CNN for Topographical Scalp Maps for Hearing Impairment Detection

Grace Wang
College of Engineering & Applied Sciences
Stony Brook University
Stony Brook, USA
grace.wang.2@stonybrook.edu

Beiyu Lin
Department of Computer Science
University of Nevada, Las Vegas
Las Vegas, USA
beiyu.lin@unlv.edu

Joel Snyder
Department of Psychology
University of Nevada, Las Vegas
Las Vegas, USA
joel.snyder@unlv.edu

Brendan Morris
Department of Electrical and Computer Engineering
University of Nevada, Las Vegas
Las Vegas, USA
brendan.morris@unlv.edu

***Abstract*—Age-related sensorineural hearing loss affects more than 50% of individuals aged 65 and older, resulting in a gradual decline in sound perception, particularly high-pitched tones. Often undetected until advanced stages, early intervention becomes crucial. Unlike prior research that primarily emphasizes the temporal aspects of electroencephalography (EEG) data, our approach utilizes EEG data to create scalp maps (topographic maps). These maps visualize the spatial distribution of electrical brain activity across the scalp, offering unique insights into the brain's topographical patterns. They enable a convolutional neural network (CNN)-based classification approach to hearing loss in individuals which achieves a 86%-91% accuracy rate in distinguishing seniors with hearing loss from healthy counterparts among 44 subjects. Through the application of image mining techniques on sequential data, our results not only enable researchers to pinpoint specific areas of heightened or diminished neural activity but also open up new avenues for early health concern detection, ultimately enhancing healthcare interventions and outcomes.**

I. Introduction

Sensorinueural hearing loss affects nearly a quarter (25%) of individuals aged 65 to 74, and it increases to half (50%) for those aged 75 and older. This hearing loss becomes more important as the population ages because people will have more trouble communicating and having conversations with others without the assistance of hearing aids or assistive devices. People may also struggle to comprehend medical instructions, respond to alerts, and hear doorbells and alarms, all which can risk both their personal well being and the safety of others [1]. The gradual deterioration in sound perception is often detected too late, emphasizing the need for early intervention. However, the landscape of hearing loss detection is marked by several obstacles, encompassing limited awareness, restricted access to routine screening, and disparities in specialized healthcare services. These challenges lead to inequalities in detection, impacting under-served communities, rural residents, individuals with lower socioeconomic status, and marginalized populations. This situation is further complicated in cases where communication is inherently challenging, such as among newborn infants and those with developmental conditions like autism or Down syndrome.

Current diagnostic methods, including otoacoustic emission testing, magnetic resonance imaging (MRI), and computerized tomography (CT) scans, have merits but also limitations. For instance, otoacoustic emission testing provides insights into cochlear functionality but may not fully capture complex neural responses [2]. MRI and CT scans offer detailed anatomical information but are resource-intensive and may not reveal functional nuances as effectively as EEG [3]. These limitations could impact the overall diagnostic accuracy, potentially leading to the oversight of subtle yet significant neural anomalies that contribute to hearing impairments.

In contrast, electroencephalography (EEG) is promising due to its affordability and non-invasiveness and has been a valuable tool for studying brain activities, cognitive processes, and neurological disorders. EEG's affordability stems from its relatively low equipment costs and minimal infrastructure requirements. Its non-invasiveness is achieved by placing electrodes on the scalp to record brain activity without the need for surgical procedures or injections, ensuring safety and comfort for patients during assessments [4]. EEG also has the capacity to rapidly generate substantial amounts of data. During EEG recordings, time-series data is captured, reflecting the brain's electrical activity over time. This data can be subjected to processing and transformation into various formats, such as spectrograms and event-related potentials (ERPs), which effectively serve to help analyze and study specific electrical brain responses. [5].

In a separate study, researchers employed an innovative approach utilizing an Auditory Evoked Potentials (AEP) dataset, which was derived from EEG data. This dataset was used in tandem with an improved-VGG16 CNN, resulting in a remarkable accuracy of 96.87% in identifying hearing deficiencies from AEP brain signals [6]. This study primarily

relied on CWT-based time-frequency images, which tracks the evolution of frequency components within the signal over time. While temporal data like the CWT-based time-frequency images provide insights into when specific events occur and how they evolve, scalp maps provide information about where these events are localized on the scalp. Also, when scalp maps are generated in a sliding window fashion, they have the potential to visually illustrate the evolution of scalp power over time similar to temporal data.

The novelty of this study lies in using EEG-derived scalp maps to classify hearing-impaired individuals from healthy ones. This approach is unique, as no other studies we are aware of related to hearing impairment have employed the scalp map feature of EEG data. Scalp maps not only enable researchers to pinpoint the precise areas of neural activity change in the brain but also provide temporal-spatial information on how neural responses evolve during different stages of information processing. This can help identify the brain areas associated with hearing impairment and help researchers in uncovering any potential underlying patterns. A deep learning model based on CNNs is employed due to its image-based nature, ideally suited for the visual nature of scalp maps. Through our CNN model experiments and data partitioning, accurate classification between hearing impairment and healthy subgroups reaches an average of 88% among a total number of 44 subjects.

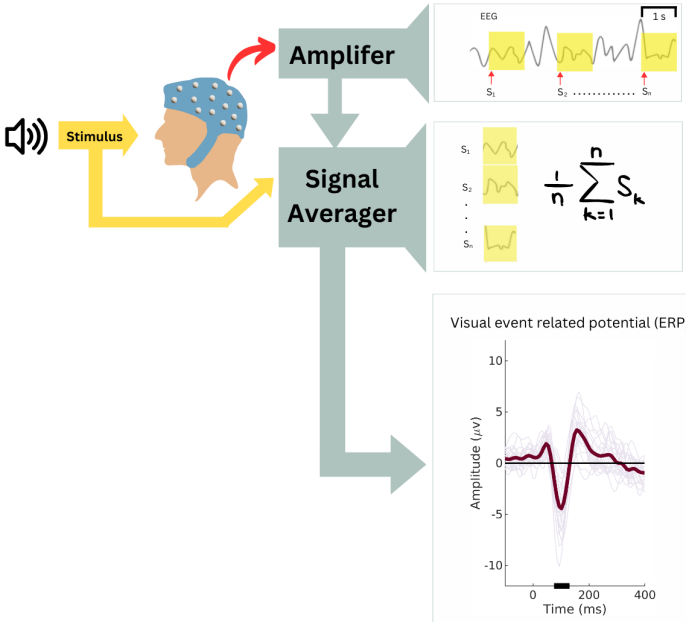

Fig. 1: Visual representation of how ERPs are derived from EEG data.

II. Methods

In this study, we use an open-source sensorineural hearing loss (SNHL) dataset that contains EEG data of 22 hearing impaired (HI) listeners with sensorineural hearing loss and 22 age matched normal hearing (NH) peers (between 51 - 76 years old) [7].

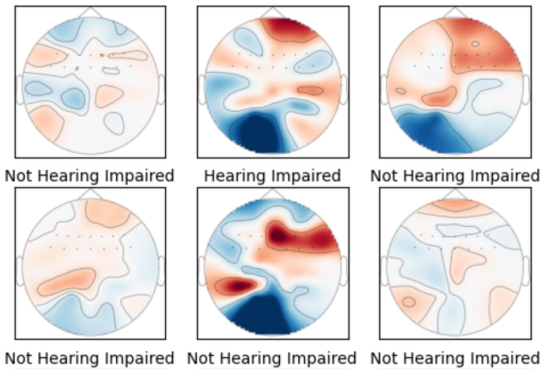

Fig. 2: Examples of scalp maps

This study consists of 4 main parts, including EEG signal processing, ERP extraction, extracting topographic brain maps, and training, validating, and testing the model.

A. EEG Preprocessing

We specifically look 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.

EEG preprocessing, which involves data cleaning and enhancement steps to remove noise, artifacts, and other sources of interference from the raw EEG signals 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 on GitLab[1].

B. ERP Extraction

We chose Event-Related Potentials (ERPs) as our EEG feature because they offer precise insights into both the brain's response to specific events or stimuli through waveform patterns and the timing of these responses due to their time-locked association with these events.

ERPs are obtained by averaging EEG recordings collected during repeated stimulus presentations, where EEG signals are initially weak and are amplified to enhance detectability. Multiple stimulus trials are recorded, segmented into epochs for baseline correction and analysis. Averaging these EEG traces across trials enhances ERP components and reduces random noise, providing a clearer representation of synchronized brain activity in response to the stimulus. An illustration of the process of extracting ERPs is depicted in Figure 1.

[1]https://gitlab.com/sfugl/snhl

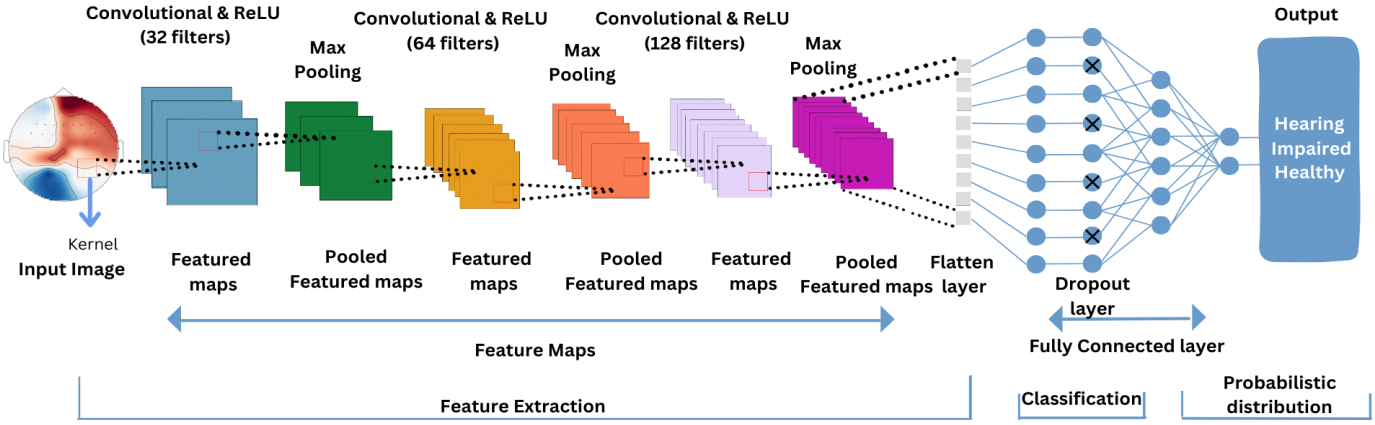

Fig. 3: Visual representation of CNN model

C. 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, a reasonable and rounded choice given that the smallest interval we could define was 8 ms. Scalp 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.

This approach of generating scalp maps resulted in the creation of 40 scalp maps for each subject (Fig. 2). Ultimately, 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.

D. Convolutional Neural Network (CNN) Architecture

We employed a CNN model to perform image classification on the dataset (Fig. 3). 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.

III. Results

A. Experiments

Our dataset consists of 1,760 images extracted from SNHL which are balanced between healthy and hearing impaired. We tested three different ways of splitting the data (into training, validation and testing data) to see how partitioning the data influences the CNN's classification accuracy and effectiveness in distinguishing between the two groups.

The approach used to construct the training, validation, and testing datasets involved the random shuffling of image files, followed by their randomized allocation across the training, testing, and validation subsets. As a result of these experiments, we obtained varying test precision, recall, F1 score and accuracy.

Precision measures the accuracy of positive predictions, indicating how often the model's positive predictions are correct. High precision suggests accurate positive predictions. Recall evaluates the model's completeness, showing how effectively it identifies actual positive instances. High recall means that the model is good at capturing most of the positive instances in the data. The F1-score balances precision and recall, offering a single performance measure. Support represents the number of images within each class being used in the dataset.

TABLE I: 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

1) Experiment 1 (30 Epochs): Experiment 1 involved training the model with 70% of the data and testing it on the remaining 30%. The classification report in Table I represents the performance of the model on the testing dataset.

From the classification report, we can see that the model achieved an accuracy of 86% on the testing dataset. Both classes, Healthy and Hearing Impaired, have identical preci-

sion, recall, and F1-score values of 0.86, indicating that the model's performance is well-balanced for both classes.

TABLE II: Classification Report (Experiment 2)

	Precision	Recall	F1-score	Support
Healthy	0.93	0.90	0.91	88
Hearing Impaired	0.90	0.93	0.92	88
Accuracy			0.91	176

2) Experiment 2 (100 Epochs): Experiment 2 involved training the model with 80% of the data, using 10% for validation, and the remaining 10% for testing. The classification report in Table II presents the performance of the model on the testing dataset. From the classification report, we can observe that the model achieved a test accuracy of 91%. The precision, recall, and F1-score values for both classes, Healthy and Hearing Impaired, are quite high, indicating a strong overall performance.

TABLE III: Classification Report (Experiment 3)

	Precision	Recall	F1-score	Support
Healthy	0.91	0.88	0.89	132
Hearing Impaired	0.88	0.91	0.90	132
Accuracy			0.89	264

3) Experiment 3 (100 Epochs): Experiment 3 involved training the model using 70% of the data for training, allocating 15% for validation, and using the remaining 15% for testing. The classification report in Table III presents the model's performance on the testing dataset.

The model achieved a test accuracy of 89%. Both classes, Healthy and Hearing Impaired, exhibit similar precision, recall, and F1-score values of around 0.89, indicating a well-balanced performance. The accuracy and F1-score are indicative of the model's ability to effectively classify instances from both classes.

Figure 4 displays the accuracy graph for Experiment 3. Experiment 2 exhibits a similar trend, so we are displaying one graph for brevity. We observe that the accuracy of our model stabilizes after approximately 30 epochs of training. During this period, the training accuracy reaches around 100%, indicating that the model has effectively learned the nuances of the training dataset. On the other hand, the validation accuracy plateaus at around 90%. This suggests that the model generalizes well to new, unseen data, performing at a consistently high level beyond its training data.

B. Experiments Summary

In Table IV, we present the results of all our experiments. Train means the ratio of data used for the training; Validation means the ratio of data used for the validation; Test means the ratio of data used for the test. We also include the test loss and test accuracy for each experiment.

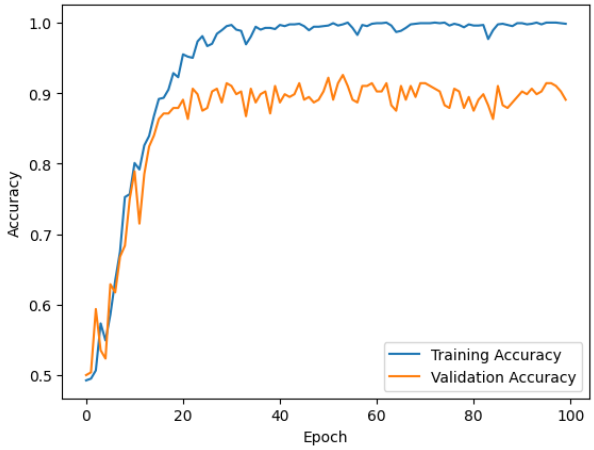

Fig. 4: Accuracy comparison for Experiment 3

The classification of hearing loss improved with the inclusion of the validation set. The results were quite close for Experiment 2 and 3 with Experiment 2 having the edge in performance perhaps due to combination of slightly larger training dataset and the smaller test set. Despite the lack of validation, Experiment 1 still had strong numbers which may be more indicative of real performance since the test set was at least twice the size of the other experiments.

These results highlight the potential for the use of CNN-based image classification on scalp maps generated from ERPs of EEG signals. Further research is needed with more participants to better characterize the generalization ability of this approach. K-fold cross-validation could be used due to the smaller dataset size or better having a larger and more diverse population since CNN classifications techniques require large volumes of images. Instead of shuffling images to training, validation and testing subsets randomly, we could partition data by individual subjects' scalp maps to ensure complete independence of subsets. Finally, the CNN architecture could be optimized by fine-tuning its hyperparameters.

TABLE IV: Experiment Results

Experiment	**Train**	**Validation**	**Test**	**Test Loss**	**Test Accuracy**
1	70%	–	30%	0.38	0.86
2	80%	10%	10%	0.57	0.91
3	70%	15%	15%	0.62	0.89

IV. Conclusions

The CNN architecture employed in this study achieves a performance of approximately 86%-91% accuracy in distinguishing between individuals with hearing impairment and healthy individuals based on scalp maps derived from spatial and temporal EEG data.

V. Acknowledgments

This work was supported by the National Science Foundation under Research Grant Number 1950872 at the University of Nevada, Las Vegas.

REFERENCES

[1] National Institute on Aging (NIA) U.S. Department of Health & Human Services (HHS), National Institutes of Health (NIH). Hearing loss: A common problem for older adults, 2023. Accessed on September 6 2023.

[2] Allen Young and Matthew Ng. Otoacoustic emissions. 2023.

[3] Mohamed Abdelazeem Hafez and Kirankumar Moholkar. Patient-specific instruments: advantages and pitfalls. *SICOT Journal*, 3:66, 2017.

[4] Luis Fernando Nicolas-Alonso and Jaime Gomez-Gil. Brain computer interfaces, a review. *Sensors (Basel)*, 12(2):1211–1279, 2012.

[5] Steven J Luck and Emily S Kappenman. Erp features and eeg dynamics: An ica perspective. In Steven J Luck and Emily S Kappenman, editors, *The Oxford Handbook of Event-Related Potential Components*, chapter 3, pages 51–76. Oxford University Press, 2012.

[6] Md Nahidul Islam, Norizam Sulaiman, Fahmid Al Farid, Jia Uddin, Salem A Alyami, Mamunur Rashid, Anwar P P Abdul Majeed, and Mohammad Ali Moni. Diagnosis of hearing deficiency using eeg based aep signals: Cwt and improved-vgg16 pipeline. *PeerJ Comput Sci*, 7:e638, Sep 2021.

[7] Søren A. Fuglsang et al. Effects of sensorineural hearing loss on cortical synchronization to competing speech during selective attention. *Journal of Neuroscience*, 40(12):2562–2572, 2020.