CNN for Topographical Scalp Maps for Hearing Impairment Detection

Grace Wang[1], Beiyu Lin[2], Brendan Morris[3], Joel Snyder[4]
grace.wang.2@stonybrook.edu, beiyu.lin@unlv.edu, brendan.morris@unlv.edu, joel.snyder@unlv.edu
[1]*College of Engineering & Applied Sciences, Stony Brook University, Stony Brook, USA*
[2]*Department of Computer Science, University of Nevada, Las Vegas, Las Vegas, USA*
[3]*Department of Electrical and Computer Engineering, University of Nevada, Las Vegas, Las Vegas, USA*
[4]*Department of Psychology, University of Nevada, Las Vegas, Las Vegas, USA*

***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 which offer insight into the brain's topographical patterns based on spatial distribution of electrical activity. A convolutional neural network (CNN) scalp map classifier was designed to distinguish seniors with hearing loss from healthy counterparts. Despite a lower accuracy rate of 55% among 44 subjects on the public SNHL dataset due to the temporal nature of our input data, our study reveals new insights into performance variations across time intervals and individual subjects. This understanding can guide us for further refinement and improvement of classification techniques in this domain.**

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 loss can hinder communication, making it harder to interact and understand important information. As people age, difficulties in hearing can lead to problems with medical instructions, alerts, and even safety signals like doorbells [1]. Unfortunately, detecting hearing loss early isn't always easy due to limited awareness, lack of screening access, and disparities in healthcare services. These challenges particularly affect underserved communities, rural areas, and those with lower socioeconomic status [2]. Special circumstances, like communicating with newborns or individuals with developmental conditions, add further complexity to the issue [3].

Current diagnostic methods like otoacoustic emission testing, MRI, and CT scans have strengths but also drawbacks. For example, while otoacoustic emission testing reveals cochlear function, it may miss complex neural responses [4]. MRI and CT scans provide detailed anatomy but are resource-intensive and may not capture functional details as effectively as EEG [5], [6]. These limitations could affect diagnostic accuracy, potentially overlooking subtle yet important neural anomalies linked to hearing loss.

EEG stands out for its affordability and non-invasiveness, making it a valuable tool for studying brain activity, cognitive processes, and neurological disorders. Its low equipment costs and simple setup make it accessible, while its non-invasive nature, through placing electrodes on the scalp, ensures patient comfort and safety [7]. EEG captures time-series data reflecting brain electrical activity, which can be processed into formats like spectrograms and and Event-Related Potentials (ERPs), aiding in the analysis of specific brain responses [8].

In prior work, researchers tried to distinguish older adults with or without mild hearing loss. They used numerical features extracted from ERP data, representing subjects' responses to clear and noisy speech stimuli. A support vector machine (SVM) served as a classifier to explore the temporal dynamics and brain regions responsible for segregating groups based on speech clarity. Utilizing 5-fold cross-validation, they achieved a classification accuracy of 81.5% for whole-brain source waveform data in response to clear speech [9].

The novelty of this study lies in using EEG-derived temporal scalp maps to classify hearing-impaired individuals from healthy ones. Scalp maps not only enable researchers to pinpoint the precise areas of neural activity change in the brain but also provide spatio-temporal information on how neural responses evolve during different stages of information processing [10]. This can potentially help identify the brain areas associated with hearing impairment and help researchers in uncovering any potential underlying patterns. A deep learning model based on convolutional neural networks (CNNs) is employed due to success in other image classification problems [11].

II. Methodology

In this study, we use the open-source sensorineural hearing loss (SNHL) dataset [12] 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). Our approach includes EEG signal processing, ERP extraction, generation of topographic brain maps, followed by model training and testing. To the best of our knowledge, no other studies have employed this specific dataset in a similar manner.

A. EEG Topographic Scalp Map

We focused on a tone stimuli experiment where ERPs were recorded during passive listening to 1 kHz pure tones

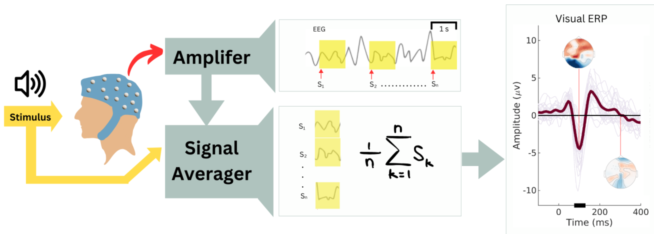

Fig. 1: Visual representation of how ERPs and scalp maps are derived from EEG data.

with random presentation fluctuations for 180 trials. EEG preprocessing, including noise removal and artifact correction, was largely completed by dataset researchers[1]. ERPs, averaged EEG responses to repeated stimuli, were chosen as key features for group classification due to their ability to reveal precise brain responses over time [13].

From each patient's ERP waveform, recorded up to 400 ms post-tone onset, we extracted topographic scalp maps at 10 ms intervals. This involves capturing the ERPs of each electrode at every time interval to create visual representations of scalp maps. These maps visualize brain electrical activity spatially, providing insight into activity magnitude and location over time. Employing a sliding window approach, 40 scalp maps per subject were generated, resulting in 1,760 images (44 subjects, 40 images each) for input into our CNN model. An illustration of the process of extracting ERPs and scalp maps is depicted in Fig. 1 with comparison of scalp maps for different subjects at different times in Fig. 2.

B. *Convolutional Neural Network (CNN) Architecture*

In this paper, we compare two CNN models: a lightweight CNN from scratch to directly learn topographic features and a pre-trained ResNet50 backbone for more general features learned from a large corpus of natural images.

We designed a custom lightweight CNN model to perform image classification on the dataset from scratch (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. These convolutional layers had filter sizes of 32, 64, and 128, respectively. ReLU activation functions were applied after each convolutional operation to introduce non-linearity. To prevent overfitting, weight decay regularization with a regularization strength of 0.0001 was incorporated into 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.

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

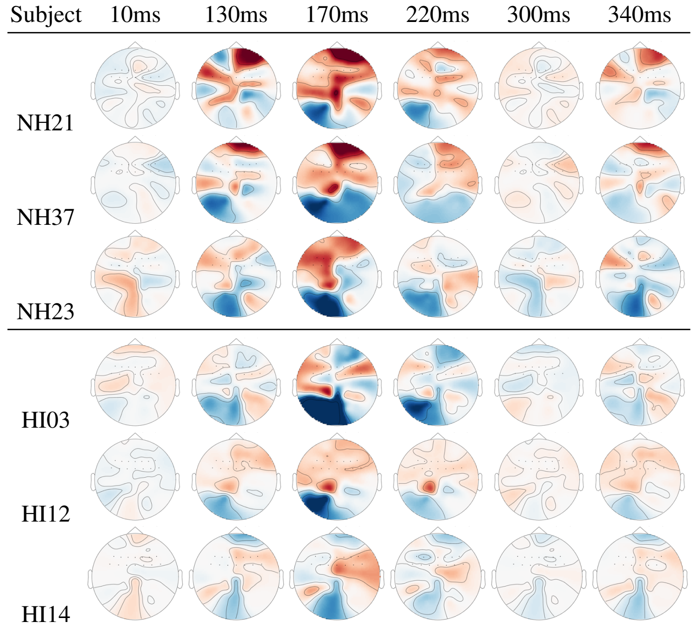

Fig. 2: Examples of topographical scalp map evolution over time. There is some structure such as the red in the top right and blue in bottom left in at 170ms. Yet, there are clear differences over time for each subject.

We also used an ImageNet pre-trained ResNet50 model, where only the convolutional layers were frozen to preserve the learned features from the ImageNet dataset [14]. Custom classification layers were then added atop this pre-trained ResNet50 base model for fine-tuning. These additional layers included a flattening layer, a fully connected dense layer with 512 units and ReLU activation, a dropout layer to mitigate overfitting, and a final dense layer with softmax activation.

Both models were compiled using the Adam optimzier, a learning rate of 0.001 and cateogrical cross-entropy loss function.

III. Results

A. *Experiments*

Our dataset consists of 1,760 images extracted from SNHL, which are balanced between healthy and hearing impaired classes. Due to the limited size of the dataset, we utilize a 10-fold cross validation evaluation to characterize the performance of both CNN classifiers. Each fold includes a random selection of four hearing-impaired (HI) and four normal hearing (NH) subjects for testing and the remaining subjects were used for training, resulting in an 80/20 split between training and testing data. To address the limited size of the dataset the training set underwent data augmentation to enhance diversity. We applied slight rotation, height and width shift, zooming, and flipping techniques. Training was conducted over 100 epochs to ensure thorough learning and adaptation.

To better understand our results, we categorized our analysis into three experiments. First we evaluated overall classification accuracy using 10-fold cross-validation. Next, we examined

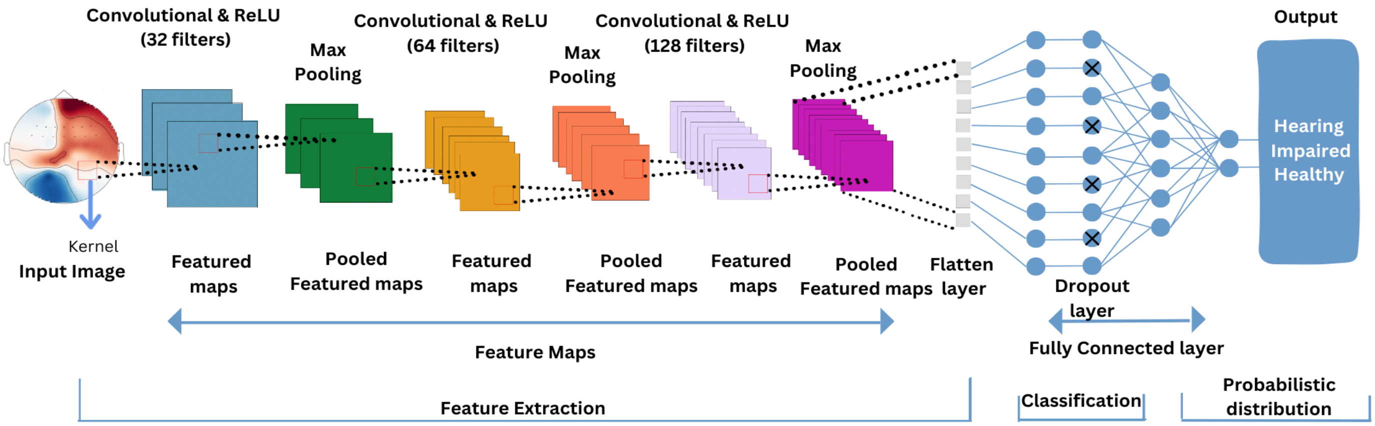

Fig. 3: Visual representation of CNN model

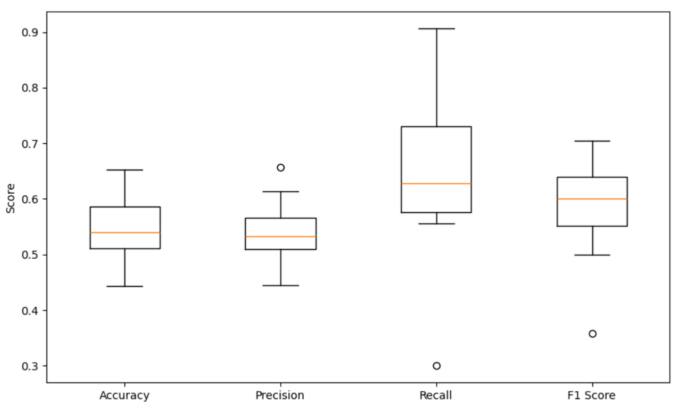

Fig. 4: Performance metrics boxplot across 10 folds (custom CNN)

TABLE I: Average 10-Fold Performance

Model	Accuracy	Precision	Recall	F1 Score
Custom	**0.55**	**0.54**	**0.65**	**0.58**
ResNet50	0.50	0.50	0.50	0.49
ResNet50+conv	0.50	0.49	0.50	0.44

classification accuracy over time. Last, we observed classification by individual subject.

1) Experiment 1: We evaluated the overall performance using 10-fold cross-validation to characterize classification performance. This approach ensures a generalized view rather than biased by the small size and specific individuals selected for the test set. Performance is measured with classification accuracy, precision, recall, and the F1-score (higher is better for all). The custom CNN results are visualized with the boxplots, in Fig. 4, and show the performance was fairly consistent across folds. Recall had the highest variability.

Table I compares the lightweight custom CNN to the ResNet pretrained model (transfer learning for classification layers) and ResNet+conv which fine tunes the last four convolutional layers to have more scalp map specific features. Fine-tuning the convolutional layers did not provide performance increase

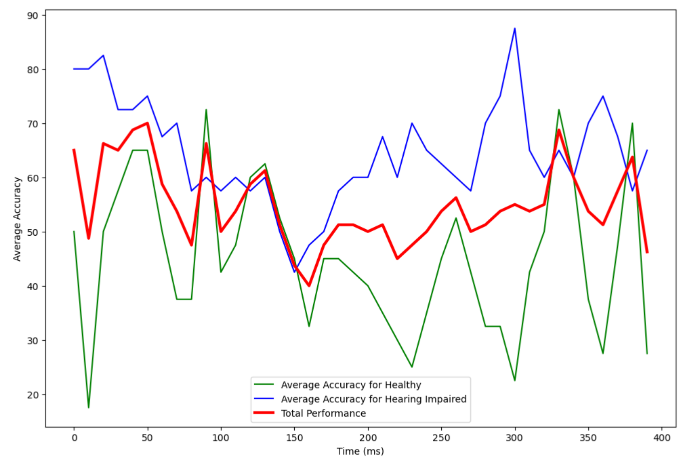

Fig. 5: Average Accuracy for Healthy and Hearing Impaired and Total Performance Over Time (custom CNN)

TABLE II: Average Accuracy % for Healthy and Hearing Impaired Subjects

	Custom	ResNet50	ResNet50+conv
Healthy	45.12	47.75	39.52
Impaired	64.50	51.88	60.38
Total	54.81	49.81	49.95

and both ResNet variants performed rather poorly. Overall, the custom network performed slightly better with a balanced F1 score almost 10% higher.

2) Experiment 2: Since the CNN classifier does not explicitly model temporal evolution of the EEG signals, the classification is assessed over time. Fig. 5 gives the average accuracy for each time slice with normal hearing in green, hearing impaired in blue, and total performance in red.

At certain time points, there's a significant difference in the average accuracy between each class, and the total performance shows low accuracy. However, at other points, the difference is smaller, and the total performance exhibits higher accuracy. For example, there are large accuracy differences

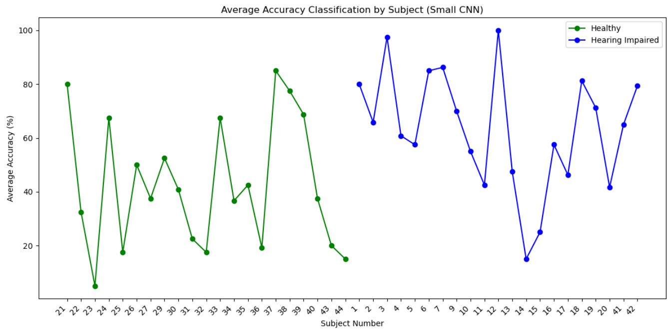

Fig. 6: Average accuracies in classification by subject (custom CNN)

between the two classes around 10, 300 and 360 ms compared to other time points (Fig. 5). This may be due to lack of consistent structure as seen in the scalp maps at 10 and 300 ms as compared to the more consistent patterns seen at 130, 170, and 340 ms (blue on the bottom left, red upper right in Fig. 2). Finally, despite the overall low accuracy rates, the CNNs do better at classifying the hearing impaired subjects (Table II).

3) Experiment 3: The final experiment focuses on individual subject classification. Instead of evaluating the average accuracy across all subjects by time, this method assesses the model's performance on each subject independently. The process involves gathering accuracy results for classifying images of each test subject across all folds and averaging them. The goal is to uncover insights into whether specific subjects present more challenges in accurate classification than others. Notably, NH subjects 23 and 44 (blue) and HI 14 and 41 (red) emerge when compared to other subjects (Figure 6) due to lower accuracies. This can also be reflected in these subject's scalp map, such as NH subject 23's scalp map at 10ms, appearing more enhanced in red as compared to other subjects.

B. *Experiments Summary*

Our experiments reveal that despite low accuracy results from the CNN model across the folds, analyzing accuracy over time reveals instances where the model encounters challenges. Certain time intervals demonstrate better performance in distinguishing between classes than others, perhaps suggesting that some times may not be optimal for classification while others may hold more value. These insights are further supported by the visual differences observed in the scalp maps corresponding to different time intervals.

The performance variation by subject and time provide motivation to explore stronger and more complex architectures such as recurrent neural networks to consider temporal variation of scalp features. The results presented provide image classification accuracy rather than a single label per individual and with explicit temporal modeling, the performance can be better compared to existing literature. Expanded datasets are also required for more complete training of larger network models.

IV. Conclusions

In this paper, we proposed a deep learning approach for hearing impairment classification using a sliding window scalp maps approach. Specifically, we explored using a temporal sequence of scalp maps derived from ERPs as input features. Through 10-fold cross-validation, we evaluated the model's performance across various folds, subjects, and time intervals. The model achieved a 55% overall accuracy rate in distinguishing between healthy and hearing-impaired subjects suggesting plenty of room for future improvement.

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] Matthew L Bush, Miriam R Kaufman, and Brian N McNulty. Disparities in access to pediatric hearing health care. *Curr Opin Otolaryngol Head Neck Surg*, 25(5):359–364, Oct 2017.

[3] Saral Mehra, Roland D Eavey, and Donald G Keamy. The epidemiology of hearing impairment in the united states: Newborns, children, and adolescents. *Otolaryngology–Head and Neck Surgery*, 140(4):461–472, 2009.

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

[5] T. Heye, R. Knoerl, T. Wehrle, D. Mangold, A. Cerminara, M. Loser, M. Plumeyer, M. Degen, R. Lüthy, D. Brodbeck, and E. Merkle. The energy consumption of radiology: Energy- and cost-saving opportunities for ct and mri operation. *Radiology*, 295(3):593–605, 2020.

[6] Abdullah Almohammadi and Yu-Kai Wang. Revealing brain connectivity: graph embeddings for eeg representation learning and comparative analysis of structural and functional connectivity. *Frontiers in Neuroscience*, 17, 2024.

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

[8] 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.

[9] MS Mahmud, F Ahmed, R Al-Fahad, KA Moinuddin, M Yeasin, C Alain, and GM Bidelman. Decoding hearing-related changes in older adults' spatiotemporal neural processing of speech using machine learning. *Front Neurosci*, 14:748, Jul 2020.

[10] V. Janiukstyte, T.W. Owen, U.J. Chaudhary, B. Diehl, L. Lemieux, J.S. Duncan, J. de Tisi, Y. Wang, and P.N. Taylor. Normative brain mapping using scalp eeg and potential clinical application. *Sci Rep*, 13(1):13442, Aug 2023.

[11] Waseem Rawat and Zenghui Wang. Deep convolutional neural networks for image classification: A comprehensive review. *Neural Computation*, 29(9):2352–2449, 2017.

[12] 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.

[13] Geoffrey F Woodman. A brief introduction to the use of event-related potentials in studies of perception and attention. *Attention, Perception, & Psychophysics*, 72(8):2031–2046, 2010.

[14] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778, 2016.