Topographical Scalp Maps for Hearing Impairment Detection: A CNN Approach

Grace Wang, Dr. Beiyu Lin

July 2023

1 Abstract

It is estimated that more than 50% of people over the age of 65 experience some form of hearing loss. The goal of this project is to utilize scalp maps derived from spatial and temporal electroencephalography (EEG) data to distinguish between individuals with hearing impairment and healthy individuals. A CNN model will be trained to recognize patterns in the spatial and temporal EEG data that can accurately classify the two groups. After experimenting with the CNN model and data splits, the model achieves an accuracy of $85\% - 95\%$. The use of advanced technology like CNNs can potentially lead to more efficient and accurate hearing impairment detection methods, benefiting individuals and healthcare professionals alike.

2 Introduction

The normal hearing threshold for healthy individuals is 0 to 20 dB. A person is said to be hearing impaired if their hearing threshold is greater than 35 decibels (dB) in the better hearing ear. This means it becomes challenging for the affected individual to communicate without the use of hearing aids or assistive devices.

Current diagnostic methods, such as otoacoustic emission testing, MRI scans, and CT scans, assist in distinguishing between healthy individuals and those with hearing loss. However, this study explores an alternative approach by utilizing scalp maps derived from spatial and temporal EEG data, which records the brain's electrical activity through electrodes on the scalp. EEG offers affordability and non-invasiveness, making it a valuable tool in studying brain activity, cognitive processes, and neurological disorders. A deep learning model based on CNNs is employed due to its image-based nature, ideally suited for the visual nature of scalp maps. The potential of CNNs in hearing impairment detection could lead to earlier intervention and cost savings in healthcare by minimizing the need for extensive treatments resulting from untreated hearing loss.

3 Terminologies

- *EEG (Electroencephalography):* EEG is a technique used to record the electrical activity of the brain. It measures the electrical potentials generated by the brain's neurons through electrodes placed on the scalp.

- *ERPs (Event-Related Potentials):* ERPs are specific components of the EEG signal that are time-locked to a particular event or stimulus, such as the presentation of a tone or visual stimulus. ERPs represent the brain's electrical response to these specific events and are characterized by distinct waveforms.

- *Scalp Maps:* Scalp maps, also known as topographical maps, are graphical representations of the spatial distribution of brain electrical activity recorded by EEG electrodes on the scalp. They provide a visual depiction of the magnitude and location of brain activity at specific time points or time intervals.

- *Spatial EEG Data:* Spatial EEG data refers to the patterns of electrical activity across different regions of the scalp. These patterns can be analyzed to identify brain regions involved in specific cognitive processes or responses to stimuli.

- *Temporal EEG Data:* Temporal EEG data refers to the patterns of electrical activity over time. It captures the dynamic changes in brain activity in response to events or stimuli.

- *EEG Preprocessing:* EEG preprocessing involves data cleaning and enhancement steps to remove noise, artifacts, and other sources of interference from the raw EEG signals. Common preprocessing steps include filtering, artifact removal, and baseline correction.

- *Feature Extraction:* Feature extraction is the process of identifying and extracting relevant features from the ERP waveforms. These features capture important characteristics of brain activity and are used as input to machine learning models for classification tasks.

4 Database

In this study, we used a dataset containing EEG of 22 hearing impaired (HI) listeners with sensorineural hearing loss and 22 age matched normal hearing (NH) peers (between 51 - 76 years old). The data was collected from a previous study by Fuglsang et al. (2020). For this study, we are specifically looking at the tone stimuli experiment where during passive listening, event-related potentials (ERPs) were recorded while subjects were exposed to 1 kHz pure tones. These tones had a duration of 100 ms and were smoothly ramped using a 10 ms Hann window. The presentation rate of the tones was around 1 second on average,

with random fluctuations of up to ± 25 ms (jittered). Each subject underwent 180 repetitions of the tone stimuli.

5   Methods

The study consists of 4 main parts, including EEG signal processing, ERP extraction, extracting topographic brain maps, and training, validating, and testing the model. The first part includes EEG signal preprocessing and extraction of event-related potential features, which are used to draw the topographic brain map.

5.1   EEG Preprocessing

EEG preprocessing and most of the ERP extraction were already completed by the researchers of the dataset. The explanation for how they preprocessed and extracted ERP is detailed in the paper, and the repository that contains code for the analysis of auditory EEG data is located at `https://gitlab.com/sfugl/snhl`.

5.2   ERP Extraction

In the additional figures section of the repository at the bottom of the 'project overview,' Fig S2 contains a graph of the individual traces of ERP data averaged over the fronto-central electrode cluster, and the thin lines reflect the data from individual subjects. Thick lines show group mean averages. Within the repository, there is a script that calculates the mean average lines within the `snhl/reports/paper/func/extract_erp.m` file path.
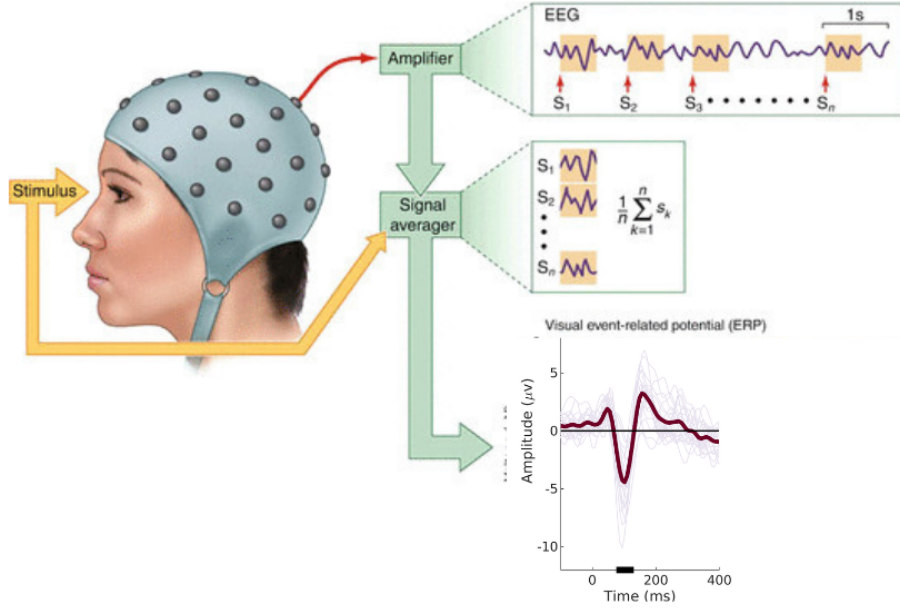
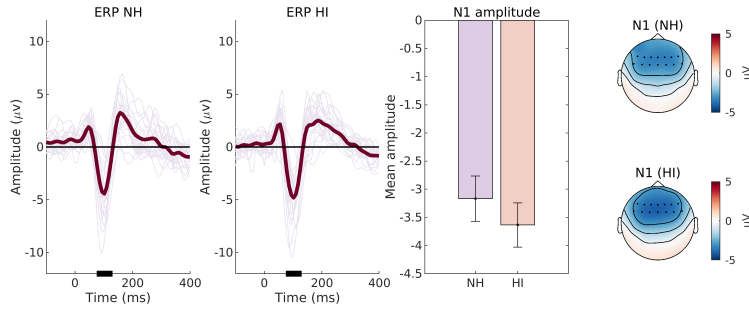Figure 1: Visual representation of how ERPs are derived from EEG data.

Figure 2: First and second panel (from left): Individual traces of ERP data averaged over the same fronto-central electrode cluster that was used for the entrainment analysis. Thin lines reflect data from individual subjects. Thick lines reflect group-mean averages. Third panel: Mean amplitude of N1 ERPs averaged over a time interval from 75 ms to 130 ms. Errorbars show s.e.m. across listeners. Fourth (right) panel: Group-mean topographies of mean N1 amplitude.

5.3 Topographic Brain Maps

ERPs serve as the key features for classifying the two groups. Using the ERP waveform of each patient, recorded up to 400 ms after the onset of tone stimuli, we developed a MATLAB script to systematically process each individual's ERP data. This script enabled the extraction of topographical scalp maps at 10 ms intervals within the time range of 0 to 400 ms. This approach resulted in the creation of 40 scalp maps for each subject.

The choice of a 10 ms interval was motivated by the fact that the smallest interval we could define was 8 ms. Opting for a 10 ms interval w s a reasonable decision due to its rounded nature. Consequently, this methodology yielded a total of 1,760 images, corresponding to 44 subjects and 40 images per subject. These images were subsequently used as input for our CNN model.
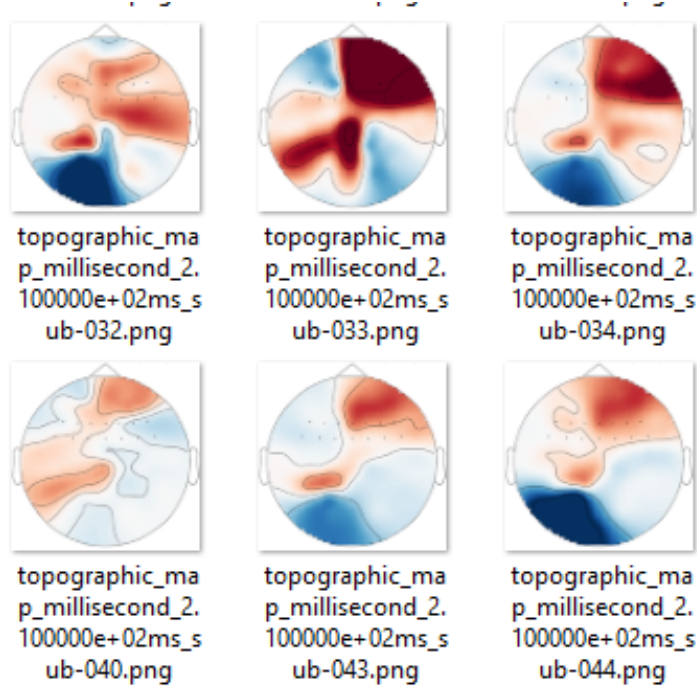

Figure 3: Examples of scalp maps

5.4 Convolutional Neural Network (CNN) Architecture

We employed a CNN model to perform image classification on the dataset. The CNN model was implemented using the TensorFlow Keras library in a python notebook. The architecture consists of three convolutional layers, each followed by a max-pooling layer to extract relevant features from the input images. Specifically, the layers had filter sizes of 32, 64, 128, respectively with

a ReLU activation function applied after each convolutional layer. Max-pooling layers with a 2x2 kernel were used to downsample the feature maps. The final layers included a flatten layer followed by a fully connected dense layer with 512 units and a ReLU activation function. A dropout layer with a dropout rate of 0.5 was applied to mitigate overfitting. The output layer consists of two units representing the two classes (Hearing Impaired and Healthy), with a softmax activation function to predict the probability distribution of each class.
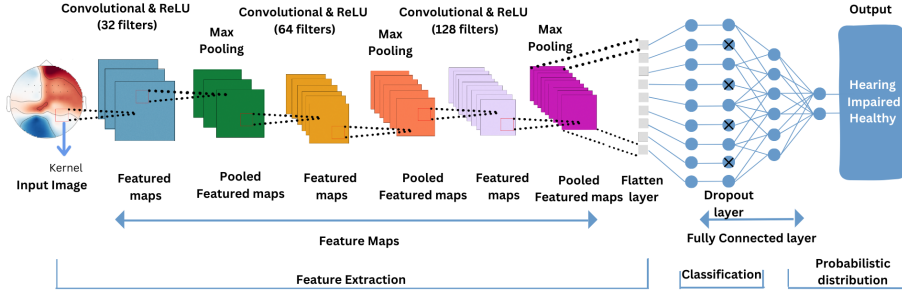

Figure 4: visual representation of CNN model

6 Results

6.1 Experiments

Due to the limited size of our dataset, consisting of only 1,760 total images, we conducted various experiments to determine the most suitable division of data for training, validation, and testing purposes. As a result of these experiments, we obtained varying test accuracies.

6

Table 1: Classification Report (Experiment 1)

	Precision	Recall	F1-score	Support
Healthy	0.86	0.86	0.86	264
Hearing Impaired	0.86	0.86	0.86	264
Accuracy			0.86	528
Macro Avg	0.86	0.86	0.86	528
Weighted Avg	0.86	0.86	0.86	528

Table 2: Classification Report (Experiment 2)

	Precision	Recall	F1-score	Support
Healthy	0.98	0.92	0.95	88
Hearing Impaired	0.92	0.98	0.95	88
Accuracy			0.95	176
Macro Avg	0.95	0.95	0.95	176
Weighted Avg	0.95	0.95	0.95	176

6.1.1 Experiment 1: 70% Training, 30% Testing

6.1.2 Experiment 2: 80% Training, 10% Validation, 10% Testing
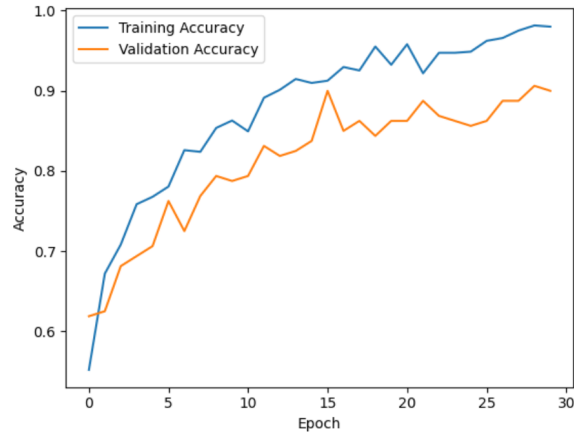

Figure 5: Accuracy of Trained CNN Model

6.1.3 Experiment 3: 70% Training, 15% Validation, 15% Testing

Table 3: Classification Report (Experiment 3)

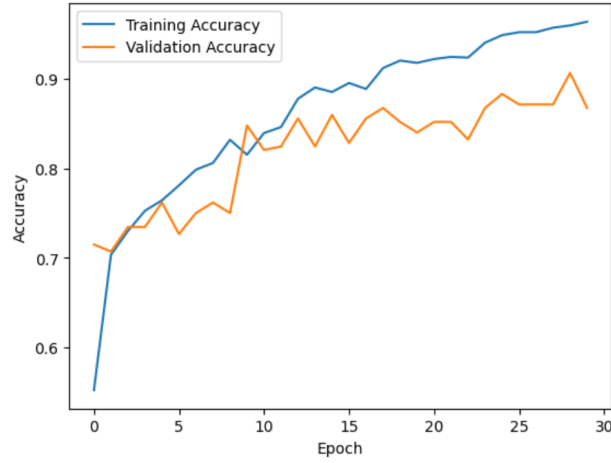|	Precision	Recall	F1-score	Support
Healthy	0.90	0.88	0.89	132
Hearing Impaired	0.88	0.90	0.89	132
Accuracy			0.89	264
Macro Avg	0.89	0.89	0.89	264
Weighted Avg	0.89	0.89	0.89	264

Figure 6: Accuracy of Trained CNN Model

6.2 Experiments Summary

Table 4: Experiment Results

Experiment	Training Data	Validation Data	Testing Data	Test Loss	Test Accuracy
1	70%	–	30%	0.379	0.865
2	80%	10%	10%	0.227	0.950
3	70%	15%	15%	0.470	0.891

7 Conclusions

The CNN architecture employed in this study achieves a satisfactory performance of approximately 85%-95% accuracy in distinguishing between individ-

uals with hearing impairment and healthy individuals based on scalp maps derived from spatial and temporal EEG data.

8   Future Work

For future investigations, several important steps can be taken to further enhance the model's accuracy and generalization. Firstly, applying this model on a different and larger dataset would be beneficial to assess its performance in a more diverse population. A larger dataset can help validate the model's effectiveness and robustness across a broader range of individuals.

Moreover, exploring alternative machine learning methods beyond CNNs, such as support vector machines (SVM), random forests, or recurrent neural networks (RNNs), could offer valuable insights and potentially lead to even higher accuracies. Comparing the performance of different machine learning algorithms can provide a better understanding of which approach is most suitable for this specific classification task.

Additionally, fine-tuning the model parameters and conducting necessary optimizations can significantly improve its performance. Techniques like cross-validation, hyperparameter tuning, and regularization can be employed to fine-tune the model, ensuring it is well-optimized and less prone to overfitting.

9   References

Søren A. Fuglsang, Jonatan Märcher-Rørsted, Torsten Dau, & Jens Hjortkjær. (2020). Selective auditory attention in normal-hearing and hearing-impaired listeners [Data set]. Zenodo. `https://doi.org/10.5281/zenodo.3618205`

10   Acknowledgments