segmentation.

The overall goal of this project is to develop and evaluate deep learning-based preprocessing method for suppressing illumination artifacts in high-speed vocal fold recordings.

Current State: A dataset of high-speed vocal fold recordings acquired under different conditions is already available, together with an established in-house software for GAW extraction [1]. Previous work [2] investigated conventional preprocessing methods for reducing illumination artifacts caused by localized reflections, demonstrating the potential of preprocessing but without using deep learning approaches.

Research Questions

1. Can neural network-based image restoration methods effectively suppress laser-induced illumination artifacts while preserving the anatomical structures of the vocal folds?
2. Which deep learning architecture provides the best trade-off between artifact removal, image fidelity and computational efficiency?
3. To what extent does neural network-based preprocessing improve the robustness and accuracy of subsequent glottal segmentation and GAW extraction compared with conventional preprocessing methods?
4. How well does the developed method generalize across different experimental conditions and integrate into the existing GAW analysis workflow?

Methods & Tasks

1. Review the literature on high-speed vocal fold imaging, biomedical image restoration and deep learning-based artifact reduction.
2. Design, implement and analyse three – five deep learning models for suppressing laser-induced illumination artifacts in high-speed vocal fold recordings.
3. Quantitatively evaluate different network architectures and compare them with existing preprocessing approaches using downstream GAW extraction performance.
4. Integrate the developed preprocessing method into a software tool for efficient processing and analysis of high-speed recordings.

Goal: Develop and validate a deep learning-based preprocessing framework for suppressing laser-induced illumination artifacts in high-speed vocal fold recordings while preserving anatomical information required for accurate glottal segmentation and reliable extraction of the glottal area waveform.

The work will be supervised by **Dr. rer. nat. Helena Vulić** and **Prof. Dr.-Ing. Michael Döllinger** (Member of Department Informatics & AIBE) at FAU.

We search for a dedicated and motivated student with

- Very good grades
- Experience in programming language (Python or MATLAB)
- Good knowledge of machine learning and image processing
- Independent and structured working style
- Good written and verbal communication skills in English
- Good written and spoken English

Application: Please send CV and current transcripts.

Contact: Dr. rer. Nat. Helena Vulić (helena.vulic@uk-erlangen.de), Prof. Dr.-Ing. Michael Döllinger (michael.doellinger@uk-erlangen.de)

References:

1. Kist et al.: A Deep Learning Enhanced Novel Software Tool for Laryngeal Dynamics Analysis, 2021.
2. Parchent J.: Filtering of High-Speed Video Recordings of the Vocal Folds for Optimisation of Glottis Segmentation. Bachelor thesis, 2016.