diagnostics

MDPI

Article

Auditory Brainstem Response Data Preprocessing Method for the Automatic Classification of Hearing Loss Patients

Jun Ma [1], Jae-Hyun Seo [2], Il Joon Moon [3], Moo Kyun Park [4], Jong Bin Lee [5], Hantai Kim [5], Joong Ho Ahn [6], Jeong Hun Jang [7], Jong Dae Lee [8], Seong Jun Choi [9] and Min Hong [10,*]

1 Department of Software Convergence, Soonchunhyang University, Asan 31538, Republic of Korea; ringring369@gmail.com
2 Department of Otorhinolaryngology-Head and Neck Surgery, Seoul St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul 06591, Republic of Korea; revivalseo@catholic.ac.kr
3 Department of Otorhinolaryngology-Head and Neck Surgery, Sungkyunkwan University School of Medicine, Samsung Medical Center, Seoul 06351, Republic of Korea; mooniljoon@gmail.com
4 Department of Otorhinolaryngology-Head and Neck Surgery, Seoul National University Hospital, Seoul National University College of Medicine, Seoul 03080, Republic of Korea; aseptic@snu.ac.kr
5 Department of Otorhinolaryngology-Head and Neck Surgery, Konyang University College of Medicine, Daejeon 35365, Republic of Korea; rogue25@kyuh.ac.kr (J.B.L.); noto.hantai@gmail.com (H.K.)
6 Department of Otorhinolaryngology-Head and Neck Surgery, Asan Medical Center, University of Ulsan College of Medicine, Seoul 05505, Republic of Korea; ucanhear@gmail.com
7 Department of Otolaryngology, Ajou University School of Medicine, Suwon 16499, Republic of Korea; jhj@ajou.ac.kr
8 Department of Otorhinolaryngology-Head and Neck Surgery, Soonchunhyang University College of Medicine, Bucheon Hospital, Bucheon 14584, Republic of Korea; ljdent@schmc.ac.kr
9 Department of Otorhinolaryngology-Head and Neck Surgery, College of Medicine, Soonchunhyang University, Cheonan Hospital, Cheonan 31151, Republic of Korea; akas9238@hanmail.net
10 Department of Computer Software Engineering, Soonchunhyang University, Asan 31538, Republic of Korea
* Correspondence: mhong@sch.ac.kr

check for updates

Citation: Ma, J.; Seo, J.-H.; Moon, I.J.; Park, M.K.; Lee, J.B.; Kim, H.; Ahn, J.H.; Jang, J.H.; Lee, J.D.; Choi, S.J.; et al. Auditory Brainstem Response Data Preprocessing Method for the Automatic Classification of Hearing Loss Patients. *Diagnostics* **2023**, *13*, 3538. https://doi.org/10.3390/diagnostics13233538

Academic Editor: Kathiravan Srinivasan

Received: 31 October 2023
Revised: 19 November 2023
Accepted: 21 November 2023
Published: 27 November 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

Abstract: Auditory brainstem response (ABR) is the response of the brain stem through the auditory nerve. The ABR test is a method of testing for loss of hearing through electrical signals. Basically, the test is conducted on patients such as the elderly, the disabled, and infants who have difficulty in communication. This test has the advantage of being able to determine the presence or absence of objective hearing loss by brain stem reactions only, without any communication. This paper proposes the image preprocessing process required to construct an efficient graph image data set for deep learning models using auditory brainstem response data. To improve the performance of the deep learning model, we standardized the ABR image data measured on various devices with different forms. In addition, we applied the VGG16 model, a CNN-based deep learning network model developed by a research team at the University of Oxford, using preprocessed ABR data to classify the presence or absence of hearing loss and analyzed the accuracy of the proposed method. This experimental test was performed using 10,000 preprocessed data, and the model was tested with various weights to verify classification learning. Based on the learning results, we believe it is possible to help set the criteria for preprocessing and the learning process in medical graph data, including ABR graph data.

Keywords: deep learning; VGG; ABR; image processing; hearing loss

1. Introduction

Hearing loss refers to a condition in which the ability to hear or understand sounds is more difficult than for people with normal hearing due to abnormalities in the inner, outer, or vestibulocochlear nerves. These symptoms may be temporary or permanent and may affect one or both ears. Auditory brainstem response (ABR) refers to an electrical response that occurs when sound is recognized in the brainstem, which is the path through

which sound (audio) reaches the auditory cortex of the brain as a neuroelectric signal of the cochlear tube. The diagnostic process used to assess the hearing threshold by measuring the waveform of the corresponding electrical response is usually referred to as an ABR test.

The ABR test, generally performed to evaluate hearing loss, is a non-invasive test that is unaffected by sleep or anesthesia. In particular, the ABR test is a means of objectively measuring the hearing of newborns, infants, and patients with congenital disabilities who have difficulty accurately measuring hearing level, and due to the objectivity of non-invasive testing, plays an important role in tracking the hearing of children at risk of late-onset or progressive hearing loss, such as those affected by congenital CMV infection [1,2]. The hearing threshold of the pure tone audiometry test is the smallest sound that the subject can hear in each frequency band at (125 to 8000) Hz. However, the ABR test uses clicks, tone pips, or tone bursts about 0.8 ms apart to audibly stimulate the brain and records changes in brain waves due to energy induced by stimuli modulation during sound transmission. An auditory stimulation ranging from approximately 10 to 100 dBHL is given, and waves l to V are detected. In normal adults, a waveform that responds to the stimulus appears within approximately 10 ms of starting a click sound stimulus. In the case of using a click sound, the auditory threshold is 35 dBnHL or more, and in the case of using a tone pip and tone burst, the auditory threshold is 40 dBnHL or more, referred to as hearing loss [3–7]. Among them, in audiological investigations, recordings are made at different levels of stimulus intensity until a response (wave V) is no longer observed, and this is taken as an estimation of the hearing threshold [8].

The ABR test is a process by which small electrodes are attached to the forehead and behind the ears to detect electrical activity in the auditory nerve and brainstem in response to sound. The patient hears a series of clicking sounds through earphones inserted into the ears, and the brain waves, which are the brain's response to the sounds, are detected and automatically recorded on a computer. Therefore, the ABR test can be measured objectively, compared to test methods based on the patient's subjective response, such as pure tone audiometry. The audiologist judges the hearing threshold of the ABR test with waveform 5, which is detected between 6 and 8 ms, which is the most reliable among waveforms 1 to 5 from each sound stimulus (dB). The ABR test is completed after measuring the EEG for each (30 to 90) dB sound stimulus [9,10].

In general, the ABR test results can determine hearing thresholds, the presence of pseudohypacusis, and retro-cochlear hearing loss [11]. Despite its many advantages, the ABR test has several limitations. The biggest limitation is that there is a high possibility that an error can occur depending on the experience and skill level of the audiologist in the hearing threshold or the hearing result value due to the subjective judgment of the audiologist. In addition, the ABR threshold corresponds reasonably well with the average subjective hearing threshold across the frequency range from 2 kHz up to 4 kHz and does not reflect hearing loss in the low-frequency region [12,13]. As shown in Figure 1, the ABR test results are displayed on the screen and stored in various forms because each hospital uses measuring equipment from different manufacturers. Because of the different measuring devices, audiologists need experience in using the specific devices, and it is difficult to collect ABR data in the same format for deep learning.

These problems also occur in the proposed hearing loss classification AI model using ABR data. When the ABR result data of each manufacturer are directly applied to the AI model and trained without refining, low-accuracy prediction results are obtained. In this paper, to develop an AI learning model that automatically classifies hearing loss, the preprocessing process for efficiently learning ABR data is proposed. After applying the proposed preprocessing process, we trained the VGG16 model for data and the classification test for hearing loss patients.

The ABR data in this paper were obtained by gathering data from Soonchunhyang University Cheonan Hospital, Asan Medical Center, Samsung Medical Center, Ajou University Medical Center, Seoul National University Hospital, Severance Hospital, Seoul St. Mary's Hospital, and Konyang University Hospital. The ABR data of 5000 people with normal

hearing and 5000 patients with hearing loss were collected and preprocessed. Using the collected data, a deep learning model was applied to analyze the classification performance and accuracy of the deep learning model. This study was reviewed and approved by the Soonchunhyang University Cheonan Hospital's Investigational Review Board (IRB number: 2021-06-040). We classified ABR image data into hearing loss and normal hearing based on the ABR hearing threshold of 40 dB. The main contribution of this paper is that the proposed method is designed to automatically distinguish between hearing loss patients and normal hearing people using the VGG16 model. The proposed method is expected to be clinically usable with an accuracy of approximately 85%.

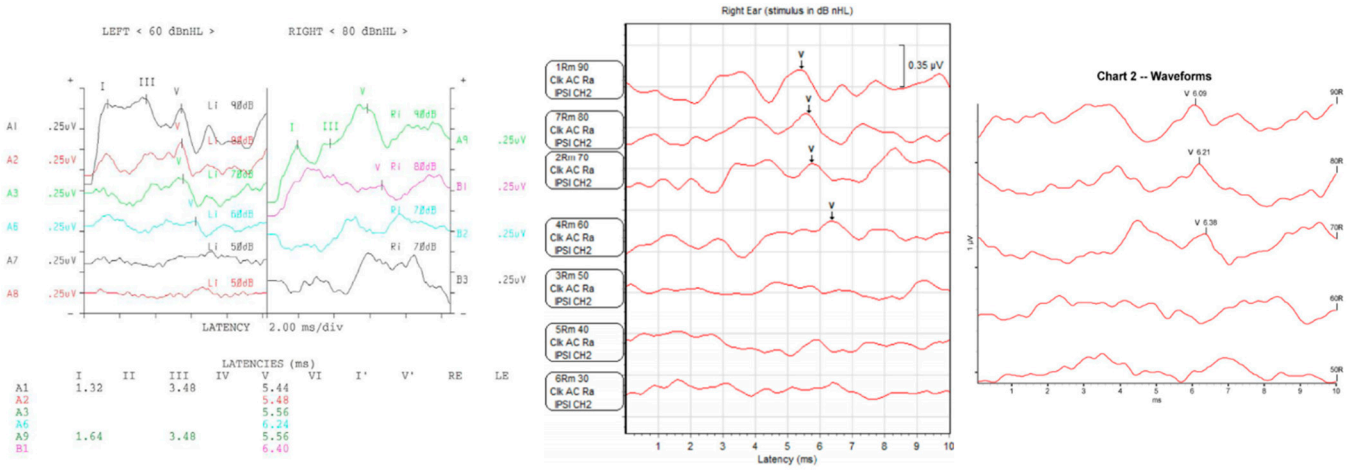


Figure 1. ABR test results of different forms from various manufacturers.

2. Materials and Methods

2.1. Medical Data for Deep Learning

Medical data is the most actively researched in the field of AI, especially in deep learning technology, and is one of the fields in which many automatic diagnosis systems are currently being developed. Recently, as various medical data have been digitized and a large amount of high-quality medical information has been generated, research on computer-based analysis and automatic diagnosis has been actively conducted [14–17]. In addition, research to apply traditional machine learning to artificial intelligence technologies is also being conducted. However, despite active research on image diagnosis through machine learning, the results have not been good. The developed algorithm had limitations for actual clinical application due to poor medical accuracy and the lack of robustness of various data. Since algorithms in classical image processing techniques or machine learning methods are based on the developer's intuition or experience, it is difficult to consider all the various situations, shapes, and transformations of data. Therefore, it is not easy to model complex and abstract transformations beyond the developer's knowledge. This is considered a limitation of machine learning methods in medical imaging and remains at the research level without leading to actual clinical application or commercialization.

With the advent of deep learning technology, this perception and situation have considerably changed. Deep learning is a concept designed based on an artificial neural network (ANN) that showed limitations due to the initial overfitting problem, but Jeffrey Hinton presented the concept of prior learning in 2006 and the concept of dropout in 2012, and it has been proven that it can overcome the existing problems [14,18–20]. In particular, among various deep learning fields, the convolutional neural network (CNN) performs highly in image recognition. Research has been conducted based on CNN in medical imaging fields, including various fields of vision [21–24]. As a representative example, preoperative medical image analysis can be used to determine the accuracy of the surgical site and the severity of the condition. An objective analysis of traumatic lesions [25] or

prescription of optimal nutrients using the patient's unique vital data [26,27] are utilized using AI.

2.2. Auditory Brainstem Responses

By attaching electrodes, the auditory brainstem response (ABR) test records electrical changes occurring in auditory conduction across the auditory nerve and brainstem after sound stimulation through the ear canal. Unlike pure tone audiometry (PTA), the ABR test is a non-invasive test that can objectively evaluate hearing threshold, is not affected by anesthesia or sleep, and is widely used in otolaryngology [28,29]. As shown in Figure 2, the ABR test determines whether the most reliable V waveform, related to hearing, is detected among the (I–V) waveforms transmitted from the brainstem. Based on dBHL, ISO 1964 [30], it is classified as normal if the V waveform is detected below 25 dBHL, mild hearing loss if the V waveform is detected in the (26 and 40) dBHL range, moderate hearing loss if the V waveform is detected in the (41 and 55) dBHL range, moderate–severe hearing loss if the V waveform is detected in the (56−70) dBHL range, and severe hearing loss if it is detected above 71 dBHL [31,32].

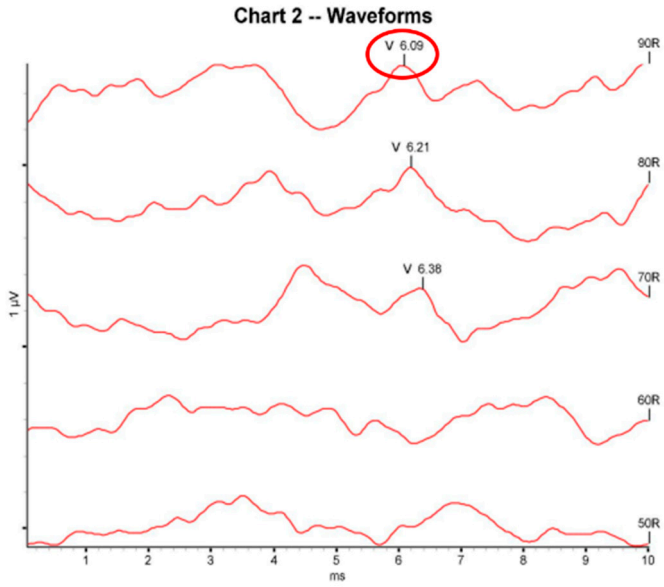
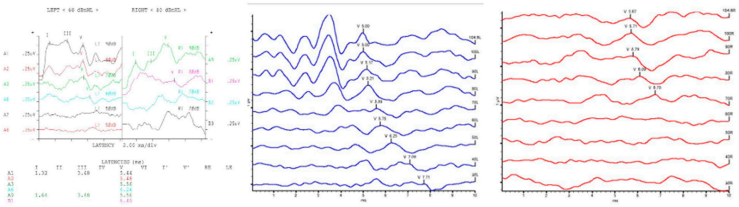
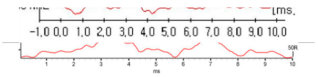
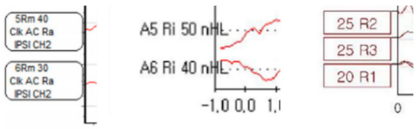


Figure 2. V waveform detection in an ABR result graph.

2.3. Data Processing and Cleansing

Hearing test results are stored in a medical information system, such as an electronic medical record (EMR), at each hospital. Patients' personal information is removed manually by using data labelers. Since ABR data is stored in graph form from ABR devices used in each hospital, they should be collected and preprocessed (as shown in Table 1). Accuracy of the data can be increased by deep learning via unifying the different data formats through data preprocessing, and improving the data quality.

Table 1. Examples of representative differences in ABR graphs obtained from each hospital.

Difference	Example of ABR Data Image
Different graph colors for each device	
Inconsistency of *X*-axis starting point	
Inconsistency of V waveform mark	
Inconsistency of the presence of graph grid	

2.4. ABR Data Preprocessing Process

Data Normalization and Preprocessing

Figure 3 is a flowchart of the preprocessing process performed to normalize the ABR data acquired from each hospital, and a detailed explanation of the proposed data preprocessing process follows.

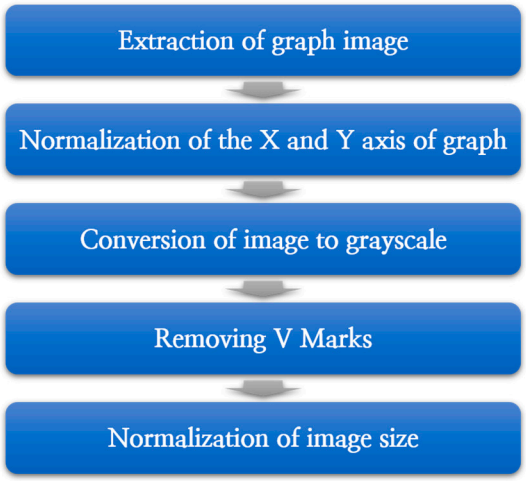


Figure 3. The proposed data preprocessing process.

1. Extraction of graph image

Figure 4 shows the process of extracting only the graph part necessary for image learning from the raw ABR data. In the collected ABR data, graph and table image data are stored together. We extracted only graph images to automatically determine whether there was hearing loss using graph images.

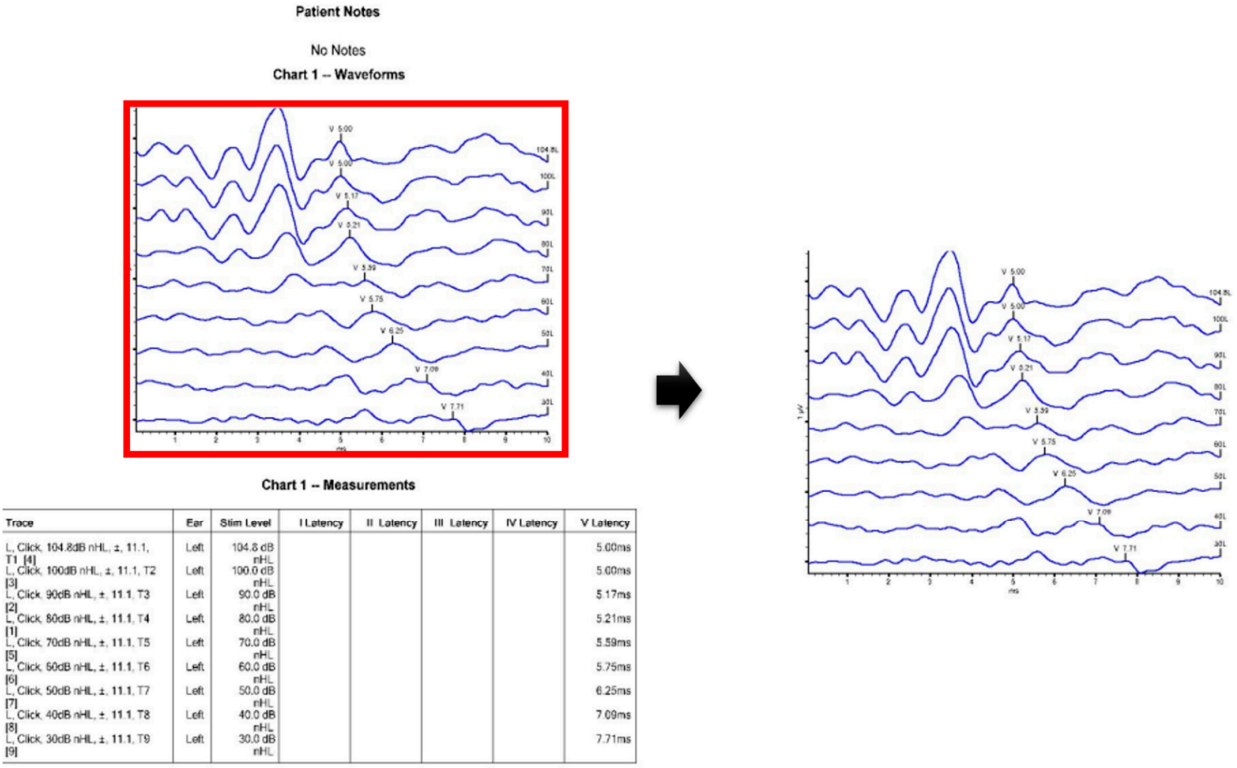


Figure 4. Extraction of graph image.

2. Normalization of the *X* and *Y* axes of the graph

Figure 5 shows the before and after normalization of the X and Y-axes for the ABR result data obtained from each hospital ABR device. Since the zero point standard is 0 or — value, and the scale standards of the axes are different from each other, which affects data learning and analysis, we proceeded to standardize the axes in the same format.

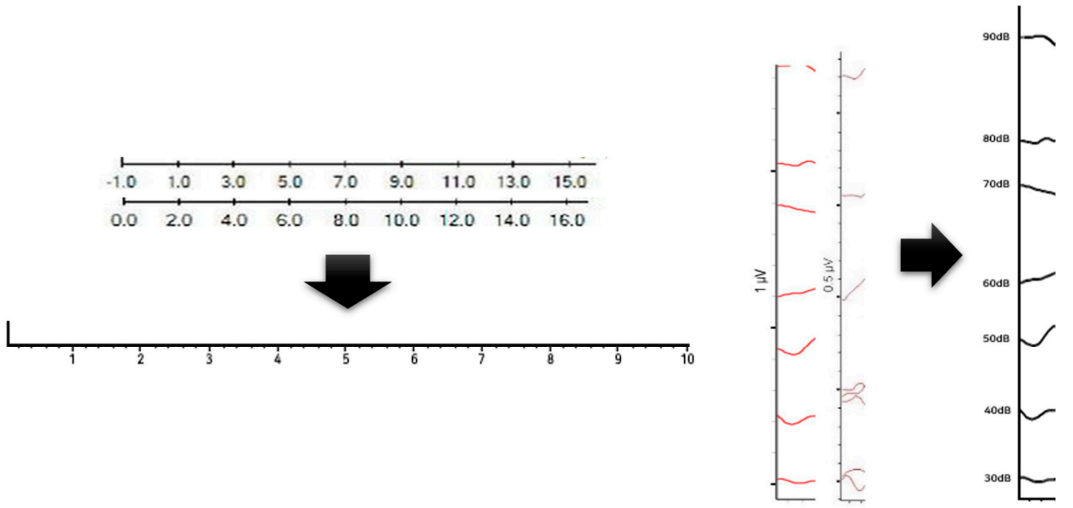


Figure 5. Normalization of the *X* and *Y* axes in a graph image.

3. Conversion of image to gray-scale

Since the ABR test basically measures both ears, the ABR device is displayed in a separate color to distinguish the left ear from the right ear. The ABR data in this paper is a model to verify the classification of simple hearing loss, and since there is no need to

distinguish between left and right ears, the image data is unified in gray-scale, as shown in Figure 6 below:

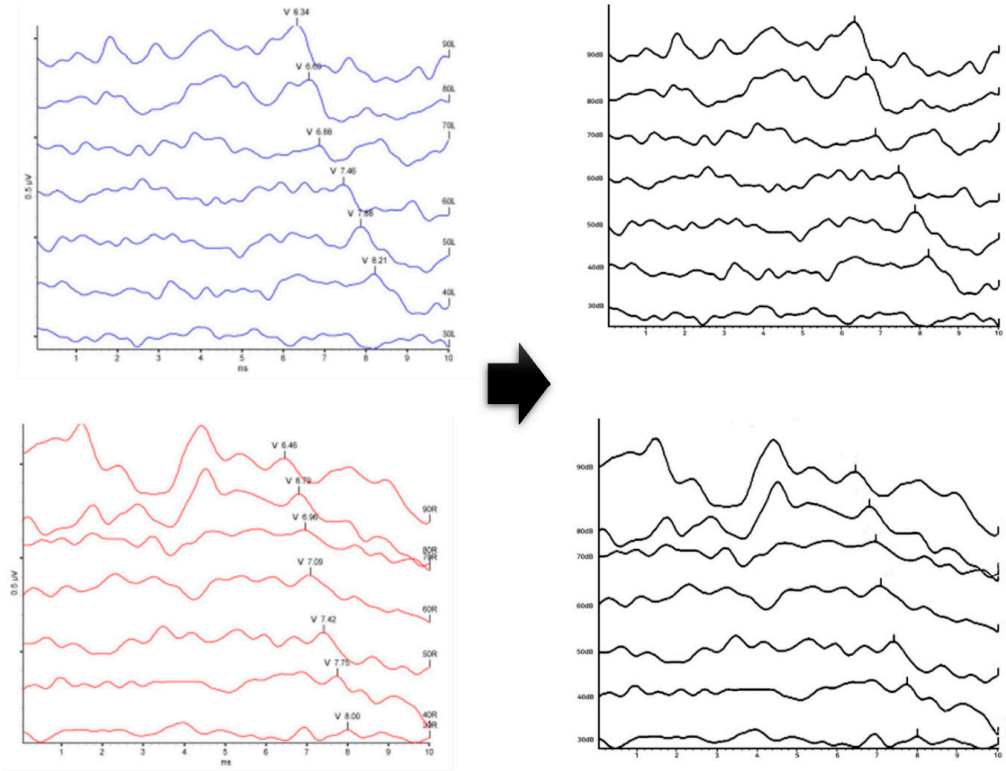


Figure 6. Image conversion to gray-scale.

4. Removing V Marks

In the obtained ABR images, the V waveform notation method is not uniform, such as some graph images having scale marks and images with a separate mark (v) for feature point marks. Therefore, as shown in Figure 7, we deleted the separate V waveform mark for each graph and left only the vertical bar shape at the corresponding V position.

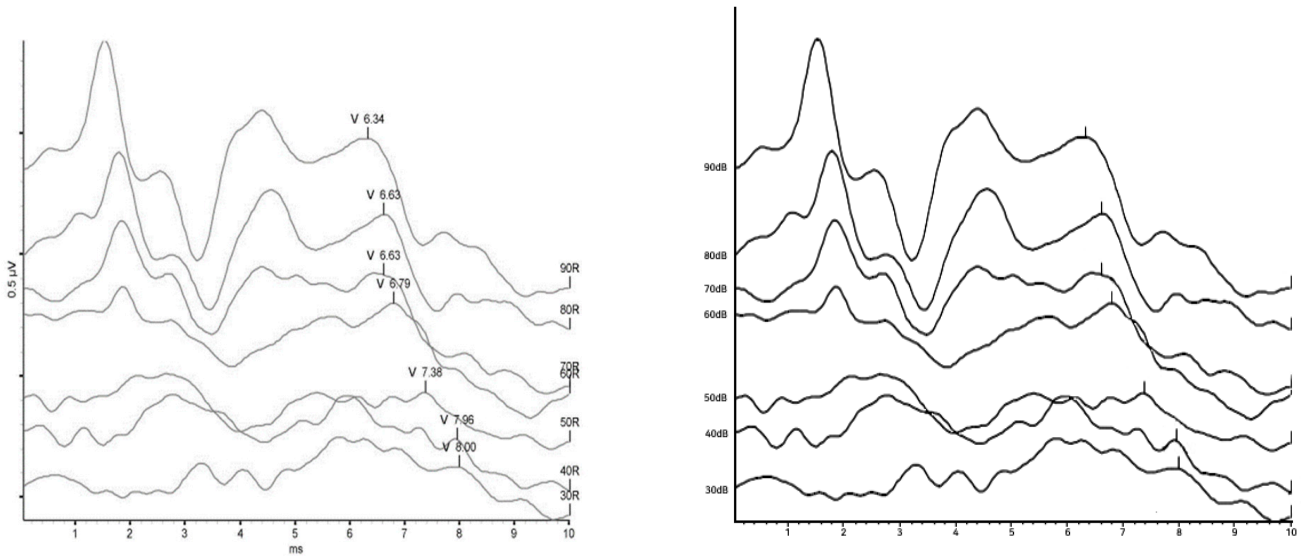


Figure 7. V Mark removal.

5. Normalization of image size

To smoothly process the deep learning for graph images, the image size was changed to 573 × 505 px for normalized and preprocessed images. Since *X* and *Y*-axis normalization and V-mark removal were already performed in the previous step, ABR image data loss for deep learning training was minimized.

In this paper, 10,000 ABR data from 8 different hospitals were collected, preprocessed, and normalized. In addition, a deep learning classification model was applied to the preprocessed ABR image data to classify patients with hearing loss automatically.

2.5. VGG 16 Model

VGG is a model proposed by the Visual Geometry Group of Oxford. There are two types of VGG structures: VGG16 and VGG19. VGG 19 has three convolution layers, one in front of the pooling layer in layers 3, 4, and 5, compared to VGG 16. For binary classification learning, the model in this paper adopted VGG 16 because too many convolution layers have a high probability of over-integration and may thus be less optimized [33–35]. Figure 8 shows the architecture of VGG16.

1. Input layer: Basically, an image of size 224 × 224 is delivered to the convolution layer;
2. Convolution layers: VGG16 has 13 convolution layers (conv 1–1~5–3). Each convolution layer consists of a small filter with a size of 3 × 3. Each filter is responsible for extracting features from the input image. After the convolutional layer, a rectified linear unit (ReLU) is used as the activation function;
3. Pooling Layers: After each convolution layer, the max pooling layer is applied. Maximum pooling is responsible for reducing space by extracting only the largest values from each area;
4. Dense Layers: These layers perform a final classification based on the features of the image. After each fully connected layer, the ReLU activation function is also used [36–38].

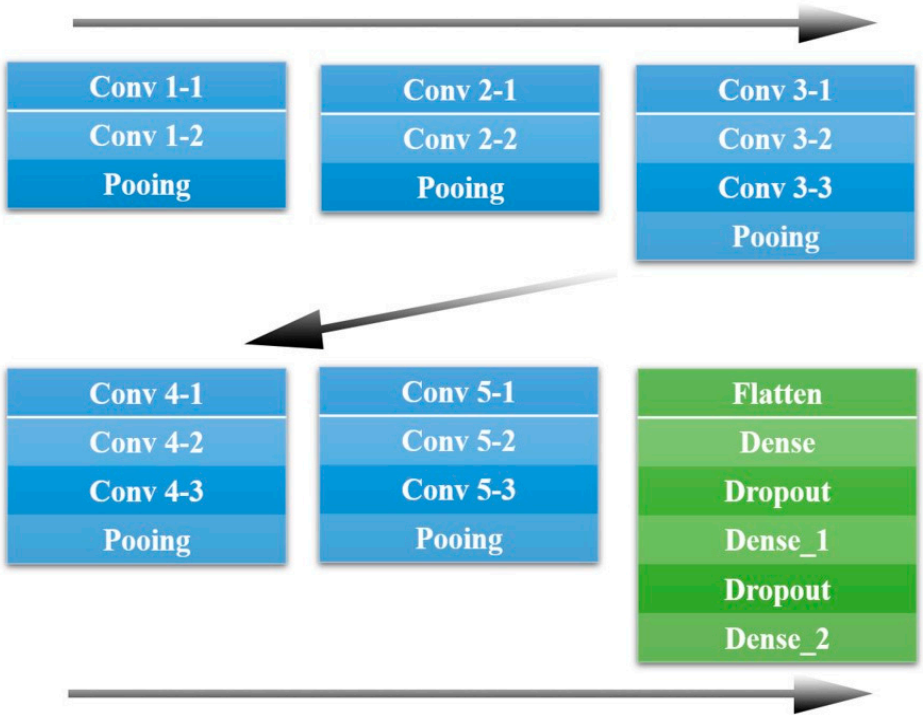


Figure 8. The VGG 16 Architecture.

2.6. Structure of the Proposed VGG16 Model by Tuning Hyperparameter

As shown in Figure 9, in the existing vgg16 model, we use 573 × 505 images rather than 224 × 224 images to perform object recognition learning by reducing the size to 286 × 252, 143 × 126, 71 × 63, 35 × 31, and 17 × 15 when passing through each convolutional layer, and the dence layer was also trained by adjusting the number of neurons to 2048, 1024, and 2 instead of 4096, 4096, and 2. A total of six cases were studied and tested, including two cases of batch size 8, 16 and three cases of epoch 50, 100, and 200 as hyperparameters of learning.

Layer (type)	Output Shape	Param #
input_2 (InputLayer)	[(None, 573, 505, 3)]	0
block1_conv1 (Conv2D)	(None, 573, 505, 64)	1792
block1_conv2 (Conv2D)	(None, 573, 505, 64)	36928
block1_pool (MaxPooling2D)	(None, 286, 252, 64)	0
block2_conv1 (Conv2D)	(None, 286, 252, 128)	73856
block2_conv2 (Conv2D)	(None, 286, 252, 128)	147584
block2_pool (MaxPooling2D)	(None, 143, 126, 128)	0
block3_conv1 (Conv2D)	(None, 143, 126, 256)	295168
block3_conv2 (Conv2D)	(None, 143, 126, 256)	590080
block3_conv3 (Conv2D)	(None, 143, 126, 256)	590080
block3_pool (MaxPooling2D)	(None, 71, 63, 256)	0
block4_conv1 (Conv2D)	(None, 71, 63, 512)	1180160
block4_conv2 (Conv2D)	(None, 71, 63, 512)	2359808
block4_conv3 (Conv2D)	(None, 71, 63, 512)	2359808
block4_pool (MaxPooling2D)	(None, 35, 31, 512)	0
block5_conv1 (Conv2D)	(None, 35, 31, 512)	2359808
block5_conv2 (Conv2D)	(None, 35, 31, 512)	2359808
block5_conv3 (Conv2D)	(None, 35, 31, 512)	2359808
block5_pool (MaxPooling2D)	(None, 17, 15, 512)	0
global_max_pooling2d (GlobalMaxPooling2D)	(None, 512)	0
flatten (Flatten)	(None, 512)	0
dense (Dense)	(None, 2048)	1050624
dropout (Dropout)	(None, 2048)	0
dense_1 (Dense)	(None, 1024)	2098176
dropout_1 (Dropout)	(None, 1024)	0
dense_2 (Dense)	(None, 2)	2050

Total params: 17,865,538
Trainable params: 3,150,850
Non-trainable params: 14,714,688

Figure 9. The structure of the VGG 16 model.

3. Results

In this paper, the VGG16 model for ABR data classification was used to classify hearing loss patients and people with normal hearing. To train the VGG16 model on the collected data, 4500 normal ABR images and 4500 hearing loss ABR images were used. To test

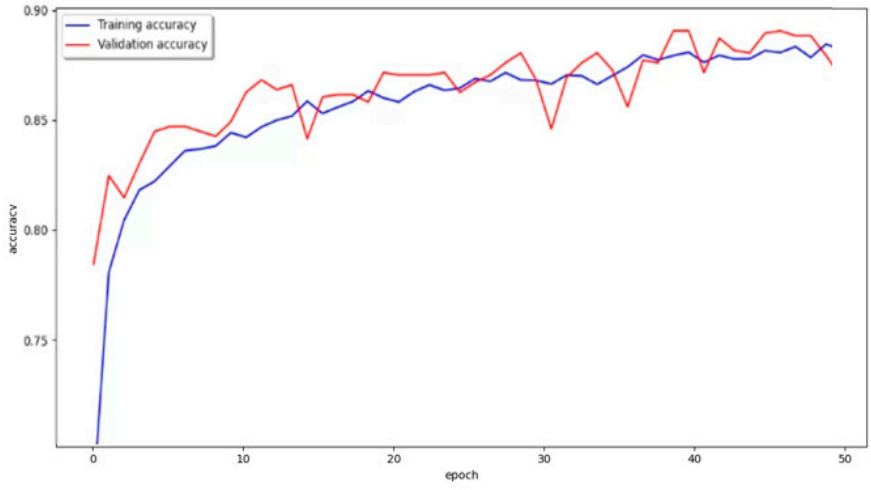
the VGG16 model, 500 normal ABR images and 500 hearing loss ABR images were used. Table 2 shows the specifications of the deep learning machine that conducted learning and classification work in this study.

Table 2. Hardware specifications.

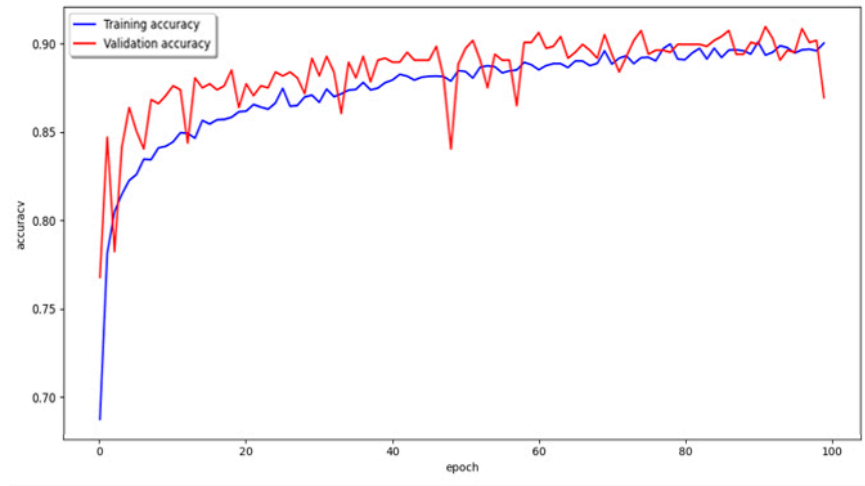
Part	Specification
OS	Windows 10 Pro
CPU	Intel Core i7−12700 2.10 GHz
GPU	NVIDIA GeForce RTX 4090 24 GB
RAM	Samsung 21400—32.0 GB*2
SSD	Samsung 970 1 TB

VGG16 Model Learning and Classification Results

To accurately evaluate the performance of the proposed VGG16 model, we tested and compared it under several conditions. In our experimental test of the proposed VGG16 model, the epochs were increased to (50, 100, and 200) for batch sizes (16 and 8). Therefore, six results were derived, and Figures 10 and 11 and Table 3 show the performance results for the hearing loss patients classification after each learning.

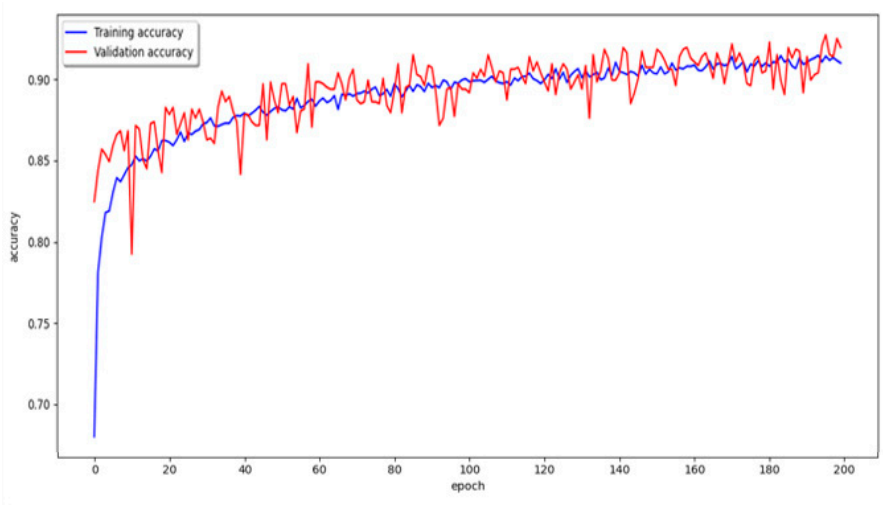


(**a**)



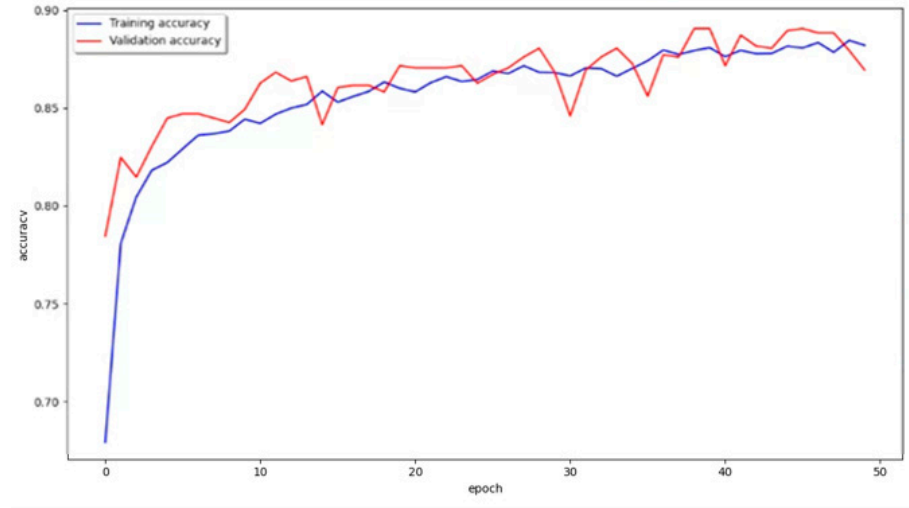
(**b**)

Figure 10. *Cont.*

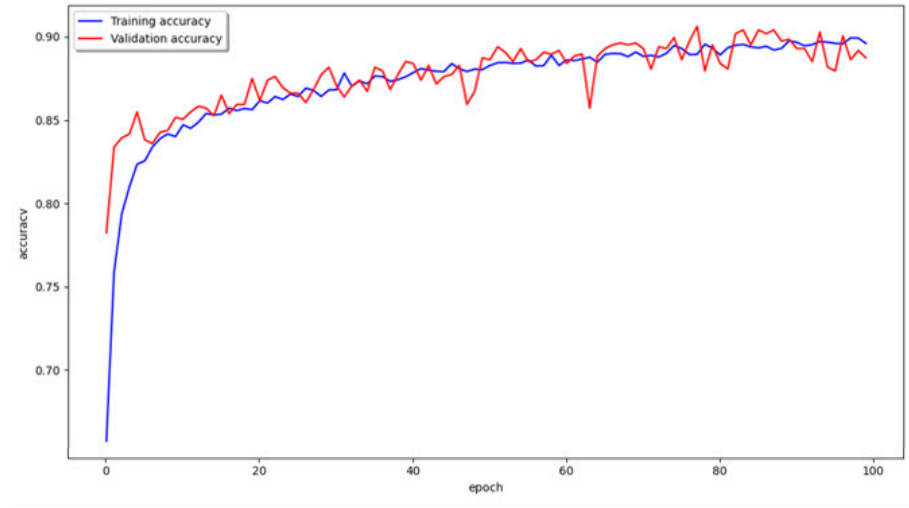


(**c**)

Figure 10. Graph of training and validation accuracy. (**a**) epoch 50, batch size 16; (**b**) epoch 100, batch size 16; (**c**) epoch 200, batch size 16. Blue line is training accuracy. Red line is validation accuracy.

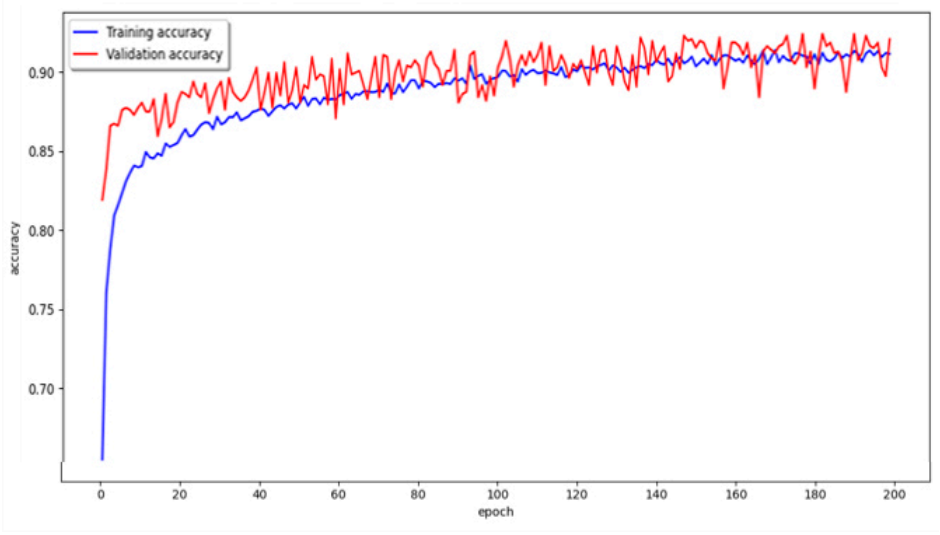


(**a**)



(**b**)

Figure 11. *Cont.*



(**c**)

Figure 11. Graph of training and validation accuracy. (**a**) epoch 50, batch size 8; (**b**) epoch 100, batch size 8; (**c**) epoch 200, batch size 8. Blue line is training accuracy. Red line is validation accuracy.

Table 3. Training accuracy of each test.

Epoch/Batch Size	**50/16**	**100/16**	**200/16**
Training accuracy	87.52%	88.20%	92.13%
Epoch/Batch Size	**50/8**	**100/8**	**200/8**
Training accuracy	88.12%	89.40%	91.96%

4. Discussion

Performance Evaluation of the VGG16 Model

Tables 4 and 5 show the performance evaluation of the proposed VGG16 model. It shows the test result of 500 normal ABR images and 500 hearing loss ABR images analyzed by calculating the accuracy, specificity (TNR: true negative rate), sensitivity (TPR: true positive rate), FPR (false positive rate), FNR (false negative rate), precision, and F1 score. Prediction results for the hearing loss patient data consisted of four types: (1) True positive (TP) is when a normal patient is classified as normal. (2) False positive (FP) is when a patient with hearing loss is classified as normal. (3) True negative (TN) is a case in which a patient with hearing loss is classified as hearing loss. (4) False negative (FN) refers to cases of normal hearing loss patients classified as hearing loss patients. Equations (1)−(7) show the calculation equations for each performance evaluation.

Table 4. Test result of batch size 16.

Epoch	**50**	**100**	**200**
Confusion matrix (tn;fp;fn;tp)	386;114;68;432	402;98;89;411	384;116;43;457
Test accuracy	81.80%	81.30%	84.10%
Specificity	77.20%	80.40%	76.80%
Sensitivity	86.40%	82.20%	91.40%
FPR	22.80%	19.60%	23.20%
FNR	13.60%	17.80%	8.60%
Precision	79.12%	80.75%	79.76%
F1 score	82.60%	81.47%	85.18%

Table 5. Test result of batch size 8.

Epoch	50	100	200
Confusion matrix (tn;fp;fn;tp)	405;95;106;394	430;70;95;405	398;102;49;451
Test accuracy	79.90%	83.50%	84.90%
Specificity	81.00%	86.00%	79.60%
Sensitivity	78.80%	81.00%	90.20%
FPR	19.00%	14.00%	20.40%
FNR	21.20%	19.00%	9.80%
Precision	80.57%	85.26%	81.56%
F1 score	79.68%	83.08%	85.66%

Accuracy indicates the rate at which the proposed model accurately predicted whether the subject was a person with normal hearing (40 dB $\leq$ hearing threshold) or a patient with hearing loss (40 dB > hearing threshold). Sensitivity represents the rate at which people with normal hearing were properly classified as normal. The false positive rate (FPR) refers to the rate of actual hearing loss patients predicted to be people with normal hearing, while the false negative rate (FNR) refers to the rate of those who were predicted to be hearing loss patients, even though they were people with normal hearing. Precision represents the ratio of cases classified as people with normal hearing to actual people with normal hearing. The F1 score represents the harmonic average of precision and sensitivity. The higher the F1 score value, the more valid the value is. The primary purpose of this paper was to determine whether patients applying for disability-grade screening are people with normal hearing or hearing loss patients. Therefore, high sensitivity, precision, and low FNR, which do not classify people with normal hearing as hearing loss patients as much as possible, are judged as valid values while judging the accuracy value that reliably distinguishes people with normal hearing from hearing loss patients. In addition, the F1 score must also be high to verify the reliability of the corresponding value.

$$\text{Accuracy} = \frac{\text{tp} + \text{tn}}{\text{tp} + \text{tn} + \text{fp} + \text{fn}} \tag{1}$$

$$\text{Specificity} = \frac{\text{tn}}{\text{tn} + \text{fp}} \tag{2}$$

$$\text{Sensitivity} = \frac{\text{tp}}{\text{tp} + \text{fn}} \tag{3}$$

$$\text{FPR} = \frac{\text{fp}}{\text{fp} + \text{tn}} \tag{4}$$

$$\text{FNR} = \frac{\text{fn}}{\text{fn} + \text{tp}} \tag{5}$$

$$\text{Precision} = \frac{\text{tp}}{\text{tp} + \text{fp}} \tag{6}$$

$$\text{F1 score} = 2 * \frac{\text{precision} \times \text{sensitivity}}{\text{precision} + \text{sensitivity}} \tag{7}$$

The experimental results showed that when the epoch was set to 200 and the batch size was set to 8, it showed the best performance with accuracy results of 84.90% and an F1 score of 85.66%. The batch size refers to the number of data belonging to a small group when learning training image data once. A smaller number requires longer data learning time, but the results of the model used in this paper showed slightly better learning results.

Epoch refers to the number of times the entire data passes through the neural network, and it simply refers to the number of learning times. Greater number epochs also require longer learning time, but a better learning rate is achieved when learning is carried out up to a certain upper limit.

To prove the effectiveness of the preprocessing performed in this paper, we conducted classification learning and testing for hearing loss patients and people with normal hearing using the VGG16 model. In addition to the best value of epoch 200, epochs 50 and 100 were also sufficiently effective for classification. Therefore, we confirmed that the preprocessing work proposed in this paper is effective in building a hearing loss patient classification model using ABR image data, and in future research, we plan to continue automating the preprocessing work. It is expected that this work will help to distinguish ABR data-based hearing loss between measurement technicians and otolaryngologists.

5. Conclusions

This paper introduced a preprocessing process to automatically classify patients with hearing loss using ABR graph image data using a deep learning model. A result of testing for automatic classification with the VGG16 model after preprocessing the ABR input image of this paper shows accuracy results of 84.90%, specificity of 79.60%, sensitivity of 90.20%, FPR of 20.40%, FNR of 9.80%, precision 81.56%, F1 score 85.66% under epoch 200, and batch size 8 were obtained. Among the 6 learning experimental tests, the FNR is low while the harmonic average F1 score is high, which explains that the rate of normal data classified as hearing loss patient data is low, while the reliability of the learning results is somewhat high. This is expected to help doctors conduct a second review after the first discrimination using the model when classifying the disabled with ABR graph data.

Currently, ABR data is measured by an audiologist using a specific ABR device, and then an otolaryngologist diagnoses hearing loss using these data. This study aims to analyze the ABR result automatically using AI technology for the primary classification of hearing loss, and then to proceed with the otolaryngologist's secondary discrimination. This technique is expected to be of great help in improving the hearing loss diagnosis and analysis process. In a future study, we will present the automation of the ABR data preprocessing process and use the proposed model as new research. In another future study, we will apply an algorithm that automatically detects V waveforms to research a deep learning model to help audiologists evaluate ABR data.

Author Contributions: Conceptualization, M.H. and S.J.C.; methodology, J.M. and M.H.; software, J.M.; validation, J.-H.S., I.J.M., M.K.P., J.B.L., H.K., J.H.J., J.D.L., J.H.A. and S.J.C.; formal analysis, J.M.; investigation, J.M.; resources, J.-H.S., I.J.M., M.K.P., J.B.L., H.K., J.H.J., J.D.L., J.H.A. and S.J.C.; data curation, J.-H.S., I.J.M., M.K.P., J.B.L., H.K., J.H.J., J.D.L., J.H.A. and S.J.C.; writing original draft preparation, J.M.; writing review and editing, M.H. and S.J.C.; visualization, J.M.; supervision, M.H.; project administration, S.J.C.; funding acquisition, M.H. and S.J.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the BK21 FOUR (Fostering Outstanding Universities for Research) Grant No. 5199990914048 and was supported by the Soonchunhyang University Research Fund.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Institutional Review Board of Soonchunhyang University Cheonan Hospital (Cheonan, Korea) (IRB number: 2021-06-040, approval date: 4 June 2021).

Informed Consent Statement: Because of the retrospective design of the study, patient consent was waived.

Data Availability Statement: No new data were created in this study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Burkard, R.F.; Eggermont, J.J.; Don, M. *Auditory Evoked Potentials: Basic Principles and Clinical Application*; Lippincott Williams & Wilkins: Philadelphia, PA, USA, 2007.
2. Aldè, M.; Binda, S.; Primache, V.; Pellegrinelli, L.; Pariani, E.; Pregliasco, F.; Di Berardino, F.; Cantarella, G.; Ambrosetti, U. Congenital cytomegalovirus and hearing loss: The state of the art. *J. Clin. Med.* **2023**, *12*, 4465. [CrossRef] [PubMed]
3. Eggermont, J.J. Auditory brainstem response. In *Handbook of Clinical Neurology*; North-Holland Publishing Company: Amsterdam, The Netherlands, 2019; Volume 160, pp. 451–464.
4. Hood, L.J. Principles and applications in auditory evoked potentials. *Ear Hear.* **1996**, *17*, 178. [CrossRef]
5. Abdala, C.; Folsom, R.C. The development of frequency resolution in humans as revealed by the auditory brain-stem response recorded with notched-noise masking. *J. Acoust. Soc. Am.* **1995**, *98*, 921–930. [CrossRef]
6. Gorga, M.P.; Johnson, T.A.; Kaminski, J.R.; Beauchaine, K.L.; Garner, C.A.; Neely, S.T. Using a combination of click-and tone burst–evoked auditory brain stem response measurements to estimate pure-tone thresholds. *Ear Hear.* **2006**, *27*, 60–74. [CrossRef] [PubMed]
7. Dowley, A.C.; Whitehouse, W.P.; Mason, S.M.; Cope, Y.; Grant, J.; Gibbin, K.P. Auditory neuropathy: Unexpectedly common in a screened newborn population. *Dev. Med. Child Neurol.* **2009**, *51*, 642–646. [CrossRef] [PubMed]
8. Mason, S.; McCormick, B.; Wood, S. Auditory brainstem response in paediatric audiology. *Arch. Dis. Child.* **1988**, *63*, 465. [CrossRef]
9. Elberling, C.; Parbo, J. Reference data for ABRs in retrocochlear diagnosis. *Scand. Audiol.* **1987**, *16*, 49–55. [CrossRef]
10. Lim, T.; Oh, J.; Joo, J.B.; Cho, J.E.; Park, P.; Kim, J.Y. Difference in Thresholds between Auditory Brainstem Response Test and Pure Tone Audiometry by Frequency. *Korean J. Otorhinolaryngol.—Head Neck Surg.* **2023**, *66*, 7–14. [CrossRef]
11. Norton, S.J.; Gorga, M.P.; Widen, J.E.; Folsom, R.C.; Sininger, Y.; Cone-Wesson, B.; Vohr, B.R.; Fletcher, K.A. Identification of neonatal hearing impairment: Summary and recommendations. *Ear Hear.* **2000**, *21*, 529–535. [CrossRef]
12. Kim, L.S.; Jung, S.W.; Heo, S.D.; Hwang, C.H. Comparison of Auditory Brainstem Response and Auditory Steady-State Response Thresholds in Infants and Young Children. *Korean J. Otorhinolaryngol.—Head Neck Surg.* **2004**, *47*, 304–310.
13. Stapells, D.R.; Gravel, J.S.; Martin, B.A. Thresholds for auditory brain stem responses to tones in notched noise from infants and young children with normal hearing or sensorineural hearing loss. *Ear Hear.* **1995**, *16*, 361–371. [CrossRef] [PubMed]
14. Shen, D.; Wu, G.; Suk, H.-I. Deep learning in medical image analysis. *Annu. Rev. Biomed. Eng.* **2017**, *19*, 221–248. [CrossRef] [PubMed]
15. Litjens, G.; Kooi, T.; Bejnordi, B.E.; Setio, A.A.A.; Ciompi, F.; Ghafoorian, M.; van der Laak, J.A.W.M.; van Ginneken, B.; Sánchez, C.I. A survey on deep learning in medical image analysis. *Med. Image Anal.* **2017**, *42*, 60–88. [CrossRef] [PubMed]
16. Yang, C.; Lu, G. Skin Lesion Segmentation with Codec Structure Based Upper and Lower Layer Feature Fusion Mechanism. *KSII Trans. Internet Inf. Syst. (TIIS)* **2022**, *16*, 60–79.
17. Jia, X.; Qian, C.; Yang, Z.; Xu, H.; Han, X.; Ren, H.; Wu, X.; Ma, B.; Yang, D.; Min, H. Boundary-aware dual attention guided liver segment segmentation model. *KSII Trans. Internet Inf. Syst. (TIIS)* **2022**, *16*, 16–37.
18. Razzak, M.I.; Naz, S.; Zaib, A. Deep learning for medical image processing: Overview, challenges and the future. In *Classification in BioApps: Automation of Decision Making*; Springer: Cham, Switzerland, 2018; pp. 323–350.
19. Suzuki, K. Overview of deep learning in medical imaging. *Radiol. Phys. Technol.* **2017**, *10*, 257–273. [CrossRef]
20. Katz, J. *Handbook of Clinical Audiology*, International ed.; Williams & Wikins: Philadelphia, PA, USA, 2014.
21. Hinton, G.E.; Salakhutdinov, R.R. Reducing the dimensionality of data with neural networks. *Science* **2006**, *313*, 504–507. [CrossRef]
22. Hinton, G.E.; Srivastava, N.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R.R. Improving neural networks by preventing co-adaptation of feature detectors. *arXiv* **2012**, arXiv:1207.0580.
23. Tajbakhsh, N.; Shin, J.Y.; Gurudu, S.R.; Hurst, R.T.; Kendall, C.B.; Gotway, M.B.; Liang, J. Convolutional neural networks for medical image analysis: Full training or fine tuning? *IEEE Trans. Med. Imaging* **2016**, *35*, 1299–1312. [CrossRef]
24. Chen, J.; Wang, Y.; Wu, Y.; Cai, C. An ensemble of convolutional neural networks for image classification based on LSTM. In Proceedings of the 2017 International Conference on Green Informatics (ICGI), Fuzhou, China, 15–17 August 2017; pp. 217–222.
25. Lan, Y.; Xie, L.; Cai, X.; Wang, L. A many-objective evolutionary algorithm based on integrated strategy for skin cancer detection. *KSII Trans. Internet Inf. Syst. (TIIS)* **2022**, *16*, 80–96.
26. Saini, R.; Semwal, P.; Jaware, T.H. Brain Tumor Classification Using VGG-16 and MobileNetV2 Deep Learning Techniques on Magnetic Resonance Images (MRI). In *International Conference on Soft Computing and its Engineering Applications*; Springer Nature: Cham, Switzerland, 2022; pp. 300–313.
27. Koo, J.H.; Hwang, L.M.; Kim, H.H.; Kim, T.H.; Kim, J.H.; Song, H.S. Machine learning-based nutrient classification recommendation algorithm and nutrient suitability assessment questionnaire. *KSII Trans. Internet Inf. Syst.* **2023**, *17*, 16–29.
28. Cone-Wesson, B.; Ramirez, G.M. Hearing sensitivity in newborns estimated from ABRs to bone-conducted sounds. *J. Am. Acad. Audiol.* **1997**, *8*, 5.
29. Sininger, Y.S. Auditory brain stem response for objective measures of hearing. *Ear Hear.* **1993**, *14*, 23–30. [CrossRef] [PubMed]
30. Goh, E.K. Diagnosis of hearing loss. *J. Clin. Otolaryngol. Head Neck Surg.* **2003**, *14*, 161–167. [CrossRef]
31. Cueva, R.A. Auditory brainstem response versus magnetic resonance imaging for the evaluation of asymmetric sensorineural hearing loss. *Laryngoscope* **2004**, *114*, 1686–1692. [CrossRef]

32. Verhulst, S.; Jagadeesh, A.; Mauermann, M.; Ernst, F. Individual differences in auditory brainstem response wave characteristics: Relations to different aspects of peripheral hearing loss. *Trends Hear.* **2016**, *20*, 2331216516672186. [CrossRef]
33. Mohan, R.; Ganapathy, K.; Rama, A. Brain tumour classification of magnetic resonance images using a novel CNN based medical image analysis and detection network in comparison with VGG16. *J. Popul. Ther. Clin. Pharmacol.* **2021**, *28*, 2.
34. Qu, X.; Lu, H.; Tang, W.; Wang, S.; Zheng, D.; Hou, Y.; Jiang, J. A VGG attention vision transformer network for benign and malignant classification of breast ultrasound images. *Med. Phys.* **2022**, *49*, 5787–5798. [CrossRef]
35. Al-Radaideh AA, H.; bin Mohd Rahim, M.S.; Ghaban, W.; Bsoul, M.; Kamal, S.; Abbas, N. Arabic Words Extraction and Character Recognition from Picturesque Image Macros with Enhanced VGG-16 Based Model Functionality Using Neural Networks. *KSII Trans. Internet Inf. Syst.* **2023**, *17*, 7.
36. Yang, L.; He, H.; Liu, T. Road Damage Detection and Classification based on Multi-Level Feature Pyramids. *KSII Trans. Internet Inf. Syst.* **2021**, *15*, 2.
37. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
38. Zhang, X.; Zou, J.; He, K.; Sun, J. Accelerating very deep convolutional networks for classification and detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **2015**, *38*, 1943–1955. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.